

---

# JMIR Public Health and Surveillance

---

Impact Factor (2023): 3.5  
Volume 8 (2022), Issue 2 ISSN 2369-2960 Editor in Chief: Travis Sanchez, PhD, MPH

---

## Contents

### Review

- Effectiveness of Mobile Phone and Web-Based Interventions for Diabetes and Obesity Among African American and Hispanic Adults in the United States: Systematic Review ([e25890](#))  
Chineme Enyioha, Matthew Hall, Christiane Voisin, Daniel Jonas. . . . . 3

### Original Papers

- Vitamin K Insufficiency in the Indian Population: Pilot Observational Epidemiology Study ([e31941](#))  
Rama Vaidya, Ashok Vaidya, Jayesh Sheth, Shashank Jadhav, Umakant Mahale, Dilip Mehta, Janusz Popko, Vladimir Badmaev, Sidney Stohs. . . . . 16
- The Use of Cremation Data for Timely Mortality Surveillance During the COVID-19 Pandemic in Ontario, Canada: Validation Study ([e32426](#))  
Gemma Postill, Regan Murray, Andrew Wilton, Richard Wells, Renee Sirbu, Mark Daley, Laura Rosella. . . . . 23
- Use of a Smartphone Self-assessment App for a Tobacco-Induced Disease (COPD, Cardiovascular Diseases, Cancer) Screening Strategy and to Encourage Smoking Cessation: Observational Study ([e19877](#))  
Edouard Stavaux, François Goupil, Guillaume Barreau, Anne Septans, Bertrand Dautzenberg, Armelle Foulet-Rogé, Norbert Padilla, Thierry Urban, Fabrice Denis. . . . . 35
- Public Reactions to the New York State Policy on Flavored Electronic Cigarettes on Twitter: Observational Study ([e25216](#))  
Li Sun, Xinyi Lu, Zidian Xie, Dongmei Li. . . . . 44
- America's HIV Epidemic Analysis Dashboard: Protocol for a Data Resource to Support Ending the HIV Epidemic in the United States ([e33522](#))  
Patrick Sullivan, Cory Woodyatt, Oskian Kouzouian, Kristen Parrish, Jennifer Taussig, Chris Conlan, Harold Phillips. . . . . 55
- Diagnostic Accuracy of an At-Home, Rapid Self-test for Influenza: Prospective Comparative Accuracy Study ([e28268](#))  
Rachel Geyer, Jack Kotnik, Victoria Lyon, Elisabeth Brandstetter, Monica Zigman Suchsland, Peter Han, Chelsey Graham, Misja Ilcisin, Ashley Kim, Helen Chu, Deborah Nickerson, Lea Starita, Trevor Bedford, Barry Lutz, Matthew Thompson. . . . . 84
- An Evaluation of the Text Illness Monitoring (TIM) Platform for COVID-19: Cross-sectional Online Survey of Public Health Users ([e32680](#))  
Heather Joseph, Susan Ingber, Chelsea Austin, Caroline Westnedge, F Strona, Leslie Lee, Ami Shah, Lauren Roper, Anita Patel. . . . . 99

---

COVID-19 Assessment and Testing in Rural Communities During the Pandemic: Cross-sectional Analysis ( <a href="#">e30063</a> ) Jonathan Fitzsimon, Oliver Gervais, Chelsea Lanos. ....	113
COVID-19 Surveillance Updates in US Metropolitan Areas: Dynamic Panel Data Modeling ( <a href="#">e28737</a> ) Theresa Oehmke, Charles Moss, James Oehmke. ....	121
Determining the Case Fatality Rate of COVID-19 in Italy: Novel Epidemiological Study ( <a href="#">e32638</a> ) Mengqing Yan, Wenjun Kang, Zhifeng Guo, Qi Wang, Peizhong Wang, Yun Zhu, Yongli Yang, Wei Wang. ....	131

## Open Source/Open Data

Physical Activity, Sedentary Behavior, and Sleep on Twitter: Multicountry and Fully Labeled Public Data Set for Digital Public Health Surveillance Research ( <a href="#">e32355</a> ) Zahra Shakeri Hossein Abad, Gregory Butler, Wendy Thompson, Joon Lee. ....	67
---	----

Review

# Effectiveness of Mobile Phone and Web-Based Interventions for Diabetes and Obesity Among African American and Hispanic Adults in the United States: Systematic Review

Chineme Enyioha<sup>1</sup>, MD, MPH; Matthew Hall<sup>1</sup>, MD, MPH; Christiane Voisin<sup>2</sup>, MSLS; Daniel Jonas<sup>3</sup>, MD, MPH

<sup>1</sup>Department of Family Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

<sup>2</sup>Cecil G Sheps Center for Health Services Research, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

<sup>3</sup>Department of Internal Medicine, The Ohio State University, Columbus, OH, United States

**Corresponding Author:**

Chineme Enyioha, MD, MPH

Department of Family Medicine

University of North Carolina at Chapel Hill

590 Manning Drive

Chapel Hill, NC, 27514

United States

Phone: 1 9849744888

Email: [chineme\\_enyioha@med.unc.edu](mailto:chineme_enyioha@med.unc.edu)

## Abstract

**Background:** Mobile health (mHealth) and web-based technological advances allow for new approaches to deliver behavioral interventions for chronic diseases such as obesity and diabetes. African American and Hispanic adults experience a disproportionate burden of major chronic diseases.

**Objective:** This paper reviews the evidence for mHealth and web-based interventions for diabetes and obesity in African American and Hispanic adults.

**Methods:** Literature searches of PubMed/Medline, The Cochrane Library, EMBASE, CINAHL Plus, Global Health, Scopus, and Library & Information Science Source were conducted for relevant English-language articles. Articles identified through searches were reviewed by 2 investigators and, if they met the inclusion criteria, were extracted and assessed for risk of bias. Findings were summarized in tabular and narrative format. The overall strength of the evidence was assessed as high, moderate, low, or insufficient on the basis of risk of bias, consistency of findings, directness, precision, and other limitations.

**Results:** Searches yielded 2358 electronic publications, 196 reports were found to be eligible for inclusion, and 7 studies met the eligibility criteria. All 7 included studies were randomized control trials. Five studies evaluated the effectiveness of an mHealth intervention for weight loss, including one that evaluated the effectiveness for diabetes and two studies focused on diabetes. Of all the studies that focused on weight loss, 3 reported significant differences in weight loss in participants in the intervention group compared with those in the usual care group. Although all studies on diabetes control showed greater improvement in glycemic control for the intervention group compared to that in the control group, only one study showed a significant difference between the 2 groups.

**Conclusions:** This analysis indicates that there are few published studies that assessed mHealth interventions among minority populations and focused on weight or diabetes. Although the overall strength of evidence was low for diabetes control, it was moderate for weight loss, and our findings suggest that mHealth and web-based interventions may provide a promising approach for interventions among African American and Hispanic adults who have obesity or diabetes.

(*JMIR Public Health Surveill* 2022;8(2):e25890) doi:[10.2196/25890](https://doi.org/10.2196/25890)

**KEYWORDS**

mHealth; mobile health; technology; diabetes; obesity; African American; Hispanic

## Introduction

Mobile health (mHealth) and web-based technological advances allow for new approaches to delivering behavioral interventions for chronic diseases such as obesity and diabetes [1,2]. These are tailored interventions that supplement or augment in-person contact or office visits through activities such as text messaging, the use of downloadable apps with real time feedback, and other mobile technologies [3,4]. mHealth and web-based interventions are easily accessible because a large proportion of Americans have access to a mobile phone and the internet [5,6].

African American and Hispanic adults experience a disproportionate burden of major chronic diseases, have a high prevalence of diabetes and obesity, and end up with worse health outcomes compared to their White counterparts [7]. For instance, African American adults are at a higher risk of obesity than any other racial or ethnic groups, and they are more likely to have uncontrolled diabetes, develop diabetic retinopathy and end stage renal disease, and have a higher diabetes-related mortality rate than non-Hispanic White adults [8-10]. The management of these health conditions can be challenging for patients since a high level of health literacy and numeracy is required for the self-administration of insulin doses, medications, and the maintenance of a healthy diet. This is compounded with limited access to care or very brief clinic visits with limited time for engagement to fully discuss and understand the disease process or treatment plan [11].

mHealth and web-based technology can mitigate some of these access issues and treatment challenges by providing pertinent information for patients with regard to their condition and offering elements such as monitoring of blood glucose levels and insulin dose adjustment, carbohydrate and calorie counting, and physical activities, which can be incorporated into their routine to improve outcomes [11]. Several studies on mHealth and web-based technology have shown promise in the management of diabetes and obesity [12-14]. A number of systematic reviews have suggested significant benefits of mobile or web-based interventions that target physical activity, diet, and diabetes [15-17]. However, these reviews did not specifically focus on underrepresented racial and ethnic minorities such as African American or Hispanic individuals [18-20]. One systematic review that examined participation of African American adults in mHealth interventions focused on recruitment and retention strategies [21]. Another systematic review that examined the effect of health information technology-based diabetes self-management education interventions on medically underserved patients with diabetes included a broader range of technology including computer software with no internet and telemedicine [22].

Furthermore, minority populations such as African American adults are known to be underrepresented in mHealth research despite existing health disparities [21]. mHealth and web-based technology interventions have the potential to reach patients who may have significant barriers to accessing health care the traditional way. A large proportion of underserved minority populations have ownership of a smartphone and access the internet on their mobile device on a daily basis [23]. African

American and Hispanic adults, compared to non-Hispanic White adults, have been shown to use mobile phones and text messages for a wider range of functions such as purchasing products, using social network sites, and web-based banking [24], which suggests that they may be more receptive to mHealth interventions, thus bridging the health care access divide.

While a growing body of literature supports the ease and effectiveness of mHealth and web-based interventions for diabetes and obesity, to our knowledge, no systematic review has evaluated the effectiveness of such interventions for improving health outcomes of African American and Hispanic patients. The objective of this study is to provide a review of the evidence for mHealth and web-based interventions for diabetes and obesity in African American and Hispanic patients.

## Methods

### Search Strategy and Data Sources

Our literature search was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [25]. We searched for published literature in electronic literature databases PubMed/Medline, The Cochrane Library, EMBASE, CINAHL Plus, Web of Science, Global Health, Scopus, and Library & Information Science Source for unpublished studies, using ClinicalTrials.gov.

A research librarian performed searches for English-language articles on human adults in scientific journals (aged  $\geq 18$  years), which combined terms for (1) minority population terms, (2) mHealth intervention terms, and (3) chronic condition terms for type 2 diabetes mellitus and obesity. The last search was conducted on September 23, 2020. Searches were limited by study design to controlled trials. [Multimedia Appendix 1](#) provides full details regarding the search terms. Studies were included if they met the eligibility criteria ([Table 1](#)). Selected studies were either randomized controlled trials or nonrandomized controlled trials, and study participants had to be either African American or Hispanic individuals in the United States. Studies were also included if at least 30% of participants were African American adults, Hispanic adults, or a mixture of both minority groups. The intervention was any mobile or web-based health intervention including cellular phone calls, text messaging (SMS or MMS), and web-based applications or downloadable mobile apps that targeted management for diabetes and obesity.

Studies with any of the aforementioned interventions in combination with other types of activities such as keeping a journal of meals or extra group or one-on-one coaching were also included. Studies that focused on patients with diabetes or overweight with other health conditions as a group such as stroke, pregnancy, or depression were excluded from this study. There was no limit on the publication time, duration of the intervention, or participants' age. For this review, primary outcomes of interest were objective measures related to obesity and diabetes, including BMI, weight change, waist circumference, or hemoglobin A<sub>1c</sub>, which is a standard measure of glycemic control in patients with diabetes.

**Table 1.** Eligibility criteria.

Category	Include	Exclude
Population	Studies focused on African American or Hispanic adults and studies in which African American or Hispanic adults comprise $\geq 30\%$ of the study population. All age groups.	Studies in which $< 30\%$ of the population comprises African American or Hispanic adults and special populations such as pregnant women.
Intervention	Controlled trials with any mobile phone and web-based intervention including text messages (SMS or MMS), downloadable apps, use of other hand-held devices, or the internet. Duration of the intervention, frequency of contact, time of the day, expected response or action from participants with each contact such as note or log recording or reply to text message.	N/A <sup>a</sup>
Comparators	Usual care including face-to-face coaching, handouts, and no intervention	Comparative effectiveness studies or studies in which both the intervention and control groups had any form of mobile phone or web-based intervention. Studies in which treatment and control groups differed by other interventions besides the intervention delivered by mobile phones or the web (for instance, the intervention group also attends a group class and control group receives phone calls)
Outcome(s)	Hemoglobin A <sub>1c</sub> , BMI, weight, and waist circumference.	N/A
Timing	No limit	N/A
Setting	Studies performed in the United States	Studies performed in other countries
Study Designs	Randomized and nonrandomized controlled trials (includes pilot studies)	Cohort studies, case-control studies, case series, and meta-analyses

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

### Study Selection and Data Extraction

Two reviewers (CE and MH) extracted the data, and each reviewer independently extracting data from all studies. Extracted elements included the following: publication date, authors, study aims and objectives, study design, number of participants in the intervention and control arms, components and duration of the intervention, and results of the study. Outcome measures were extracted at all timepoints for studies that included multiple assessments. Reviewers checked each other's extractions for accuracy and completeness. Discrepancies were discussed until arrival at an agreement and when necessary, a third reviewer was involved.

### Risk of Bias Assessment, Data Synthesis, and Analysis

Each included study was assessed for risk of bias, using the Cochrane tool for risk of bias in randomized trials [26]. The following domains were assessed for risk of bias: selection bias (including the method of randomization and allocation concealment), detection bias, attrition bias, and reporting bias. Studies were rated as low risk, high risk, or unclear risk. Unclear

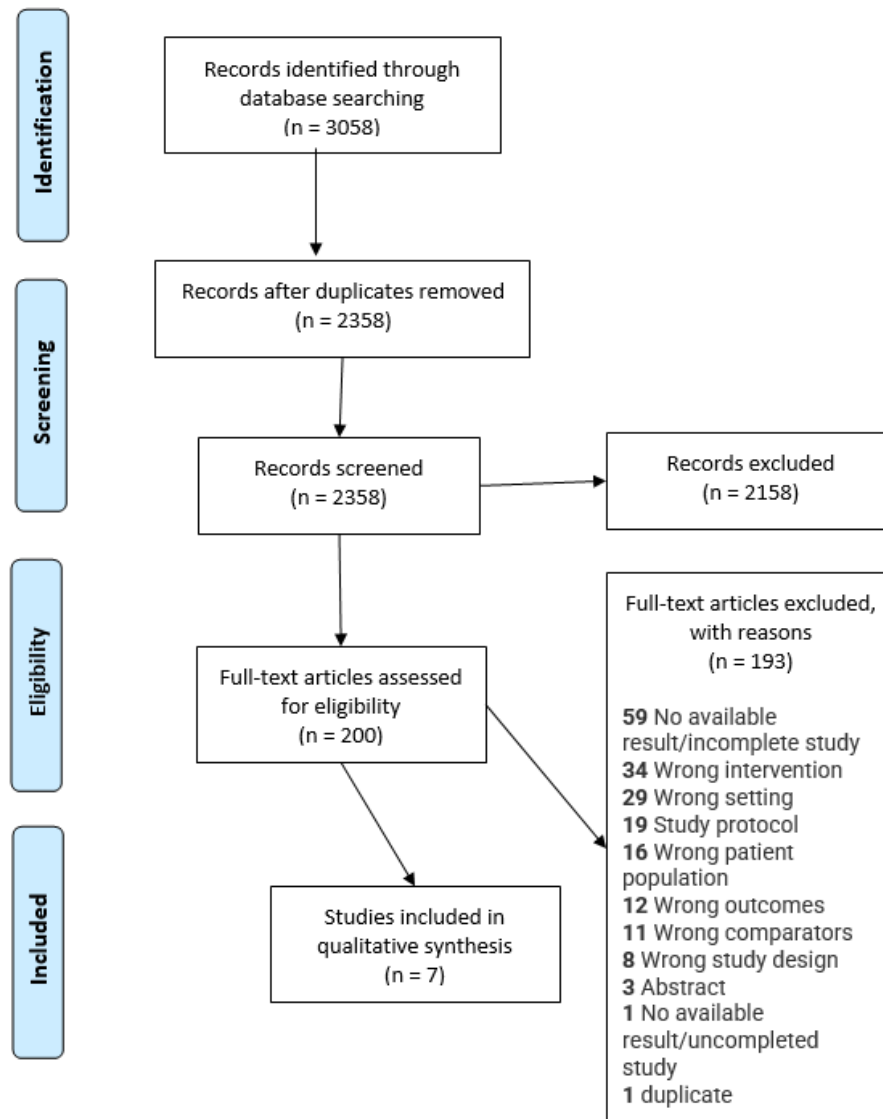
risk of bias was assigned to a study if there was uncertainty or lack of information.

We summarized our findings for each question in tabular and narrative format. We did not conduct a meta-analysis owing to the limited number of studies that met the inclusion criteria. The overall strength of the evidence (SOE) was assessed as high, moderate, low, or insufficient on the basis of the overall risk of bias of included studies, the consistency of findings across studies, directness, precision, and other limitations [27].

## Results

### Search Strategy and Data Sources

The combined search strategies yielded 2358 electronic citations, which were screened to assess for eligibility (Figure 1). In total, 200 studies were found to be potentially eligible, and their full texts were obtained for further assessment. Of these, 7 studies met the eligibility criteria. The top 4 reasons for rejecting other studies were primarily unavailability of results or incompleteness of studies, incorrect patient population, incorrect intervention, and incorrect setting.

**Figure 1.** PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart for study selection.

### Study Selection and Data Extraction

The primary aim of 5 studies was to evaluate the effectiveness of an mHealth intervention for weight loss or diabetes, while 2 studies included an evaluation of the feasibility and acceptability of the mHealth intervention as their primary aim. All included studies were randomized control trials [1,2,12,28-31]. Included studies focused on diabetes control (2 trials) [12,28], obesity and weight loss (4 trials) [2,29-31], or both diabetes and obesity (1 trial) [1].

The total number of participants in all studies was 942, with sample sizes ranging from 18 to 371. Three studies had all African American or Hispanic participants, and the other 3 studies had greater than 30% of participants who identified as

African American or Hispanic minorities. Study durations ranged from 3 months to 12 months. Mean participant ages across studies ranged from 24 to 53 years. BMI of the participants ranged from 31.5 to 38.0. Mean hemoglobin A<sub>1c</sub> values for the diabetes studies ranged from 9.02% to 10.2%.

All studies had mobile phones and SMS text messaging as the main mHealth device and medium of communication (Table 2). SMS text messages were sent to participants anywhere from 2-3 times a day to 3-4 times a week. Mobile technology was applied in several ways including medication reminders, prompts for blood glucose check, provision of motivational messages, education on healthy eating patterns, eating cues, and links to other resources and social support.

**Table 2.** Details of intervention and control activities for the included studies.

Study	Study aim	Intervention	Control
Agboola, 2016 [12]	To evaluate the effectiveness on sending daily physical activity–focused text messages in patients with diabetes compared with no text message on physical activity. Furthermore, an evaluation of the effectiveness of the intervention on hemoglobin A <sub>1c</sub> level, weight change, physical activity, engagement, usability, and satisfaction with the intervention	Participants received a minimum of two automated text messages per day in addition to pedometers—one message in the morning and the other in the evening for 6 months. Messages provided coaching that was dependent on pedometer-captured step counts and physical activity goals target set during the initial visit. Messages received in the morning gave feedback based on activities of the previous day, and messages on other times of the day focused on different coaching themes. Some messages were interactive and focused on elements such as food intake, health status, physical activity, and satisfaction with the program	Pedometers
Arora, 2014 [28]	To evaluate an mHealth intervention for resource-poor emergency department patients with diabetes	Patients received 2 SMS text messages delivered at 9 AM and 5 PM to their mobile phones daily for 6 months. Messages of four categories were sent: one educational/motivational message per day, 3 medication reminders per week, 2 healthy living challenges per week, and 2 trivia messages per week and sent out in question form with the answer sent out an hour later	Usual care (details not provided)
Fortmann, 2017 [1]	To investigate the glycemic effectiveness of a culturally tailored SMS-based diabetes self-management education and support intervention (Dulce Digital)	Participants received information on how to receive and send SMS text messages. Those without a mobile phone were given one. Content of text messages were derived from culturally appropriate Diabetes Self-Management Education and Support curriculum. Participants received 2-3 messages a day initially, which was tapered over 6 months. They received ongoing motivational messages, medication reminders, and prompts for blood glucose measurement	Usual care, which includes visits with a primary care physician, certified diabetes educator, and group diabetes self-management education, dependent on patient or provider initiative
Herring, 2014 [29]	To examine the feasibility, acceptability, and initial efficacy of a technology-based weight loss intervention for urban low-income mothers	Six behavioral health strategies were implemented, one at a time for 2-4 weeks. Participants set realistic goals for each strategy and received 15 minutes biweekly calls from a health counselor. They also receive 3-4 text messages weekly, which probed into their adherence with set goals. Participants also received membership to a Facebook group, which provided access to social support and videos/websites for additional resources.	Regular postpartum care, which is typically one visit with their primary care provider or with a provider through the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC). This visit is usually between 6 and 8 weeks post partum and involves counseling on lactation, birth control, and depression screening. Participants in the study also received counseling on nutrition and vouchers for food and beverages through WIC.
Lin, 2015 [2]	To investigate a behavioral theory-based mobile health intervention to enhance weight loss in patients with obesity	Automated SMS text message program tailored to participants' selection of 3 relevant goals out of 8 options. Messages were customized to participants' wake, lunch, and sleep times	Initial assessment including 20-minute individual sessions with a dietician, health status review with a study physician, and receiving educational material on diet and activity. They also received a digital pedometer
Phelan, 2017 [30]	To evaluate the effect of an internet-based weight loss program in addition to the WIC program on weight loss for low-income postpartum women	Internet-based weight loss program with setting of caloric and physical activity goals. Provision of weekly lessons, web diary, weight and physical activity tracker, and instructional and inspirational videos. Participants received 4 SMS text messages per week with notification of new website content and provision of motivation, support, and feedback. This was in addition to all elements of the WIC program	Participants received all aspects of the standard WIC program and a newsletter every 2 months with information on exercise, nutrition, and wellness

Study	Study aim	Intervention	Control
Steinberg, 2013 [31]	To evaluate the feasibility of daily text messages for self-monitoring behavioral goals for weight loss among African American women with obesity	Shape plan-tracking of tailored behavior change goals through SMS text messaging. Daily feedback through SMS text messages and weekly feedback by email. At 3 months, participants received skills training information including healthy eating patterns and eating cues to reduce face-to-face contact. At 6 months, a 1-hour face-to-face session that focused on problem-solving, progress assessment, and behavior change	Participants received a health education lesson at the start of the study and at 6 months. They also received a set of videos covering topics on healthy eating and exercise at 3 months, in addition to a pedometer and a prescription to walk 10,000 steps a day

mHealth technology was also used to check on adherence to or progress on set physical activity and dietary goals.

In addition to SMS text messages, 3 studies that focused on obesity had additional activities as part of the intervention. One study included all elements of the Women, Infants, and Children (WIC) program [30], another included 15 minutes biweekly calls from health counselors [29], while the other had a face-to-face 1-hour session at the end of the study, which focused on problem-solving, progress assessment, and behavior change [31].

Obesity trials reported mean weight loss and weight change from baseline, percentage weight loss, and change in BMI. One study reported the proportion of participants who achieved a  $\geq 5\%$  weight loss and a  $\geq 10\%$  weight loss from baseline. The diabetes trial reported hemoglobin A<sub>1c</sub>, weight, and BMI of participants at 3 months and 6 months.

### Risk of Bias Assessment and Data Synthesis and Analysis

With regard to the risk of bias of studies included, a randomization sequence generation process was reported by all studies. Bias for allocation concealment and selective outcome

reporting were determined to be low for all studies. Only two studies provided information on blinding of personnel and outcome assessors [2,30]. One trial was determined to have a high risk of bias for incomplete data (up to 63.7% missing at 3 months and 41.1% at 6 months after the start of the intervention) [2,29] and another study was unable to assess differences in provider contact time between intervention and control groups [29].

Overall, 3 studies reported significant differences in weight loss between participants in the intervention group and those in the usual care group [2,29,30] (Table 3). Improvement in weight was noted at 3 months, 14 weeks, 6 months, and 12 months. Two studies found no significant difference in weight loss between groups [1,31].

For the studies that focused on diabetes, all reported an improvement in glycemic control in participants in intervention groups compared to those in control groups [1,12,28], but only one recorded a significant improvement at 6 months [1].

Overall, some evidence supports the efficacy of mHealth and web-based interventions for weight loss among obese African American and Hispanic adults (moderate SOE; Table 4).



**Table 3.** Outcomes of the included studies.

Study participants and sample sizes	Outcome	Baseline data	Result	P value
<b>Agboola, 2016 [12]</b>				
126 English- or Spanish-speaking adults with type 2 diabetes and a hemoglobin A <sub>1c</sub> value of >7.0%	Hemoglobin A <sub>1c</sub> value at 6 months (SD)	<ul style="list-style-type: none"> <li>Control: mean hemoglobin A<sub>1c</sub> 8.38% (SD 1.37%)</li> <li>Intervention: mean hemoglobin A<sub>1c</sub> 9.02% (SD 1.63%)</li> </ul>	<ul style="list-style-type: none"> <li>Control: mean hemoglobin A<sub>1c</sub> 8.17% (SD 1.6%)</li> <li>Intervention: mean hemoglobin A<sub>1c</sub> 8.59% (SD 1.6%)</li> </ul>	.14
<b>Arora, 2014 [28]</b>				
128 adults with poorly controlled diabetes	Median change in hemoglobin A <sub>1c</sub> at 6 months (95% CI)	<ul style="list-style-type: none"> <li>Control: mean hemoglobin A<sub>1c</sub> 10.0% (SD 1.7%)</li> <li>Intervention: mean hemoglobin A<sub>1c</sub> 10.2% (SD 1.7%)</li> </ul>	<ul style="list-style-type: none"> <li>Control: median change in hemoglobin A<sub>1c</sub> -0.60 (95% CI -6.8 to 2.11)</li> <li>Intervention: median change in hemoglobin A<sub>1c</sub> -1.05 (95% CI -5.9 to 2.8)</li> </ul>	.23
<b>Fortmann, 2017 [1]</b>				
126 Hispanic adults with poor glycemic control	<ul style="list-style-type: none"> <li>Mean hemoglobin A<sub>1c</sub> at 3 months and 6 months (SD)</li> <li>Mean weight (lbs) and BMI at 3 months and 6 months (SD)</li> <li>Mean BMI (kg/m<sup>2</sup>) at 3 months and 6 months (SD)</li> </ul>	<ul style="list-style-type: none"> <li>Control: mean hemoglobin A<sub>1c</sub> 9.6% (SD 1.4%)</li> <li>Intervention: mean hemoglobin A<sub>1c</sub> 9.5% (SD 1.2%)</li> <li>Control: mean weight 176.4 (SD 41.6) lbs</li> <li>Intervention: mean weight 173.1 (SD 34.6) lbs</li> <li>Control: mean BMI 32.2 (SD 6.6)</li> <li>Intervention: mean BMI 31.5 (SD 6.0)</li> </ul>	<ul style="list-style-type: none"> <li>3 months <ul style="list-style-type: none"> <li>Control: mean hemoglobin A<sub>1c</sub> 9.3% (SD 1.9%)</li> <li>Intervention: mean hemoglobin A<sub>1c</sub> 8.5% (SD 1.2%)</li> </ul> </li> <li>6 months <ul style="list-style-type: none"> <li>Control: mean hemoglobin A<sub>1c</sub> 9.4% (SD 2.0%)</li> <li>Intervention: mean hemoglobin A<sub>1c</sub> 8.5% (SD 1.2%)</li> </ul> </li> <li>3 months <ul style="list-style-type: none"> <li>Control: mean weight 174.2 (SD 39.7) lbs</li> <li>Intervention: mean weight 176.2 (SD 33.0) lbs</li> </ul> </li> <li>6 months <ul style="list-style-type: none"> <li>Control: mean weight 175.2 (SD 41.6) lbs</li> <li>Intervention: mean weight 174.1 (SD 27.8) lbs</li> </ul> </li> <li>3 months <ul style="list-style-type: none"> <li>Control: mean BMI 32 (SD 6.1)</li> <li>Intervention: mean BMI 31.7 (SD 5.2)</li> </ul> </li> <li>6 months <ul style="list-style-type: none"> <li>Control: mean BMI 32.1 (SD 6.6)</li> <li>Intervention: mean BMI 31.9 (SD 5.4)</li> </ul> </li> </ul>	.03
<b>Herring 2014 [29]</b>				
18 adult women with singleton infants delivered within the last 2 weeks to 12 months	Mean weight loss (kg) at 14 weeks (SD)	<ul style="list-style-type: none"> <li>No baseline weight data</li> </ul>	<ul style="list-style-type: none"> <li>Control: mean weight loss 0.5 (SD 2.3) kg</li> <li>Intervention: mean weight loss 2.9 (SD 3.6) kg</li> </ul>	.04
<b>Lin, 2015 [2]</b>				

Study participants and sample sizes	Outcome	Baseline data	Result	P value
124 African Americans adults	Mean weight loss (kg) from baseline (95% CI)	<ul style="list-style-type: none"> <li>Control: mean weight loss 101.2 kg (95% CI 95.7 kg to 106.7 kg)</li> <li>Intervention: mean weight loss 101.8 kg (95% CI 96.4 kg to 107.2 kg)</li> </ul>	<ul style="list-style-type: none"> <li>3 months               <ul style="list-style-type: none"> <li>Control: mean weight loss -0.2 kg (95% CI -1.0 kg to 0.7 kg)</li> <li>Intervention: mean weight loss -2.6 kg (95% CI -3.8 kg to -1.5 kg)</li> </ul> </li> <li>6 months               <ul style="list-style-type: none"> <li>Control: mean weight loss -0.2 kg (95% CI -1.4 kg to 1.0 kg)</li> <li>Intervention: mean weight loss -3.7 kg (95% CI -5.3 kg to -2.1 kg)</li> </ul> </li> </ul>	<.001 (3 months); .001 (6 months)
<b>Phela, 2017 [30]</b>				
371 Hispanic adult women monitored 6 weeks to 12 months post partum	Mean weight change (kg) at 6 and 12 months (95% CI)	<ul style="list-style-type: none"> <li>Control: mean weight change 82.4 kg (95% CI 77.9 kg to 87.1 kg)</li> <li>Intervention: mean weight change 82.5 kg (95% CI 77.5 kg to 87.5 kg)</li> </ul>	<ul style="list-style-type: none"> <li>6 months               <ul style="list-style-type: none"> <li>Control: mean weight change -1.0 kg (95% CI -1.8 kg to -0.2 kg)</li> <li>Intervention: mean weight change -3.1 kg (95% CI -4.0 kg to -2.3 kg)</li> </ul> </li> <li>12 months               <ul style="list-style-type: none"> <li>Control: mean weight change -0.9 kg (95% CI -1.7 kg to -0.1 kg)</li> <li>Intervention: mean weight change -3.2 kg (95% CI -4.1 kg to -2.4 kg)</li> </ul> </li> </ul>	<.001
<b>Steinberg, 2013 [31]</b>				
50 African American women with obesity	<ul style="list-style-type: none"> <li>Mean weight change (kg) at 6 months (SD)</li> <li>Mean change in BMI (kg/m<sup>2</sup>) (SD)</li> </ul>	<ul style="list-style-type: none"> <li>Control: mean weight change 96.0 (SD 23.1) kg</li> <li>Intervention: mean weight change 102.0 (SD 16.6) kg</li> <li>Control: mean BMI change 34.6 (SD 5.8)</li> <li>Intervention: mean BMI change 36.9 (SD 6.2)</li> </ul>	<ul style="list-style-type: none"> <li>Control: mean weight change 1.14 (SD 2.53) kg</li> <li>Intervention: mean weight change -1.27 (SD 6.51) kg</li> <li>Control: mean BMI change 0.42 (SD 0.90)</li> <li>Intervention: mean BMI change -0.47 (SD 2.42)</li> </ul>	.09

**Table 4.** Overall strength of the evidence.

Outcome	Studies (observations), n	Summary of findings	Consistency, directness, and precision	Limitations (including reporting bias)	Overall strength of evidence	Applicability
Weight	5 [1,2,29-31] (689)	Increased weight loss noted in all intervention groups but significant in only 3 studies. One high-quality study [30] (n=371) reported a significant difference in weight change between participants in the intervention and those in the control group	Consistent, indirect, and precise	Two studies had a high risk of bias for nonblinding of study personnel, participants, and outcome assessors. One study had a high risk of bias for incomplete data	Moderate	African American and Hispanic adults with obesity or morbid obesity, who are young and middle-aged adults, and have access to a mobile phone
Hemoglobin A <sub>1c</sub>	3 [1,12,28] (380)	Improvement in hemoglobin A <sub>1c</sub> in the intervention groups but only significant in one study	Consistent, indirect, and imprecise	Studies had a high risk of bias from the nonblinding of study personnel and participants. One study also had a high risk of bias from the nonblinding of outcome assessors and the other study had an unclear risk of bias for the same	Low	African American or Hispanic adults with poorly controlled diabetes and access to a mobile phone

## Discussion

### Principal Findings

We identified 7 trials that reported the benefit of mHealth interventions directed toward obesity and diabetes among African American and Hispanic adults. We found an association between receiving the mHealth intervention and weight loss compared to those in control groups, which suggests that mHealth and web-based interventions are effective for minority patients who are overweight. Of the three studies that addressed diabetes, only one reported an improvement in hemoglobin A<sub>1c</sub> among those who received mHealth interventions compared to their control group counterparts, suggesting that the evidence may be unclear for the role of mHealth interventions in diabetes control among patients in ethnic minorities. This systematic review shows that mHealth and web-based interventions for 6 months to a year can be effective for weight loss in African American and Hispanic adults.

Our findings in support of mHealth interventions for weight loss among patients in ethnic minorities are similar to those of other studies [32-34]. A systematic review by Ryan et al [32] reported that mHealth intervention for weight loss is beneficial, while the systematic review and meta-analysis by Sorgente et al [33] was more specific in suggesting that web-based interventions are more effective than usual care or minimal intervention, but the evidence is conflicting with regard to the benefit of web-based interventions compared to non-web-based interventions. In our systematic review, of the 3 studies that reported significant weight loss in the intervention group, one of them had a control group that received usual care, while control group participants in the other 2 studies had comparable

non-mHealth interventions. Beilegoli et al [34] suggest that web-based interventions only lead to a short-term weight loss but provide no evidence for weight loss over a long period of time. The longest study duration in our review was 12 months. Studies beyond 12 months, which met our inclusion criteria, were not found, and it is unclear if participants will remain in an intervention of this kind for longer than a period of a year. Changes in diet and activity patterns during holiday periods, the challenge or burden of adherence to self-monitoring over a long period of time, and the need for coaching in addition to self-monitoring have been suggested as reasons for the difficulty in maintaining the long-term effects of mHealth interventions for obesity and weight loss [35,36].

We also observed an improvement in hemoglobin A<sub>1c</sub> with mHealth interventions compared to those with usual care. Our findings are similar to those of other studies that have shown improved glycemic control with mHealth interventions in the general population [12,14,36], except that only one study in this review showed a significant difference at 6 months. Another systematic review showed an improvement in 42% of studies in the review [37]. One major finding noted in the diabetes studies in this review was that the study with significant results incorporated culturally appropriate messages as part of their intervention [1]. Even with the improvements noted, it is unclear if outcomes would be significantly better if mHealth and web-based interventions have improved usability. African American and Hispanic adults have been found to be less adherent to interventions with technology [38,39]. Better outcomes have been noted in studies where the intervention was developed to be more appealing to the target minority population [40]. The lack of cultural consistency and application has been noted as a major challenge with mHealth and

web-based interventions, which can influence results and outcomes [37,41]. Furthermore, some studies have shown that diabetes mHealth programs do not always meet health literacy guidelines, and they are not as user-friendly [42]. The design of some mHealth apps makes it challenging for some groups of people. For instance, Hispanic patients with diabetes who use apps are more likely to use certain functionalities such as medication and blood glucose diaries and less likely to use diaries developed for insulin and hemoglobin A<sub>1c</sub> [43].

Another important finding from this study is the limited number of studies with African American and Hispanic adults in mHealth interventions for obesity and diabetes. The underrepresentation of participants from minority groups in research studies has been well-documented [44-46] and is reflective in this review. Several reasons have been cited for this, including barriers related to misconceptions or lack of information, mistrust, concerns about outcomes, stigmatization, insurance coverage, and immigration status in the United States [46-48]. More studies focused on mHealth interventions for African American and Hispanic adults with diabetes and obesity will augment the current evidence on the role of mHealth interventions for chronic disease management. This may also contribute to reducing the disparities gap that currently exist in the health care system.

### Limitations

The limited number of studies with primarily African American and Hispanic participants is a limitation of this study, which has also been observed in several other studies [21]. Another limitation is our study inclusion criteria. This review focused on African American and Hispanic participants and only

included studies with  $\geq 30\%$  of the participants of interest. The team considered a set of less rigid inclusion criteria but decided against it since greater the proportion of African American or Hispanic participants, greater the ability to make conclusions specifically about the effectiveness of mHealth interventions on these racial groups in these studies. While all included studies involved mobile phones and SMS text messages, they differed in the content of the intervention, sample sizes, and study duration. The absence of studies with data beyond 12 months is also a limitation. Obesity and diabetes are chronic conditions that require long-term management and lifestyle modification. Furthermore, although some studies that met our inclusion criteria showed significant changes, it is unclear whether these translate to clinical significance.

### Conclusions

Although the overall strength of the evidence was moderate for weight loss and low for diabetes, our findings suggest that mHealth and web-based interventions may provide a promising approach. mHealth technology can be used to meet different health needs such as the delivery of motivational messages, health education, prompts or reminders, and personal coaching [49]. Further studies are required to fully understand the role of mHealth interventions in chronic disease management in minority patients. These findings are significant because minority patients bear a heavy burden of diabetes and obesity in the United States. mHealth and web-based interventions have the potential to assist medical providers as they provide care to patients with chronic disease and improve health disparities. Further research to better understand how mHealth and web-based interventions can help overcome challenges or barriers to chronic disease management and will be very useful.

---

### Acknowledgments

This work was supported by the UNC Department of Family Medicine and the Cecil G. Sheps Center for Health Disparities, University of North Carolina. We would also like to acknowledge Jennifer Middleton for her assistance with the data analysis.

---

### Conflicts of Interest

None declared.

---

Multimedia Appendix 1

Search terms.

[[PDF File \(Adobe PDF File\), 611 KB - publichealth\\_v8i2e25890\\_app1.pdf](#)]

---

### References

1. Fortmann AL, Gallo L, Garcia M, Taleb M, Euyoque JA, Clark T, et al. Dulce Digital: An mHealth SMS-Based Intervention Improves Glycemic Control in Hispanics With Type 2 Diabetes. *Diabetes Care* 2017 Oct;40(10):1349-1355 [FREE Full text] [doi: [10.2337/dc17-0230](https://doi.org/10.2337/dc17-0230)] [Medline: [28600309](https://pubmed.ncbi.nlm.nih.gov/28600309/)]
2. Lin M, Mahmooth Z, Dedhia N, Frutchey R, Mercado CE, Epstein DH, et al. Tailored, interactive text messages for enhancing weight loss among African American adults: the TRIMM randomized controlled trial. *Am J Med* 2015 Aug;128(8):896-904. [doi: [10.1016/j.amjmed.2015.03.013](https://doi.org/10.1016/j.amjmed.2015.03.013)] [Medline: [25840035](https://pubmed.ncbi.nlm.nih.gov/25840035/)]
3. Ybarra M, Bağci Bosi AT, Korchmaros J, Emri S. A text messaging-based smoking cessation program for adult smokers: randomized controlled trial. *J Med Internet Res* 2012 Dec 27;14(6):e172 [FREE Full text] [doi: [10.2196/jmir.2231](https://doi.org/10.2196/jmir.2231)] [Medline: [23271159](https://pubmed.ncbi.nlm.nih.gov/23271159/)]
4. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. *Cochrane Database Syst Rev* 2016 Apr 10;4:CD006611 [FREE Full text] [doi: [10.1002/14651858.CD006611.pub4](https://doi.org/10.1002/14651858.CD006611.pub4)] [Medline: [27060875](https://pubmed.ncbi.nlm.nih.gov/27060875/)]

5. Shaw R, Bosworth H. Short message service (SMS) text messaging as an intervention medium for weight loss: A literature review. *Health Informatics J* 2012 Dec;18(4):235-250 [[FREE Full text](#)] [doi: [10.1177/1460458212442422](https://doi.org/10.1177/1460458212442422)] [Medline: [23257055](https://pubmed.ncbi.nlm.nih.gov/23257055/)]
6. Stephens J, Allen J. Mobile phone interventions to increase physical activity and reduce weight: a systematic review. *J Cardiovasc Nurs* 2013;28(4):320-329 [[FREE Full text](#)] [doi: [10.1097/JCN.0b013e318250a3e7](https://doi.org/10.1097/JCN.0b013e318250a3e7)] [Medline: [22635061](https://pubmed.ncbi.nlm.nih.gov/22635061/)]
7. Weinstock RS, Teresi J, Goland R, Izquierdo R, Palmas W, Eimicke JP, IDEATel Consortium. Glycemic control and health disparities in older ethnically diverse underserved adults with diabetes: five-year results from the Informatics for Diabetes Education and Telemedicine (IDEATel) study. *Diabetes Care* 2011 Feb;34(2):274-279 [[FREE Full text](#)] [doi: [10.2337/dc10-1346](https://doi.org/10.2337/dc10-1346)] [Medline: [21270184](https://pubmed.ncbi.nlm.nih.gov/21270184/)]
8. Petersen R, Pan L, Blanck HM. Racial and Ethnic Disparities in Adult Obesity in the United States: CDC's Tracking to Inform State and Local Action. *Prev Chronic Dis* 2019 Apr 11;16:E46 [[FREE Full text](#)] [doi: [10.5888/pcd16.180579](https://doi.org/10.5888/pcd16.180579)] [Medline: [30974071](https://pubmed.ncbi.nlm.nih.gov/30974071/)]
9. Golden SH, Brown A, Cauley J, Chin MH, Gary-Webb TL, Kim C, et al. Health disparities in endocrine disorders: biological, clinical, and nonclinical factors--an Endocrine Society scientific statement. *J Clin Endocrinol Metab* 2012 Sep;97(9):E1579-E1639 [[FREE Full text](#)] [doi: [10.1210/jc.2012-2043](https://doi.org/10.1210/jc.2012-2043)] [Medline: [22730516](https://pubmed.ncbi.nlm.nih.gov/22730516/)]
10. Lanting LC, Joung I, Mackenbach J, Lamberts SWJ, Bootsma AH. Ethnic differences in mortality, end-stage complications, and quality of care among diabetic patients: a review. *Diabetes Care* 2005 Sep;28(9):2280-2288. [doi: [10.2337/diacare.28.9.2280](https://doi.org/10.2337/diacare.28.9.2280)] [Medline: [16123507](https://pubmed.ncbi.nlm.nih.gov/16123507/)]
11. Shan R, Sarkar S, Martin SS. Digital health technology and mobile devices for the management of diabetes mellitus: state of the art. *Diabetologia* 2019 Jun;62(6):877-887. [doi: [10.1007/s00125-019-4864-7](https://doi.org/10.1007/s00125-019-4864-7)] [Medline: [30963188](https://pubmed.ncbi.nlm.nih.gov/30963188/)]
12. Agboola S, Jethwani K, Lopez L, Searl M, O'Keefe S, Kvedar J. Text to Move: A Randomized Controlled Trial of a Text-Messaging Program to Improve Physical Activity Behaviors in Patients With Type 2 Diabetes Mellitus. *J Med Internet Res* 2016 Nov 18;18(11):e307 [[FREE Full text](#)] [doi: [10.2196/jmir.6439](https://doi.org/10.2196/jmir.6439)] [Medline: [27864165](https://pubmed.ncbi.nlm.nih.gov/27864165/)]
13. Bennett GG, Steinberg D, Askew S, Levine E, Foley P, Batch BC, et al. Effectiveness of an App and Provider Counseling for Obesity Treatment in Primary Care. *Am J Prev Med* 2018 Dec;55(6):777-786 [[FREE Full text](#)] [doi: [10.1016/j.amepre.2018.07.005](https://doi.org/10.1016/j.amepre.2018.07.005)] [Medline: [30361140](https://pubmed.ncbi.nlm.nih.gov/30361140/)]
14. Stephens JD, Yager AM, Allen J. Smartphone Technology and Text Messaging for Weight Loss in Young Adults: A Randomized Controlled Trial. *J Cardiovasc Nurs* 2017;32(1):39-46 [[FREE Full text](#)] [doi: [10.1097/JCN.0000000000000307](https://doi.org/10.1097/JCN.0000000000000307)] [Medline: [26646593](https://pubmed.ncbi.nlm.nih.gov/26646593/)]
15. Palmer M, Sutherland J, Barnard S, Wynne A, Rezel E, Doel A, et al. The effectiveness of smoking cessation, physical activity/diet and alcohol reduction interventions delivered by mobile phones for the prevention of non-communicable diseases: A systematic review of randomised controlled trials. *PLoS One* 2018;13(1):e0189801 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0189801](https://doi.org/10.1371/journal.pone.0189801)] [Medline: [29304148](https://pubmed.ncbi.nlm.nih.gov/29304148/)]
16. Afshin A, Babalola D, Mclean M, Yu Z, Ma W, Chen C, et al. Information Technology and Lifestyle: A Systematic Evaluation of Internet and Mobile Interventions for Improving Diet, Physical Activity, Obesity, Tobacco, and Alcohol Use. *J Am Heart Assoc* 2016 Aug 31;5(9):e003058 [[FREE Full text](#)] [doi: [10.1161/JAHA.115.003058](https://doi.org/10.1161/JAHA.115.003058)] [Medline: [27581172](https://pubmed.ncbi.nlm.nih.gov/27581172/)]
17. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. *Int J Behav Nutr Phys Act* 2016 Dec 07;13(1):127 [[FREE Full text](#)] [doi: [10.1186/s12966-016-0454-y](https://doi.org/10.1186/s12966-016-0454-y)] [Medline: [27927218](https://pubmed.ncbi.nlm.nih.gov/27927218/)]
18. Isaacs T, Hunt D, Ward D, Rooshenas L, Edwards L. The Inclusion of Ethnic Minority Patients and the Role of Language in Telehealth Trials for Type 2 Diabetes: A Systematic Review. *J Med Internet Res* 2016 Sep 26;18(9):e256 [[FREE Full text](#)] [doi: [10.2196/jmir.6374](https://doi.org/10.2196/jmir.6374)] [Medline: [27670360](https://pubmed.ncbi.nlm.nih.gov/27670360/)]
19. Jackson CL, Bolen S, Brancati FL, Batts-Turner ML, Gary TL. A systematic review of interactive computer-assisted technology in diabetes care. *Interactive information technology in diabetes care. J Gen Intern Med* 2006 Feb;21(2):105-110 [[FREE Full text](#)] [doi: [10.1111/j.1525-1497.2005.00310.x](https://doi.org/10.1111/j.1525-1497.2005.00310.x)] [Medline: [16390512](https://pubmed.ncbi.nlm.nih.gov/16390512/)]
20. Ford JG, Howerton MW, Lai GY, Gary TL, Bolen S, Gibbons MC, et al. Barriers to recruiting underrepresented populations to cancer clinical trials: a systematic review. *Cancer* 2008 Jan 15;112(2):228-242 [[FREE Full text](#)] [doi: [10.1002/cncr.23157](https://doi.org/10.1002/cncr.23157)] [Medline: [18008363](https://pubmed.ncbi.nlm.nih.gov/18008363/)]
21. James DC, Harville C, Sears C, Efunbumi O, Bondoc I. Participation of African Americans in e-Health and m-Health Studies: A Systematic Review. *Telemed J E Health* 2017 May;23(5):351-364. [doi: [10.1089/tmj.2016.0067](https://doi.org/10.1089/tmj.2016.0067)] [Medline: [27792475](https://pubmed.ncbi.nlm.nih.gov/27792475/)]
22. Heitkemper EM, Mamykina L, Travers J, Smaldone A. Do health information technology self-management interventions improve glycemic control in medically underserved adults with diabetes? A systematic review and meta-analysis. *J Am Med Inform Assoc* 2017 Sep 01;24(5):1024-1035 [[FREE Full text](#)] [doi: [10.1093/jamia/ocx025](https://doi.org/10.1093/jamia/ocx025)] [Medline: [28379397](https://pubmed.ncbi.nlm.nih.gov/28379397/)]
23. Stockman M, Modzelewski K, Steenkamp D. Mobile Health and Technology Usage by Patients in the Diabetes, Nutrition, and Weight Management Clinic at an Urban Academic Medical Center. *Diabetes Technol Ther* 2019 Jul;21(7):400-405. [doi: [10.1089/dia.2018.0369](https://doi.org/10.1089/dia.2018.0369)] [Medline: [31045447](https://pubmed.ncbi.nlm.nih.gov/31045447/)]
24. Zickuhr K, Smith A. Digital differences. Pew Research Center. 2012 Apr 13. URL: <https://www.pewresearch.org/internet/2012/04/13/digital-differences/> [accessed 2020-06-24]

25. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015 Jan 01;4:1 [[FREE Full text](#)] [doi: [10.1186/2046-4053-4-1](https://doi.org/10.1186/2046-4053-4-1)] [Medline: [25554246](#)]
26. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011 Oct 18;343:d5928 [[FREE Full text](#)] [doi: [10.1136/bmj.d5928](https://doi.org/10.1136/bmj.d5928)] [Medline: [22008217](#)]
27. Berkman ND, Lohr KN, Ansari M, McDonagh M, Balk E, Whitlock E, et al. Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update. In: *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. Rockville, MD: Agency for Healthcare Research and Quality; 2008.
28. Arora S, Peters AL, Burner E, Lam CN, Menchine M. Trial to examine text message-based mHealth in emergency department patients with diabetes (TEXT-MED): a randomized controlled trial. *Ann Emerg Med* 2014 Jun;63(6):745-54.e6. [doi: [10.1016/j.annemergmed.2013.10.012](https://doi.org/10.1016/j.annemergmed.2013.10.012)] [Medline: [24225332](#)]
29. Herring SJ, Cruice JF, Bennett GG, Davey A, Foster GD. Using technology to promote postpartum weight loss in urban, low-income mothers: a pilot randomized controlled trial. *J Nutr Educ Behav* 2014;46(6):610-615 [[FREE Full text](#)] [doi: [10.1016/j.jneb.2014.06.002](https://doi.org/10.1016/j.jneb.2014.06.002)] [Medline: [25069621](#)]
30. Phelan S, Hagobian T, Brannen A, Hatley KE, Schaffner A, Muñoz-Christian K, et al. Effect of an Internet-Based Program on Weight Loss for Low-Income Postpartum Women: A Randomized Clinical Trial. *JAMA* 2017 Jun 20;317(23):2381-2391 [[FREE Full text](#)] [doi: [10.1001/jama.2017.7119](https://doi.org/10.1001/jama.2017.7119)] [Medline: [28632867](#)]
31. Steinberg DM, Levine EL, Askew S, Foley P, Bennett GG. Daily text messaging for weight control among racial and ethnic minority women: randomized controlled pilot study. *J Med Internet Res* 2013 Nov 18;15(11):e244 [[FREE Full text](#)] [doi: [10.2196/jmir.2844](https://doi.org/10.2196/jmir.2844)] [Medline: [24246427](#)]
32. Ryan K, Dockray S, Linehan C. A systematic review of tailored eHealth interventions for weight loss. *Digit Health* 2019;5:2055207619826685 [[FREE Full text](#)] [doi: [10.1177/2055207619826685](https://doi.org/10.1177/2055207619826685)] [Medline: [30783535](#)]
33. Sorgente A, Pietrabissa G, Manzoni GM, Re F, Simpson S, Perona S, et al. Web-Based Interventions for Weight Loss or Weight Loss Maintenance in Overweight and Obese People: A Systematic Review of Systematic Reviews. *J Med Internet Res* 2017 Jun 26;19(6):e229 [[FREE Full text](#)] [doi: [10.2196/jmir.6972](https://doi.org/10.2196/jmir.6972)] [Medline: [28652225](#)]
34. Beleigoli AM, Andrade AQ, Cançado AG, Paulo MN, Diniz MDFH, Ribeiro AL. Web-Based Digital Health Interventions for Weight Loss and Lifestyle Habit Changes in Overweight and Obese Adults: Systematic Review and Meta-Analysis. *J Med Internet Res* 2019 Jan 08;21(1):e298 [[FREE Full text](#)] [doi: [10.2196/jmir.9609](https://doi.org/10.2196/jmir.9609)] [Medline: [30622090](#)]
35. Burke LE, Styn MA, Sereika SM, Conroy MB, Ye L, Glanz K, et al. Using mHealth technology to enhance self-monitoring for weight loss: a randomized trial. *Am J Prev Med* 2012 Jul;43(1):20-26 [[FREE Full text](#)] [doi: [10.1016/j.amepre.2012.03.016](https://doi.org/10.1016/j.amepre.2012.03.016)] [Medline: [22704741](#)]
36. Pretlow RA, Stock CM, Allison S, Roeger L. Treatment of child/adolescent obesity using the addiction model: a smartphone app pilot study. *Child Obes* 2015 Jun;11(3):248-259 [[FREE Full text](#)] [doi: [10.1089/chi.2014.0124](https://doi.org/10.1089/chi.2014.0124)] [Medline: [25760813](#)]
37. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. *J Med Internet Res* 2015 Feb 24;17(2):e52 [[FREE Full text](#)] [doi: [10.2196/jmir.3951](https://doi.org/10.2196/jmir.3951)] [Medline: [25803266](#)]
38. Trief PM, Izquierdo R, Eimicke JP, Teresi JA, Goland R, Palmas W, et al. Adherence to diabetes self care for white, African-American and Hispanic American telemedicine participants: 5 year results from the IDEATel project. *Ethn Health* 2013;18(1):83-96. [doi: [10.1080/13557858.2012.700915](https://doi.org/10.1080/13557858.2012.700915)] [Medline: [22762449](#)]
39. Wolin KY, Steinberg DM, Lane IB, Askew S, Greaney ML, Colditz GA, et al. Engagement with eHealth Self-Monitoring in a Primary Care-Based Weight Management Intervention. *PLoS One* 2015;10(10):e0140455 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0140455](https://doi.org/10.1371/journal.pone.0140455)] [Medline: [26469065](#)]
40. Turnbull S, Cabral C, Hay A, Lucas PJ. Health Equity in the Effectiveness of Web-Based Health Interventions for the Self-Care of People With Chronic Health Conditions: Systematic Review. *J Med Internet Res* 2020 Jun 05;22(6):e17849 [[FREE Full text](#)] [doi: [10.2196/17849](https://doi.org/10.2196/17849)] [Medline: [32459632](#)]
41. Wu X, Guo X, Zhang Z. The Efficacy of Mobile Phone Apps for Lifestyle Modification in Diabetes: Systematic Review and Meta-Analysis. *JMIR Mhealth Uhealth* 2019 Jan 15;7(1):e12297 [[FREE Full text](#)] [doi: [10.2196/12297](https://doi.org/10.2196/12297)] [Medline: [30664494](#)]
42. Fu H, McMahon SK, Gross CR, Adam TJ, Wyman JF. Usability and clinical efficacy of diabetes mobile applications for adults with type 2 diabetes: A systematic review. *Diabetes Res Clin Pract* 2017 Sep;131:70-81. [doi: [10.1016/j.diabres.2017.06.016](https://doi.org/10.1016/j.diabres.2017.06.016)] [Medline: [28692830](#)]
43. Williams JP, Schroeder D. Popular Glucose Tracking Apps and Use of mHealth by Latinos With Diabetes: Review. *JMIR Mhealth Uhealth* 2015 Aug 25;3(3):e84 [[FREE Full text](#)] [doi: [10.2196/mhealth.3986](https://doi.org/10.2196/mhealth.3986)] [Medline: [26307533](#)]
44. Miranda J, Nakamura R, Bernal G. Including ethnic minorities in mental health intervention research: a practical approach to a long-standing problem. *Cult Med Psychiatry* 2003 Dec;27(4):467-486. [doi: [10.1023/b:medi.0000005484.26741.79](https://doi.org/10.1023/b:medi.0000005484.26741.79)] [Medline: [14727681](#)]

45. Hussain-Gambles M, Atkin K, Leese B. Why ethnic minority groups are under-represented in clinical trials: a review of the literature. *Health Soc Care Community* 2004 Sep;12(5):382-328. [doi: [10.1111/j.1365-2524.2004.00507.x](https://doi.org/10.1111/j.1365-2524.2004.00507.x)] [Medline: [15373816](https://pubmed.ncbi.nlm.nih.gov/15373816/)]
46. George S, Duran N, Norris K. A systematic review of barriers and facilitators to minority research participation among African Americans, Latinos, Asian Americans, and Pacific Islanders. *Am J Public Health* 2014 Feb;104(2):e16-e31. [doi: [10.2105/AJPH.2013.301706](https://doi.org/10.2105/AJPH.2013.301706)] [Medline: [24328648](https://pubmed.ncbi.nlm.nih.gov/24328648/)]
47. BeLue R, Taylor-Richardson K, Lin J, Rivera A, Grandison D. African Americans and participation in clinical trials: differences in beliefs and attitudes by gender. *Contemp Clin Trials* 2006 Dec;27(6):498-505. [doi: [10.1016/j.cct.2006.08.001](https://doi.org/10.1016/j.cct.2006.08.001)] [Medline: [16962382](https://pubmed.ncbi.nlm.nih.gov/16962382/)]
48. Katz RV, Green BL, Kressin NR, Kegeles SS, Wang MQ, James SA, et al. The legacy of the Tuskegee Syphilis Study: assessing its impact on willingness to participate in biomedical studies. *J Health Care Poor Underserved* 2008 Nov;19(4):1168-1180 [FREE Full text] [doi: [10.1353/hpu.0.0067](https://doi.org/10.1353/hpu.0.0067)] [Medline: [19029744](https://pubmed.ncbi.nlm.nih.gov/19029744/)]
49. Lyles C, Schillinger D, Sarkar U. Connecting the Dots: Health Information Technology Expansion and Health Disparities. *PLoS Med* 2015 Jul;12(7):e1001852 [FREE Full text] [doi: [10.1371/journal.pmed.1001852](https://doi.org/10.1371/journal.pmed.1001852)] [Medline: [26172977](https://pubmed.ncbi.nlm.nih.gov/26172977/)]

## Abbreviations

**mHealth:** mobile health

**SOE:** strength of the evidence

**WIC:** Women, Infants, and Children

*Edited by T Sanchez; submitted 19.11.20; peer-reviewed by H Fu, D Schroeder; comments to author 04.01.21; revised version received 05.02.21; accepted 14.10.21; published 04.02.22.*

*Please cite as:*

*Enyioha C, Hall M, Voisin C, Jonas D*

*Effectiveness of Mobile Phone and Web-Based Interventions for Diabetes and Obesity Among African American and Hispanic Adults in the United States: Systematic Review*

*JMIR Public Health Surveill* 2022;8(2):e25890

URL: <https://publichealth.jmir.org/2022/2/e25890>

doi: [10.2196/25890](https://doi.org/10.2196/25890)

PMID: [35119368](https://pubmed.ncbi.nlm.nih.gov/35119368/)

©Chineme Enyioha, Matthew Hall, Christiane Voisin, Daniel Jonas. Originally published in *JMIR Public Health and Surveillance* (<https://publichealth.jmir.org>), 04.02.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

# Vitamin K Insufficiency in the Indian Population: Pilot Observational Epidemiology Study

Rama Vaidya<sup>1</sup>, MD, PhD; Ashok D B Vaidya<sup>2\*</sup>, MD, PhD, FAIM; Jayesh Sheth<sup>3</sup>, PhD; Shashank Jadhav<sup>4</sup>, BHMS, PGDPM; Umakant Mahale<sup>4</sup>, MSc; Dilip Mehta<sup>4\*</sup>, PhD; Janusz Popko<sup>5</sup>, MD, PhD; Vladimir Badmaev<sup>4</sup>, BA, MD, PhD; Sidney J Stohs<sup>6</sup>, BSc, MS, PhD

<sup>1</sup>Division of Endocrine and Metabolic Disorders, Medical and Research Centre, Kasturba Health Society, Mumbai, Maharashtra, India

<sup>2</sup>Medical Research Centre, Kasturba Health Society, Mumbai, India

<sup>3</sup>The Foundation for Research in Genetics and Endocrinology, Ahmedabad, Gujarat, India

<sup>4</sup>Synergia Life Sciences Pvt Ltd, Mumbai, Maharashtra, India

<sup>5</sup>Medical University of Bialystok, Bialystok, Poland

<sup>6</sup>Creighton University Medical Center, Frisco, TX, United States

\* these authors contributed equally

**Corresponding Author:**

Sidney J Stohs, BSc, MS, PhD

Creighton University Medical Center

7068 Maumee Valley Court

Frisco, TX, 75036

United States

Phone: 1 2142156655

Email: [sid.stohs9@gmail.com](mailto:sid.stohs9@gmail.com)

## Abstract

**Background:** The fat-soluble K vitamins K1 and K2 play an essential role in the blood coagulation cascade and are made available predominantly through selective dietary intakes. They are less known for their nonessential roles in a family of vitamin K-dependent proteins that promote various functions of organs and systems in the body. A lack of vitamin K can characterize vitamin and nutritional element insufficiency, which is different from a clinically apparent vitamin deficiency.

**Objective:** This epidemiological study evaluated the nutritional status of vitamin K in a sample of the Indian population and vitamin K content in staple Indian foods.

**Methods:** Serum levels of vitamin K1 and vitamin K2 in the form of menaquinone-7 (MK-7) were assessed via high-performance liquid chromatography coupled with fluorescence detection in 209 patients with type 2 diabetes, 50 healthy volunteers, and common staple foods in India.

**Results:** After comparing populations with high and low serum vitamin K levels from various geographical regions, our results indicated that the sample of healthy Indian individuals and the sample of Indian patients with type 2 diabetes had low (insufficient) levels of vitamin K2 (MK-7; range 0.3-0.4 ng/mL). No significant differences existed in vitamin K1-related and MK-7-related values between healthy male and female subjects, between male and female subjects with diabetes, and between the healthy sample and the sample of patients with diabetes. The staple, commonly consumed Indian foods that were tested in this study had undetectable levels of vitamin K2, while levels of vitamin K1 varied widely (range 0-37 µg/100 g).

**Conclusions:** Based on our sample's low serum levels of vitamin K2 (MK-7) as well as the low levels of vitamin K2 in their typical diet, we propose that the general Indian population could benefit from the consumption of vitamin K2 in the form of MK-7 supplements.

**Trial Registration:** Clinical Trials Registry - India CTRI/2019/05/014246; <http://ctri.nic.in/Clinicaltrials/showallp.php?mid1=21660&EncHid=&userName=014246>; Clinical Trials Registry - India CTRI/2019/03/018278; <http://ctri.nic.in/Clinicaltrials/showallp.php?mid1=32349&EncHid=&userName=018278>

(JMIR Public Health Surveill 2022;8(2):e31941) doi:[10.2196/31941](https://doi.org/10.2196/31941)



**KEYWORDS**

phyloquinone; menaquinone-7; vitamin K1; vitamin K2; insufficiency; deficiency; Indian population; diabetes; healthy people

## Introduction

Clinically apparent nutritional deficiency diseases, such as scurvy (vitamin C deficiency), rickets (vitamin D deficiency), beriberi (vitamin B1 deficiency), and pellagra (vitamin B3 deficiency), differ from nutritional insufficiency—a relatively new finding. The effects of vitamin insufficiency with regard to a number of vitamins are well known. However, the roles of vitamin K, which are exemplified by the insufficiency of this vitamin, are much less well known [1,2].

Recently, it has been found that the mechanism of nutritional insufficiency can be explained by a triage (sort and select) mechanism, and nutritional insufficiency can be exemplified by the nutritional insufficiency of vitamin K [2]. The term *triage* is used on the battlefield to prioritize treatments for the survival of the wounded. In a similar way, the human body prioritizes nutrient use (ie, micronutrients, nutrients, and vitamins) for immediate needs (eg, survival) over nutrient use for long-term needs (eg, sustaining life) by borrowing nutrients from less critical depots in the body, always securing the emergency requirements.

In healthy people who consume a varied diet, the clinical deficiency of vitamin K is rarely encountered because the body prioritizes available vitamins for use with essential, life-sustaining, vitamin K-dependent proteins (eg, the carboxylation of such proteins and the activation of vitamin K-dependent blood coagulation proteins, such as coagulation factors II [thrombin], VII, IX, and X). However, vitamin K may be insufficient for at least 18 vitamin K-dependent, nonessential, calcium-dependent proteins that are responsible for healthy cardiovascular, immune, skeletal, and neuromuscular systems [1,3].

Due to the recent discovery of its many biological functions, vitamin K is known as a “multitasking” vitamin. In the long term, the insufficient status of this multitasking vitamin may prevent vitamin K-dependent nonessential proteins from functioning optimally and result in the development of chronic degenerative conditions, such as osteoporosis, cardiovascular disease, metabolic conditions (eg, diabetes), and neurodegenerative conditions, that ultimately diminish quality of life and shorten the life span [1].

Due to the growing awareness of the importance of vitamin K nutritional status, the objectives of this epidemiological study were to evaluate serum levels of vitamin K in healthy men and women and patients with diabetes selected from the Indian

population and to assess the average vitamin K content in the indigenous diet of India. This study evaluated 2 fat-soluble vitamins—vitamin K1 (phyloquinone), which is commonly found in a diet of green vegetables and plant margarine, and vitamin K2 (menaquinone; specifically, menaquinone-7 [MK-7]), which is derived mainly from meat; liver; butter; egg yolks; fermented foods (eg, cheese and curd); and indigenous Indian dietary products, including dosa, dhokla, and yogurt.

Besides food sources, vitamin K2 is also synthesized in the human gut microbiome predominantly by *Bacteroides* and *Veillonella* bacteria [4]. However, gut bacteria that generate menaquinones reside mostly in the large intestine. As such, gut bacteria-derived menaquinones are less bioavailable, since the absorption of menaquinones occurs predominantly in the small intestine. Therefore, serum levels of menaquinones depend largely on dietary sources [4].

## Methods

### Subject Selection

The Department of Medicine at Kokan Hospital in Mumbai, India, selected a sample of the native population based on the admission criteria for this study, which are provided in [Textbox 1](#).

The Inter System BioMedica Ethics Committee—an independent ethics committee—approved the protocol of this study and registered this study with Clinical Trials Registry - India (trial CTRI/2019/05/014246 for the healthy population and trial CTRI/2019/03/018278 for the population with diabetes). The trial registration information can be found on the Clinical Trials Registry - India website and the World Health Organization portal. The Department of Medicine at Kokan Hospital in Mumbai, India, approved the study protocol according to the International Conference on Harmonization of Good Clinical Practice Guidelines [5].

The sample of patients with diabetes (n=209) was randomly selected from patients who visited the hospital diabetes clinic for treatment, and the healthy subjects (n=50) were screened for their health status and criteria of inclusion upon responding to the advertisement for study participants. Those who met the inclusion criteria were selected for this study. The data obtained were recorded in ethics committee-approved case record forms. The participants agreed with and signed the informed consent form prior to their admission to the study groups, and blood samples were collected following a fasting period of at least 8 hours and after the consumption of a standard meal.

**Textbox 1.** Inclusion and exclusion criteria for the healthy population and the population with diabetes.

#### Inclusion criteria

- Healthy population
  - Healthy males and females aged 28 to 45 years
  - Having a BMI of 18.5-24.9 kg/m<sup>2</sup>
- Population with diabetes
  - Males and females aged ≥25 years
  - A duration of diabetes of ≥6 months from the date of diagnosis
  - Fasting plasma glucose levels of ≥126.0 mg/dL

#### Exclusion criteria

- Healthy population
  - People with any systemic illness
  - People who are on corticosteroids and oral contraceptives
  - People on antibiotics within the last week
  - People with a seropositive status
  - Pregnant and lactating women
  - Participation in clinical trials evaluating investigational pharmaceuticals or biologics within 3 months of admission to our study
  - Participation in clinical trials evaluating devices within 30 days of admission to our study
  - People who are on coumarin analogues or quinine hydrochloride
  - People who have a history of smoking, alcohol abuse, or substance abuse within the last week
- Population with diabetes
  - Subjects with type 1 diabetes
  - Lactating and pregnant mothers
  - Participants with concomitant chronic illness

## Analytical Methods

The serum levels of vitamin K1 and vitamin K2 in the form of MK-7 were assessed via high-performance liquid chromatography (HPLC) coupled with fluorescence detection. The vitamin K1 and vitamin K2 (MK-7) content in food was evaluated with the same HPLC analytical method. A reverse-phase HPLC method [6-8] involving postcolumn derivatization and fluorescent detection was used, with some modifications. This method was validated according to the International Conference on Harmonization guidelines [5]. Our laboratory works with the UK Vitamin K External Quality Assurance Scheme (KEQAS) to ensure quality control.

Vitamin K1 and vitamin K2 (MK-7) analytical standards, a vitamin K2-6 internal standard, and zinc dust (<10µM) for the postcolumn reduction of vitamin K were purchased from Sigma Chemical Co. All of the other chemicals used were analytical reagent grade, and all solvents for HPLC were HPLC grade.

An HPLC system (Shimadzu Corporation) was used. This system was equipped with a degasser (DGU-20A5R; Shimadzu Corporation), pump (LC-20AD; Shimadzu Corporation), auto-sampler (SIL-20AC HT; Shimadzu Corporation), column

oven (CTO-10 AS; Shimadzu Corporation), and fluorescence detector (RF 20A; Shimadzu Corporation). A reverse-phase C18 column (Kinetex C18 [Phenomenex Inc]; size: 100×4.6 mm; volume: 2.6 µL; angle: 100°) was used, and the column oven was adjusted to 25 °C. A postcolumn (30×4 mm) assembly was inserted between the analytical column, which was packed with zinc dust for the postcolumn reduction of vitamin K, and the fluorescence detector. The mobile phase contained 2 mL of zinc solution (0.136% zinc chloride, 0.04% sodium acetate, and 30 µL of acetic acid in 75% methanol) in 2 L of methanol and was filtered through 0.45-µm filter paper. Injected samples or standards were eluted at an isocratic flow rate of 1.2 mL per minute over a 12-minute run time. Eluted peaks were monitored with the fluorescent detector, which was set at an excitation wavelength and emission wavelength of 248 nm and 430 nm, respectively.

Six-point calibration curves of vitamin K were created by spiking plasma samples with 0.16- to 20-ng/mL concentrations of the vitamins. All samples were extracted as follows. To each plasma sample, 500 µL of the internal standard (20-ng/mL vitamin K2-6 in methanol) was added, followed by 3.5 mL of ethanol and 0.5 mL of sodium chloride (0.9% in water). Each

mixture was vortexed for 2 minutes, and 10 mL of hexane was added. Afterward, the tubes were again vortexed. The hexane layers were separated via centrifugation, collected in separate tubes, and evaporated under nitrogen. The residues were dissolved in 100  $\mu$ L of methanol, and 50  $\mu$ L of this solution was injected into the column. For all food samples, about 1000 mg of each test sample was placed in a 100-mL volumetric flask, to which 10 mL of tetrahydrofuran, 1 mL of the internal standard, and 70 mL of ethanol were added. The contents of these flasks were sonicated for 1 hour and diluted to volume with ethanol. Afterward, the samples were filtered, and aliquots of the filtrate were injected into the column.

### Statistical Methods

All data in this study were analyzed by using the 2-tailed unpaired *t* test method. This method was used to determine the statistical significance between 2 sets of data (set at  $P < .05$ ). In addition, the data were analyzed via an analysis of variance to determine their statistical significance between and across all groups. All statistical analyses were conducted by using the built-in functions of Microsoft Excel. Statistical analyses were conducted between healthy male and female subjects, between male and female subjects with diabetes, and between the healthy sample and the sample of patients with diabetes.

### Data Sharing

The data described in this paper, the code book, and the analytic code can be made available upon request.

**Table 1.** Baseline demographic and clinical data of the study population.

Characteristic	Study 1 (volunteer patients with type 2 diabetes mellitus: n=209)		Study 2 (healthy volunteers: n=50)	
	Males	Females	Males	Females
Subjects, n	100	109	25	25
Age (years), mean (SD)	50.3 (13.1)	48.6 (14.0)	41.6 (12.8)	36.0 (5.4)
BMI ( $\text{kg}/\text{m}^2$ ), mean (SD)	25.4 (4.2)	26.9 (5.4)	23.8 (2.3)	22.7 (1.9)
HbA <sub>1c</sub> <sup>a</sup> level (%), mean (SD)	7.6 (1.7) <sup>b</sup>	7.2 (1.7) <sup>b</sup>	4.9 (0.4) <sup>b</sup>	4.7 (0.4) <sup>b</sup>
Systolic blood pressure (mm Hg), mean (SD)	128.6 (15.6)	126.8 (16.7)	114.6 (10.4)	122.6 (12.9)
Diastolic blood pressure (mm Hg), mean (SD)	89.8 (21.5)	86.4 (22.9)	80.0 (4.8)	80.8 (6.8)
Fasting blood sugar level (mg/dL), mean (SD)	123.2 (36.9) <sup>b</sup>	117.2 (47.3) <sup>b</sup>	82.1 (7.7) <sup>b</sup>	82.5 (8.2) <sup>b</sup>
Postprandial blood sugar level (mg/dL), mean (SD)	193.6 (66.0) <sup>b</sup>	179.9 (83.1) <sup>b</sup>	102.5 (7.1) <sup>b</sup>	104.9 (7.0) <sup>b</sup>

<sup>a</sup>HbA<sub>1c</sub>: glycosylated hemoglobin.

<sup>b</sup>Values are significant at the  $P < .05$  level.

### Vitamin K Blood Levels

The levels of vitamin K2 (MK-7) and phylloquinone (vitamin K1) in the healthy population and the population with type 2 diabetes mellitus are presented in Table 2. The differences in the levels of vitamin K1 and MK-7 between male and female subjects with diabetes and between healthy male and female subjects were not statistically significant ( $P < .05$ ) based on the analysis of variance. There was a trend toward higher MK-7 levels in both male and female subjects with diabetes. The

## Results

### Statistical Analysis

Based on a power calculation and the use of a power calculation table [9,10], a sample size of 25 for each group had 90% power to detect a 0.38 difference between the means of the two groups at a significance level ( $\alpha$ ) of  $< .05$  via a 2-tailed *t* test. A sample size of 100 subjects for each group had 90% power to detect a 0.18 difference between the means of the two groups at a significance level of  $< .05$ .

### Demographic and Clinical Data

Baseline demographic and clinical data are presented in Table 1 for male and female subjects in the group of subjects with diabetes and the healthy volunteer comparison group. Table 1 provides the basic demographic details for all subjects with respect to age, gender, and BMI as well as baseline values for systolic and diastolic blood pressure, fasting blood sugar level, postprandial blood sugar level, and glycosylated hemoglobin (HbA<sub>1c</sub>) level in the two study groups based on gender.

Statistically significant differences ( $P < .05$ ) existed between the group of subjects with diabetes and the healthy volunteers with respect to blood sugar levels and HbA<sub>1c</sub> levels, as would be expected (Table 1). There were no significant differences between male and female subjects within the two groups with respect to these markers. Blood pressure values and BMIs tended to be higher among the subjects with diabetes.

vitamin K1 serum levels in the studied healthy population and the population with diabetes were within the normal range, based on the physiological levels of vitamin K1 in their serum samples. The determination of a low vitamin K2 (MK-7) level (range: 0.116-1.056 ng/mL) was based on limited epidemiological data [11,12] from an adult population [13]. However, the serum levels of vitamin K2 in the form of MK-7 in both the healthy sample and the sample of patients with diabetes were determined to be in the low range.

**Table 2.** Levels of menaquinone-7 (MK-7; vitamin K2) and phyloquinone (vitamin K1) in the healthy population and the population with type 2 diabetes mellitus.

Characteristic	Patients with type 2 diabetes mellitus		Healthy volunteers	
	Males	Females	Males	Females
Subjects, n	100	109	25	25
MK-7 level (ng/mL), mean (SD)	0.41 (0.37)	0.42 (0.49)	0.31 (0.23)	0.39 (0.20)
Phylloquinone (vitamin K1) level (ng/mL), mean (SD)	0.55 (0.50)	0.53 (0.46)	0.70 (0.65)	0.48 (0.39)

### Vitamin K Levels in Selected Foods

**Table 3** provides data regarding the analysis of vitamin K2 (MK-7) and vitamin K1 content ( $\mu\text{g}/100\text{ g}$  or  $\mu\text{g}/100\text{ mL}$ ) in staple foods of India. As can be seen in **Table 3**, vitamin K2 (MK-7) was absent from or below the limits of detection in all studied foods, whereas the vitamin K1 content varied widely. Therefore, the dietary intake of these foods would result in wide variations in vitamin K1 serum levels.

It should be noted that the analytical laboratory involved in this study works with the UK KEQAS to ensure quality control. To ensure the accuracy of the analytical procedures for vitamin K, on 6 occasions, the laboratory received 2 serum samples and 1 standard vitamin K sample from KEQAS for the analysis of vitamin K via the HPLC method. On all 6 occasions, the analytical method that was used yielded satisfactory results, and a "Green" certification was awarded to the laboratory. The target for results was a 20% deviation from the all-laboratory trimmed mean representing the target concentrations of the unknown samples that were provided by KEQAS.

**Table 3.** Staple foods in India that were evaluated for menaquinone-7 (MK-7; vitamin K2) and phyloquinone (vitamin K1) content.

Food item <sup>a</sup>	Description	MK-7 content, mean (SD)	Phylloquinone (vitamin K1) content, mean (SD)
Dhokla	Fermented batter derived from rice and split chickpeas	— <sup>b</sup>	2.76 (0.17) <sup>c</sup>
Naan	Bread leavened with yeast or with a bread starter (flatbread)	—	—
Jalebi	A sweet snack made with deep-fried wheat flour	—	—
Idli	A savory cake made of fermented lentils and rice	—	1.67 (2.13) <sup>d</sup>
Handvo	A cake based on gram flour with vegetables and peanuts	—	37.14 (3.02) <sup>d</sup>
Yogurt	A dairy product made by coagulating milk with any culinary acid (eg, lemon juice)	—	—
Cheese	Processed cheddar cheese	—	4.71 (0.68) <sup>d</sup>
Buttermilk	A fermented dairy drink made by churning the butter out of cultured cream	—	0.51 (0.05) <sup>c</sup>
Butter	Made of pure milk fat	—	4.60 (0.82) <sup>d</sup>
Milk	Cow milk	—	0.97 (0.19) <sup>c</sup>
Charoli	Almond-flavored seeds of a bush ( <i>Buchanania ianzan</i> )	—	2.85 (0.40) <sup>d</sup>

<sup>a</sup>Four samples from each food item were analyzed.

<sup>b</sup>Not available (not detected).

<sup>c</sup> $\mu\text{g}/100\text{ mL}$

<sup>d</sup> $\mu\text{g}/100\text{ g}$

## Discussion

### Vitamin K Blood Levels in India

The primary outcome of this study was the determination that vitamin K2 insufficiency occurred in both a healthy cohort and a cohort of people with diabetes from India. A secondary outcome was that vitamin K2 was not restricted to individuals with a disease such as diabetes; it was also detected in a group of apparently healthy individuals.

Although the typical nutritional intake in India may provide sufficient vitamin K for supporting the life-sustaining blood coagulation cascade, the vitamin K content of staple Indian foods and the intake of vitamin K from these foods may be insufficient for this vitamin to fulfill its multitasking role in preventing disease and sustaining health in a growing number of people with vitamin K-dependent health conditions. As such, the general consumption of vitamin K2 in the form of MK-7 supplements may be justified.

## Vitamin K Blood Levels Throughout the World

Studies have indicated that serum levels of vitamin K vary throughout the world and are largely dependent upon dietary intake [11]. For example, the serum levels of vitamin K2 (MK-7) in Japan have been reported to differ by region, varying from 5.26 (SD 6.13) ng/mL among women in eastern Japan (eg, Kanto region, Tokyo) to 1.22 (SD 1.85) ng/mL among women in western Japan (eg, Kasai region, Hiroshima) [11]. Further, the mean vitamin K2 (MK-7) level in British women has been reported to be 0.37 (SD 0.20) ng/mL [11].

The recommendations for the daily intake of vitamin K have been inconsistent, in part due to the unquantified vitamin K contribution of intestinal bacteria. For example, the recommended levels of vitamin K that may be adequate for blood clotting are insufficient for other functions [12]. A daily intake of 1 µg of vitamin K per 1 kg of body weight may be adequate for blood clotting [12]. In a Western diet, the intake of vitamin K is estimated to range from 60 µg to 200 µg per day, of which phylloquinone (vitamin K1) constitutes about 90% and menaquinones (vitamin K2) constitute about 10% [14].

Findings for vitamin K2 (MK-7) serum levels indicate wide variations in dietary vitamin K2 content resulting from different dietary staples with varying vitamin K2 contents [11]. For example, fermented beans known as *natto* are consumed regularly at breakfast in eastern Japan, and approximately 1000 µg of MK-7 is provided per 100 g of *natto*. However, according to an epidemiological study, consuming *natto* is an infrequent dietary practice in the western part of Japan [11]. This provides a possible explanation for the occurrence of higher vitamin K2 (MK-7) serum levels in women from eastern Japan than those in women from western Japan and may explain the lower incidence of cardiovascular and skeletal morbidity in the former group [11].

## Vitamin K and COVID-19

One of the recently discovered health conditions that has been linked with vitamin K insufficiency is SARS-CoV-2 infection, which results in the viral disease known as *COVID-19* [13]. The scientists behind this epidemiological finding proposed that insufficient serum levels of vitamin K could result in the enhancement of the inflammatory response associated with COVID-19, which contributes to multi-organ failure in patients with COVID-19. This hypothesis is supported by another epidemiological study on vitamin K, in which patients with COVID-19 exhibited reduced vitamin K status and had poor prognoses [11]. Further studies are needed to explore and affirm the possible role of vitamin K in susceptibility to and recovery from COVID-19 as well as other viral diseases, such as influenza.

## Conclusions

The results showed that the serum levels of vitamin K2 in the form of MK-7 among both the healthy sample and the sample with diabetes from the population that was studied were in the low range. As such, the results of this study suggest but do not prove that vitamin K insufficiency may be a common occurrence. The results provide a basis and template for future studies and may also serve as a guide for health care practitioners when it comes to vitamin K2 supplementation. Based on the results of this study, which indicated low serum levels of vitamin K2 (MK-7) as well as low vitamin K2 levels in the typical Indian diet, we propose that the general Indian population could benefit from the consumption of vitamin K2 in the form of MK-7 supplements. Patients with diabetes, elevated blood pressure (a harbinger of cardiovascular disease), and compromised immune systems may especially benefit from MK-7 supplementation. Although this study was conducted in India, the results may be extrapolated to other countries worldwide.

---

## Acknowledgments

This study was funded by Synergia Life Sciences Pvt Ltd, Mumbai, India.

---

## Conflicts of Interest

DM, UM, and SJ are associated with Synergia Life Sciences. RV, ADBV, JS, and VB serve on an advisory board of Synergia Life Sciences Pvt Ltd. SJS and JP have no potential conflicts to disclose.

---

## References

1. Karpiński M, Popko J, Maresz K, Badmaev V, Stohs SJ. Roles of vitamins D and K, nutrition, and lifestyle in low-energy bone fractures in children and young adults. *J Am Coll Nutr* 2017 Jul;36(5):399-412. [doi: [10.1080/07315724.2017.1307791](https://doi.org/10.1080/07315724.2017.1307791)] [Medline: [28686548](https://pubmed.ncbi.nlm.nih.gov/28686548/)]
2. McCann JC, Ames BN. Vitamin K, an example of triage theory: is micronutrient inadequacy linked to diseases of aging? *Am J Clin Nutr* 2009 Oct;90(4):889-907 [FREE Full text] [doi: [10.3945/ajcn.2009.27930](https://doi.org/10.3945/ajcn.2009.27930)] [Medline: [19692494](https://pubmed.ncbi.nlm.nih.gov/19692494/)]
3. El Asmar MS, Naoum JJ, Arbid EJ. Vitamin K dependent proteins and the role of vitamin K2 in the modulation of vascular calcification: a review. *Oman Med J* 2014 May;29(3):172-177 [FREE Full text] [doi: [10.5001/omj.2014.44](https://doi.org/10.5001/omj.2014.44)] [Medline: [24936265](https://pubmed.ncbi.nlm.nih.gov/24936265/)]
4. Walther B, Karl JP, Booth SL, Boyaval P. Menaquinones, bacteria, and the food supply: the relevance of dairy and fermented food products to vitamin K requirements. *Adv Nutr* 2013 Jul 01;4(4):463-473 [FREE Full text] [doi: [10.3945/an.113.003855](https://doi.org/10.3945/an.113.003855)] [Medline: [23858094](https://pubmed.ncbi.nlm.nih.gov/23858094/)]

5. ICH M10 on bioanalytical method validation. European Medicines Agency. URL: <https://www.ema.europa.eu/en/ich-m10-bioanalytical-method-validation> [accessed 2022-01-14]
6. Paroni R, Faioni EM, Razzari C, Fontana G, Cattaneo M. Determination of vitamin K1 in plasma by solid phase extraction and HPLC with fluorescence detection. *J Chromatogr B Analyt Technol Biomed Life Sci* 2009 Jan 15;877(3):351-354. [doi: [10.1016/j.jchromb.2008.12.044](https://doi.org/10.1016/j.jchromb.2008.12.044)] [Medline: [19129011](https://pubmed.ncbi.nlm.nih.gov/19129011/)]
7. Wang LY, Bates CJ, Yan L, Harrington DJ, Shearer MJ, Prentice A. Determination of phylloquinone (vitamin K1) in plasma and serum by HPLC with fluorescence detection. *Clin Chim Acta* 2004 Sep;347(1-2):199-207. [doi: [10.1016/j.cccn.2004.04.030](https://doi.org/10.1016/j.cccn.2004.04.030)] [Medline: [15313159](https://pubmed.ncbi.nlm.nih.gov/15313159/)]
8. Davidson KW, Sadowski JA. Determination of vitamin K compounds in plasma or serum by high-performance liquid chromatography using post-column chemical reduction and fluorometric detection. In: *Methods in Enzymology*. Amsterdam, Netherlands: Elsevier Inc; 1997:408-409.
9. Dupont WD. Power calculations for matched case-control studies. *Biometrics* 1988 Dec;44(4):1157-1168. [Medline: [3233252](https://pubmed.ncbi.nlm.nih.gov/3233252/)]
10. Bullard F. Power in tests of significance. AP Central. URL: <https://apcentral.collegeboard.org/courses/ap-statistics/classroom-resources/power-in-tests-of-significance> [accessed 2022-01-14]
11. Kaneki M, Hodges SJ, Hosoi T, Fujiwara S, Lyons A, Crean SJ, et al. Japanese fermented soybean food as the major determinant of the large geographic difference in circulating levels of vitamin K2: possible implications for hip-fracture risk. *Nutrition* 2001 Apr;17(4):315-321. [doi: [10.1016/s0899-9007\(00\)00554-2](https://doi.org/10.1016/s0899-9007(00)00554-2)] [Medline: [11369171](https://pubmed.ncbi.nlm.nih.gov/11369171/)]
12. Fusaro M, Gallieni M, Rizzo MA, Stucchi A, Delanaye P, Cavalier E, et al. Vitamin K plasma levels determination in human health. *Clin Chem Lab Med* 2017 May 01;55(6):789-799. [doi: [10.1515/ccim-2016-0783](https://doi.org/10.1515/ccim-2016-0783)] [Medline: [27732556](https://pubmed.ncbi.nlm.nih.gov/27732556/)]
13. Sadowski JA, Hood SJ, Dallal GE, Garry PJ. Phylloquinone in plasma from elderly and young adults: factors influencing its concentration. *Am J Clin Nutr* 1989 Jul;50(1):100-108. [doi: [10.1093/ajcn/50.1.100](https://doi.org/10.1093/ajcn/50.1.100)] [Medline: [2750682](https://pubmed.ncbi.nlm.nih.gov/2750682/)]
14. Booth SL, Suttie JW. Dietary intake and adequacy of vitamin K. *J Nutr* 1998 May;128(5):785-788. [doi: [10.1093/jn/128.5.785](https://doi.org/10.1093/jn/128.5.785)] [Medline: [9566982](https://pubmed.ncbi.nlm.nih.gov/9566982/)]

## Abbreviations

**HbA<sub>1c</sub>**: glycosylated hemoglobin

**HPLC**: high-performance liquid chromatography

**KEQAS**: Vitamin K External Quality Assurance Scheme

**MK-7**: menaquinone-7

*Edited by Y Khader; submitted 13.07.21; peer-reviewed by M Salimi, YL Leung; comments to author 16.08.21; revised version received 07.09.21; accepted 20.10.21; published 03.02.22.*

*Please cite as:*

Vaidya R, Vaidya ADB, Sheth J, Jadhav S, Mahale U, Mehta D, Popko J, Badmaev V, Stohs SJ

Vitamin K Insufficiency in the Indian Population: Pilot Observational Epidemiology Study

*JMIR Public Health Surveill* 2022;8(2):e31941

URL: <https://publichealth.jmir.org/2022/2/e31941>

doi: [10.2196/31941](https://doi.org/10.2196/31941)

PMID: [35113033](https://pubmed.ncbi.nlm.nih.gov/35113033/)

©Rama Vaidya, Ashok D B Vaidya, Jayesh Sheth, Shashank Jadhav, Umakant Mahale, Dilip Mehta, Janusz Popko, Vladimir Badmaev, Sidney J Stohs. Originally published in *JMIR Public Health and Surveillance* (<https://publichealth.jmir.org>), 03.02.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

# The Use of Cremation Data for Timely Mortality Surveillance During the COVID-19 Pandemic in Ontario, Canada: Validation Study

Gemma Postill<sup>1,2,3</sup>, BMSch; Regan Murray<sup>2,4</sup>, BSc, MPH; Andrew S Wilton<sup>5</sup>, MSc; Richard A Wells<sup>2</sup>, MD, DPhil; Renee Sirbu<sup>2,3</sup>, BSc; Mark J Daley<sup>1</sup>, PhD; Laura Rosella<sup>3,5,6,7</sup>, HBSc, MHSc, PhD

<sup>1</sup>Department of Epidemiology and Biostatistics, Western University, London, ON, Canada

<sup>2</sup>Office of the Chief Coroner for Ontario, Toronto, ON, Canada

<sup>3</sup>Epidemiology Division, Dalla Lana School of Public Health, Toronto, ON, Canada

<sup>4</sup>Public Health Agency of Canada, Toronto, ON, Canada

<sup>5</sup>Population and Public Health Research Program, Institute for Clinical and Evaluative Sciences, Toronto, ON, Canada

<sup>6</sup>The Vector Institute for Artificial Intelligence, Toronto, ON, Canada

<sup>7</sup>Institute for Better Health, Trillium Health Partners, Mississauga, ON, Canada

**Corresponding Author:**

Laura Rosella, HBSc, MHSc, PhD

Epidemiology Division

Dalla Lana School of Public Health

155 College Street, Suite 600

Toronto, ON, M5T 3M7

Canada

Phone: 1 416 978 0901

Email: [laura.rosella@utoronto.ca](mailto:laura.rosella@utoronto.ca)

## Abstract

**Background:** Early estimates of excess mortality are crucial for understanding the impact of COVID-19. However, there is a lag of several months in the reporting of vital statistics mortality data for many jurisdictions, including across Canada. In Ontario, a Canadian province, certification by a coroner is required before cremation can occur, creating real-time mortality data that encompasses the majority of deaths within the province.

**Objective:** This study aimed to validate the use of cremation data as a timely surveillance tool for all-cause mortality during a public health emergency in a jurisdiction with delays in vital statistics data. Specifically, this study aimed to validate this surveillance tool by determining the stability, timeliness, and robustness of its real-time estimation of all-cause mortality.

**Methods:** Cremation records from January 2020 until April 2021 were compared to the historical records from 2017 to 2019, grouped according to week, age, sex, and whether COVID-19 was the cause of death. Cremation data were compared to Ontario's provisional vital statistics mortality data released by Statistics Canada. The 2020 and 2021 records were then compared to previous years (2017-2019) to determine whether there was excess mortality within various age groups and whether deaths attributed to COVID-19 accounted for the entirety of the excess mortality.

**Results:** Between 2017 and 2019, cremations were performed for 67.4% (95% CI 67.3%-67.5%) of deaths. The proportion of cremated deaths remained stable throughout 2020, even within age and sex categories. Cremation records are 99% complete within 3 weeks of the date of death, which precedes the compilation of vital statistics data by several months. Consequently, during the first wave (from April to June 2020), cremation records detected a 16.9% increase (95% CI 14.6%-19.3%) in all-cause mortality, a finding that was confirmed several months later with cremation data.

**Conclusions:** The percentage of Ontarians cremated and the completion of cremation data several months before vital statistics did not change meaningfully during the COVID-19 pandemic period, establishing that the pandemic did not significantly alter cremation practices. Cremation data can be used to accurately estimate all-cause mortality in near real-time, particularly when real-time mortality estimates are needed to inform policy decisions for public health measures. The accuracy of this excess mortality estimation was confirmed by comparing it with official vital statistics data. These findings demonstrate the utility of cremation data as a complementary data source for timely mortality information during public health emergencies.

**KEYWORDS**

excess deaths; real-time mortality; cremation; COVID-19; SARS-CoV-2; mortality; estimate; impact; public health; validation; pattern; trend; utility; Canada; mortality data; pandemic; death; cremation data; cause of death; vital statistics; excess mortality

## Introduction

Quantifying the impact of COVID-19 on all-cause mortality in a timely manner is critical for understanding the full effect of the COVID-19 pandemic and for enabling evidence-based policy responses. While COVID-19 mortality in Ontario has been routinely tracked and reported through public health databases in near real-time, the reporting of Canadian vital statistics all-cause mortality data for Ontario, prior to the COVID-19 pandemic, was delayed by over a year [1]. Several challenges, including the need to centralize data, verify records, and categorize causes of death, impede using vital statistics data for real-time mortality surveillance [2]. Despite efforts to accelerate mortality reporting in the COVID-19 pandemic, significant data lags persist [3].

Cremation records, however, can be used to provide interim estimates of all-cause mortality [4,5]. This is because, before a cremation can occur, a coroner's certification is required. As a result, cremation data are available in real-time, offering a consistent data source to examine mortality trends in a timelier manner. Early in the first wave (April to June 2020), cremation data detected an increase in all-cause mortality in Ontario, Canada [5]. These findings parallel the increases in mortality observed by several other countries during the first wave of the COVID-19 pandemic [6-10].

Several months later, Statistics Canada published mortality data that also demonstrated an increase in mortality in Ontario during the first wave (April to June 2020) [11]. Despite the detection of excess mortality, few studies systematically examine the performance of cremation data and specifically assess its utility as a real-time mortality surveillance tool. Lacking real-time mortality data posed a challenge for policy response since the magnitude of the mortality impact was unknown and thus influenced mitigation strategies that were implemented. With official mortality statistics available for the first wave of the COVID-19 pandemic, it is now possible to examine the extent to which cremation data accurately estimated the increase in all-cause mortality. This knowledge will be critical in informing whether cremation data can be leveraged as a timelier source of mortality information. Therefore, the objective of this study was to validate the use of cremation data as a surveillance tool for all-cause mortality during a public health emergency in a jurisdiction with delays in vital statistics data. Specifically, this study aimed to validate cremation records by determining (1) the stability of the percent cremated (ie, whether the percent cremated fluctuates by season and/or changes during the pandemic), (2) the timeliness of cremation records, and (3) the robustness/predictive ability of cremation records, measured by their ability to provide accurate estimations of all-cause mortality. The choice of such a measure aligns with the Centers for Disease Control and Prevention's guidelines for evaluating surveillance tools [12]. The framework includes more metrics

than the 3 operationalized here; however, this paper focuses on the aforementioned 3 components, given that they were previously unknown and are key to validating the use of cremation data as a surveillance tool.

## Methods

In Ontario—Canada's most populous province—a cremation certificate must be provided by a coroner to authorize the cremation of a deceased person [13]. A licensed crematorium operator cannot proceed without a cremation certificate. The law in Ontario requires that a coroner review the circumstances surrounding the death before cremation takes place [2,13]. The certificate's review and authorization are documented and kept on file at the crematorium [13]. Since 2017, these records have been collected and stored electronically by the Office of the Chief Coroner for Ontario. This database contains names, dates of birth and death, location of death, and cause of death for every person cremated in the province.

The Canadian Vital Statistics Death database, coordinated by Statistics Canada, collects demographic and cause of death information from all Canadian provinces and territories [14]. The cause of death is classified using the underlying cause of death according to the International Classification of Diseases 10th revision (ICD-10) [14]. The deaths captured in the data set comprise Canadian residents and nonresidents whose deaths occurred in Canada [14]. Routine data processes, including the submission and centralization of death certificates, verification of data, and coding of the cause of death, result in lags in the publication of mortality information [14].

All 323,988 cremations that occurred between January 1, 2017, and May 25, 2021 (with dates of death before April 30, 2021) were deidentified and maintained in an electronic database for the purpose of the analysis. The following analysis was done using Python 3.8.0. A small number of records (n=74) had the cause of death specified as "test" or age at death greater than 120 years, as they were false data used to set up the data set in 2017; these were identified and excluded from the analysis. The records were categorized according to the month of death and subcategorized by age and sex. Age is recorded as a numerical field, but it was converted into a categorical vector (0-44 years, 45-64 years, 65-84 years, and 85+ years) for the purpose of this analysis.

First, the utility of cremation data for surveillance was assessed by determining the percent cremated for the entire population and then by age and sex ([Multimedia Appendix 1](#)). In order to identify any seasonality in the proportion of death cremation, the percentage of deaths cremated was calculated for each week and for the annual quarters over the time period of 2017 to 2021. Standardized differences were used to assess the effect size of any variability in the percent cremated, given that they are independent of sample size [15]. The standardized difference



is the difference in the mean of a variable between 2 groups divided by an estimate of the standard deviation of that variable [15]; relatively low values of the standardized difference show stability in the percent cremated.

Second, time lags in the completeness of cremation and vital statistics data were assessed to determine the predictive value of cremation data in estimating all-cause mortality. The data completion time lags were defined and calculated as the number of weeks between the date of death and when the data source contained >95% or >99% of the deaths that occurred in that week. Statistics Canada releases both provisional estimates of mortality and provisional counts of mortality; the latter was used in this study so that the analysis equivalents (ie, counts of cremation records to counts of vital statistics records) and findings could be generalizable to moments when there is uncertainty in a public health emergency's effect on the reporting of vital statistics data. The data lag for Ontario's vital statistics death data was assessed by comparing the provincial mortality reports released monthly by Statistics Canada [11] between July 24, 2020, and May 14, 2021, as seen in [Multimedia Appendix 2](#). The percentage of the weekly mortality captured in each release was calculated using the weekly totals in the May 14, 2021, release as the denominator. The average time lags for both 95% and 99% completeness were calculated. The same method was applied to the weekly totals of the weekly data cuts of cremation data between May 1, 2020, and October 23, 2020, which were compared to the total number of deaths during that time as available in May 2021.

Third, to assess the predictive ability of cremation records, deviations from provincial mortality trends were quantified and compared both quantitatively and qualitatively to the available vital statistics data [3]. Excess mortality was defined and calculated as the population standardized percentage increase in the number of cremations, which was calculated using risk ratios (RRs) (percent increase =  $RR - 1$ ), where the risk was cremation within the population of Ontario. Cremation records during 2020 and the first half of 2021 were compared to historical records from 2017 to 2019, grouped according to the month of death, and the age and sex of the decedent.

Cremation rates, analogous to mortality rates, were calculated using Statistics Canada's quarterly population estimates [3,16]. Both absolute differences and relative differences (estimated using rate ratios), as compared to historical data, were calculated. The quarterly incident rate ratio was calculated to determine whether the cremation rate changed significantly during the COVID-19 pandemic relative to previous years [3,16,17]. Following the release of vital statistics data, the same methodology of calculating excess mortality in cremation data was applied to the data.

The weekly number of cremations and vital statistics deaths [3] were plotted on side-by-side graphs with 2020-2021 data and baseline data (2017-2019). While the trends were initially congruent, exponential smoothing was used to reduce some of the noise (weekly variability) so that the trends were more comparable; exponential smoothing was specifically chosen as

the method of smoothing given that the data were demonstrated to be nonstationary with the Augmented Dickey-Fuller test [18]. The Statsmodel Holt package was used to exponentially smooth the trends of all graphs (and all subsequent graphs created in the analysis), and the default additive model was changed to an exponential model with a fixed smoothing slope (=2) and smoothing level (=6) [18].

To assess whether confirmed COVID-19 deaths accounted for the entirety of the increase in deaths, excess mortality was calculated, for which COVID-19-classified deaths were removed from the cremation and vital statistics records. Deaths due to COVID-19 were isolated from the records by the presence of the terms "COVID," "novel coronavirus," "Sars-CoV-2," and "coved-19" (a spelling typo) in the *cause of death*, *antecedent cause*, and *other cause of death* categories of the cremation records. Records that matched the above criteria but also contained the phrases "test results pending," "possible," "not," "non," or "negative" were excluded from the classification of death due to COVID-19.

A database containing cremation certificates is held at the Office of Chief Coroner for Ontario. The data-sharing agreement for this study prohibits the data from being publicly available. Data requests may be granted provided there is an appropriate data-sharing agreement in place.

### Ethics Approval

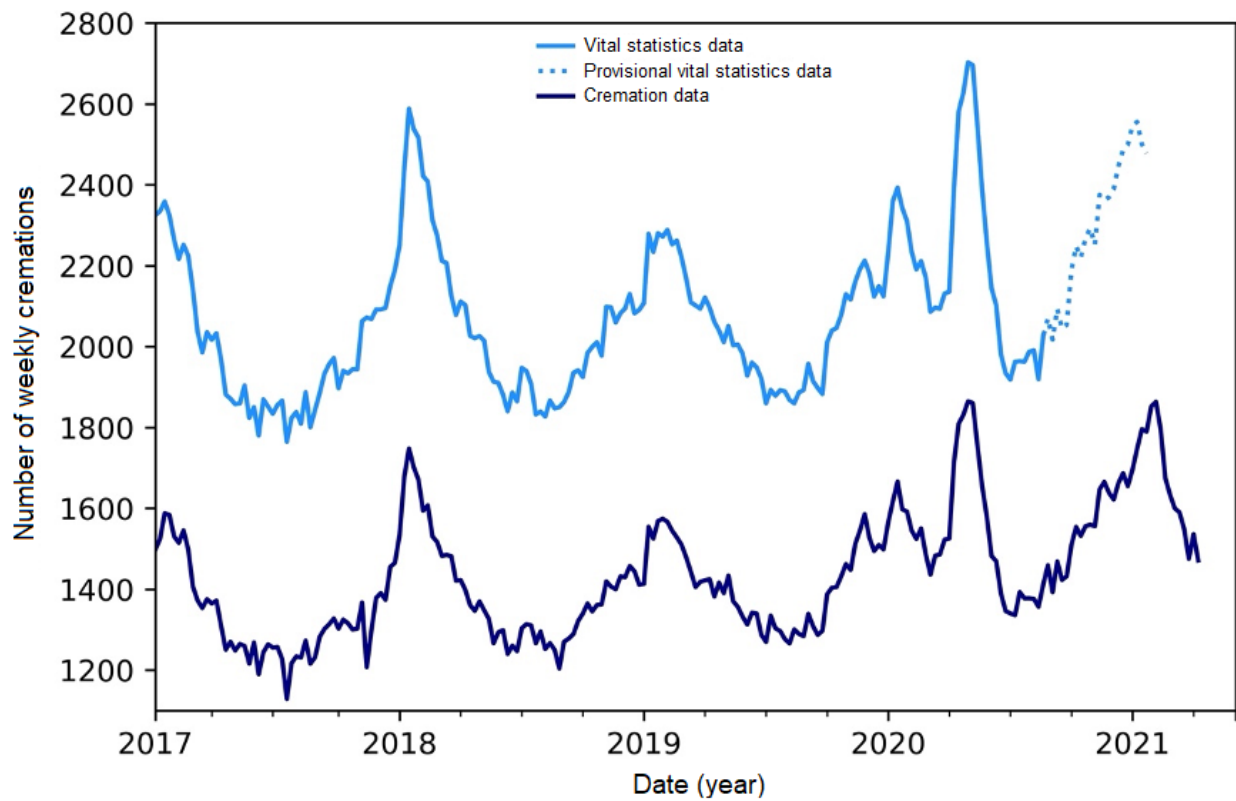
This study was approved by the Research Ethics Board of Western University (project ID: 112478).

## Results

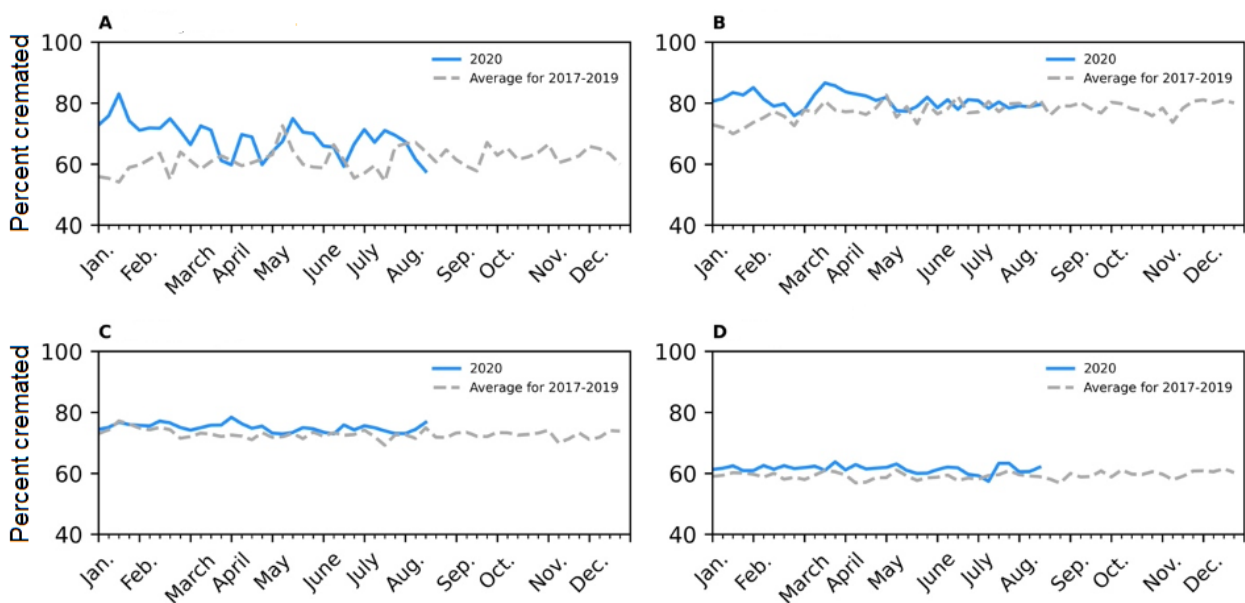
Since the cremation records became electronic (in 2017) and prior to the COVID-19 pandemic, the majority of Ontarians were cremated upon death, with 67.4% (95% CI 67.3%-67.5%) of Ontarians being cremated following death ([Figure 1](#)). There is no seasonality in the percent cremated, with the percent cremated remaining stable throughout the year. Thus, Ontario's cremation data can capture patterns in all-cause mortality between 2017 and 2019, as evidenced in [Figure 1](#). The percent cremated does not differ by sex ([Multimedia Appendix 1](#)). However, there are slight differences by age. The percent cremated is the greatest for those aged 45-64 years and 65-84 years, with approximately 75%-80% and 70%-75%, respectively, being cremated ([Figure 2](#)). The percent cremated for those aged 0-44 years and 85 years or over is approximately 60%.

The percentage of the population cremated remained stable during the COVID-19 pandemic, as seen in [Table 1](#) and [Figure 2](#). Stability was evident in the standardized difference calculations that compared the COVID-19 period to the historical period and obtained a value less than 10% difference. Stability in the percent cremated was also observed within each age group studied (0-44 years, 45-64 years, 65-84 years, and 85 years or over), as seen in [Figure 2](#) and [Multimedia Appendix 1](#).

**Figure 1.** The weekly number of deaths in Ontario, Canada, as reported in Ontario’s cremation records (January 2017 to April 2021, considered >99% complete) and vital statistics records (January 2017 to December 2020, released May 2021). Given that vital statistics records from mid-August (August 16, 2020) and onwards are <95% complete, they are considered provisional. Their respective trends have been smoothed using the Statsmodel Holt package; the default additive model has been changed to an exponential model with a fixed smoothing slope ( $\beta=.2$ ) and smoothing level ( $\alpha=.6$ ).



**Figure 2.** The weekly percent cremated in Ontario, Canada by age group for the average of 2017, 2018, and 2019, in comparison to the percent cremated in 2020. The age groups are as follows: 0-44 years (A), 45-64 years (B), 65-84 years (C), and 85 years or over (D). The respective trends have been smoothed using the Statsmodel Holt package; the default additive model has been changed to an exponential model with a fixed smoothing slope ( $\beta=.2$ ) and smoothing level ( $\alpha=.6$ ). As a single year is compared to the average of 3 years, it was expected that 2020 will display a greater level of weekly fluctuation.



**Table 1.** Stability in the percentage of Ontarians cremated, 2017-2020.

Variable	January to March <sup>a</sup>	April to June <sup>a</sup>	July to September <sup>a</sup>	October to December <sup>a</sup>	January to December
<b>Number of deaths, n</b>					
<b>Baseline (2017-2019)<sup>b</sup></b>					
Cremation records	19,045	17,146	16,884	18,568	71,644
Vital statistics records <sup>c</sup>	28,540	25,443	24,947	27,405	106,335
<b>2020</b>					
Cremation records	20,032	20,737	18,776	21,209	80,754
Vital statistics records <sup>c</sup>	28,675	29,750	N/A <sup>d</sup>	N/A	N/A
<b>Percent cremated (%), value (95% CI)<sup>e</sup></b>					
2017-2019	66.7 (66.4-67.0)	67.4 (67.1-67.7)	67.7 (67.4-68.0)	67.8 (67.5-68.1)	67.4 (67.3-67.5)
2020	69.9 (69.6-70.2)	69.7 (69.4-70.0)	N/A	N/A	N/A
Standardized differences	6.88%	4.96%	N/A	N/A	N/A

<sup>a</sup>For 2020, January to March was the pre-pandemic period, April to June was the first wave of the pandemic, July to September was summer, and October to December was the second wave of the pandemic.

<sup>b</sup>The average number of deaths in 2017, 2018, and 2019 during the same time period.

<sup>c</sup>The number of deaths in Ontario as reported by Statistics Canada in May 2021; at this time, Statistics Canada considers these numbers complete up to the end of July 2020 [3].

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

<sup>e</sup>The 95% CI is calculated using the standard error for population proportions.

In addition, cremation data also provide a timelier source of mortality information, given that cremation records are available much sooner than vital statistics mortality records ([Multimedia Appendix 2](#)). On average, cremation records are >95% complete within 1 week of the date of death and >99% complete within 3 weeks. In contrast, the vital statistics data had an average delay of 27 weeks (range 23-31 weeks) for reporting 95% completeness and a 39-week delay for 99% completeness (range 34-43) ([Multimedia Appendix 2](#)). Thus, the vital statistics data published by Statistics Canada for August 2020 and onwards is <95% complete and, as Statistics Canada states, is provisional [11].

The real-time mortality estimates with cremation data showed, as early as June 2020, that there was excess all-cause mortality during the first wave of the COVID-19 pandemic (April to June 2020; [Table 2](#)). Specifically, when standardized for the provincial population, there was an increase of 16.9% (95% CI 14.6%-19.3%; n=+3591) during April to June when compared with the number of cremations from the previous years (2017-2019), as seen in [Table 2](#). This is in contrast to the 1.7%

increase (95% CI -0.3% to 3.7%; n=+987) seen in January to March of 2020 ([Table 2](#)). Using the provisional vital statistics data released by Statistics Canada [11] and the same methodology for calculating excess mortality that was used for cremation data, there was a 13.1% increase (95% CI 11.2%-15.0%; n=+4307) in mortality during the first wave (April to June 2020).

Cremation data even captured the trends in excess mortality at an age-specific level. When broken down by age, excess mortality during the first wave (April to June 2020) was observed among all age groups in both the cremation and vital statistics data ([Figure 3](#); [Multimedia Appendix 1](#)). Cremation data accurately captured the magnitude of excess in each age group. In both data sets, the largest absolute increase occurred in the older age groups. In fact, 80.3% (1734/2373) of the excess deaths in the cremation records and 83.4% (3587/4302) of the vital statistics deaths during this time period were observed among individuals aged 65 years or older. It is also important to note that the greatest relative change in both cremation and vital statistics records occurred in those less than 45 years of age ([Figure 3](#); [Multimedia Appendix 1](#)).

**Table 2.** Magnitude of excess mortality in Ontario, Canada identified with Ontario's cremation records during the COVID-19 pandemic, January 2020 to March 2021.

Variable	January to March <sup>a</sup>	April to June <sup>a</sup>	July to September <sup>a</sup>	October to December <sup>a</sup>	January to December
<b>Baseline (2017-2019)<sup>b</sup></b>					
Number of cremations	19,045	17,146	16,884	18,568	71,644
Rate of cremations per 100,000, value (95% CI) <sup>c</sup>	134 (132 to 136)	120 (119 to 122)	118 (116 to 120)	129 (127 to 131)	501 (497 to 504)
<b>2020</b>					
Number of cremations	20,032	20,737	18,776	21,209	80,754
Absolute change in the number of cremations <sup>d</sup>	987	3591	1892	2641	9110
Population standardized percentage increase (%) <sup>e</sup> , value (95% CI) <sup>f</sup>	1.7 (-0.3 to 3.7)	16.9 (14.6 to 19.3)	8.0 (5.8 to 10.3)	11.6 (9.4 to 13.8)	12.7 (8.4 to 10.6)
Rate of cremations per 100,000, value (95% CI) <sup>c</sup>	136 (134 to 138)	140 (139 to 143)	127 (126 to 129)	144 (142 to 146)	548 (544 to 552)
Incident rate ratio <sup>g</sup> , value (95% CI)	1.02 (1.00 to 1.04)	1.17 (1.15 to 1.19)	1.08 (1.06 to 1.10)	1.12 (1.09 to 1.14)	1.09 (1.08 to 1.11)
<b>2021</b>					
Number of cremations	21,418	N/A <sup>h</sup>	N/A	N/A	N/A
Absolute change in the number of cremations <sup>d</sup>	2373	N/A	N/A	N/A	N/A
Population standardized percentage increase (%) <sup>e</sup> , value (95% CI) <sup>f</sup>	8.2 (6.1 to 10.3)	N/A	N/A	N/A	N/A
Rate of cremations per 100,000, value (95% CI) <sup>c</sup>	145 (143 to 147)	N/A	N/A	N/A	N/A
Incident rate ratio <sup>g</sup> , value (95% CI)	1.08 (1.06 to 1.10)	N/A	N/A	N/A	N/A

<sup>a</sup>For 2020, January to March was the pre-pandemic period, April to June was the first wave of the pandemic, July to September was summer, and October to December was the second wave of the pandemic. For 2021, January to March and April to June involved the third wave.

<sup>b</sup>The average of the number of deaths in 2017, 2018, and 2019 during the same time period.

<sup>c</sup>Cremation rates, analogous to mortality rates, were calculated as the number of cremations divided by the provincial quarterly population estimates published by Statistics Canada [16].

<sup>d</sup>Absolute change refers to the difference in the number between 2020/2021 and the baseline (2017-2019).

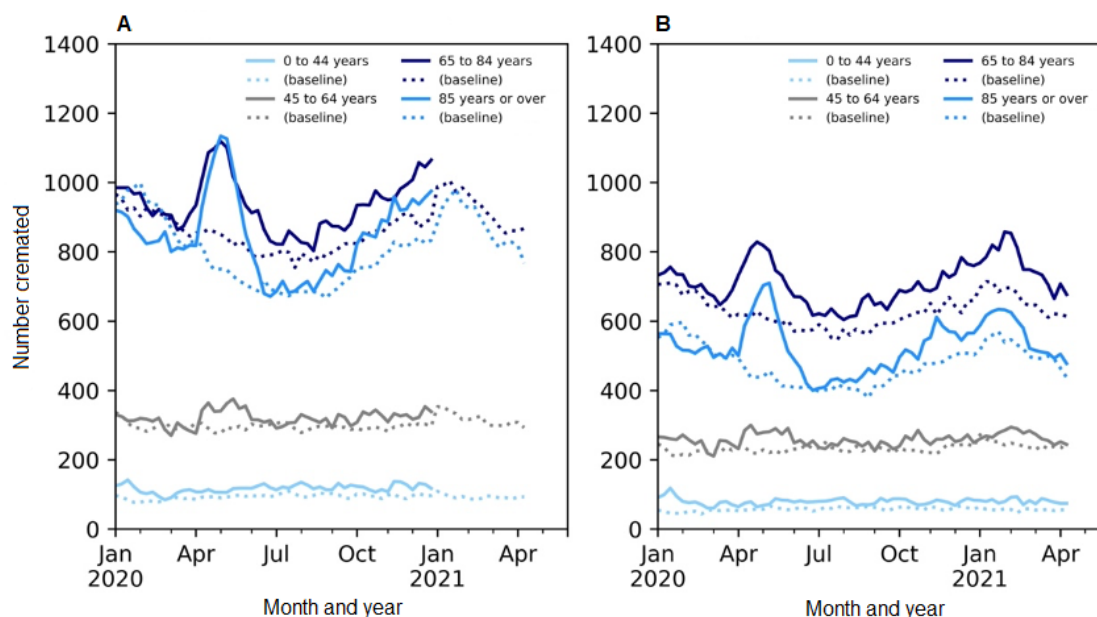
<sup>e</sup>The population standardized percentage increase is calculated as risk ratio (RR) - 1, where RR is the incidence of death (measured as the number of cremation) in the quarterly population estimates. The Q3 population estimate was used for the January-December RR.

<sup>f</sup>The 95% CI for the percentage increase is calculated as  $(RR \text{ lower bound} - 1) \times 100\%$  to  $(RR \text{ upper bound} + 1) \times 100\%$ . The RR CI is calculated as  $=\text{EXP}(\text{LN}(\text{RR}) - (1.96 \times \text{SE}))$ , with  $\text{SE}(\text{ln}(\text{rr})) = \sqrt{(1/\text{Ncrem}_{(2017-19)} - 1/\text{Npop}_{(2017-19)} + 1/\text{Ncrem}_{(2020)} - 1/\text{Npop}_{(2020)})}$ .

<sup>g</sup>The quarterly incident rate ratio was calculated by dividing the rate of cremations in 2020 to that of the baseline. The 95% CI was calculated as  $=(\text{Events} \pm 1.96 \times \text{SE}) / \text{population} \times 100,000$ , where SE is the standard error equal to the square root of the number of cremations [17].

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

**Figure 3.** Side-by-side comparison of the weekly number of deaths in Ontario, Canada during the pandemic by age group as reported in (A) vital statistics data, which contains all provincial deaths and is released by Statistics Canada [3], and (B) Ontario's cremation data for January 2020 to April 2021, which refers to the baseline data (the average of data in 2017-2019). The annual trends have been smoothed using the Statsmodel Holt package; the default additive model has been changed to an exponential model with a fixed smoothing slope ( $\beta=.2$ ) and smoothing level ( $\alpha=.6$ ).

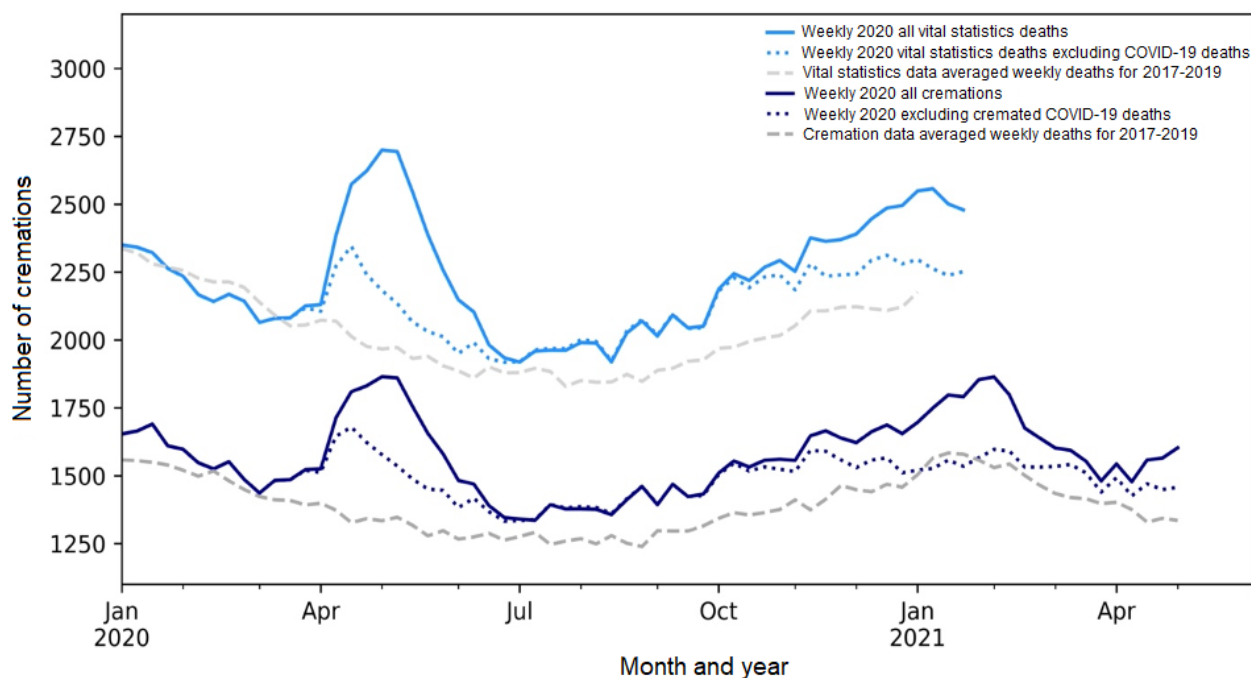


Most recently, in the latter half of 2020 (the second wave) and the beginning of 2021 (the third wave), there was still significant excess mortality (Table 2). Specifically, after adjusting for population, there was an 8.0% increase (95% CI 5.8%-10.3%) in mortality in the summer months (July to September 2020), a 11.6% increase (95% CI 9.4%-13.8%) in the fall (October to December 2020), and a 9.5% increase (95% CI 8.4%-10.6%) during January to March 2021, relative to the same period for previous years. During this time, the absolute unadjusted excess mortality among those aged 0-44 years accounted for a greater percentage of the overall excess mortality (22%-27% of the excess) than in the first wave (20% of the absolute unadjusted excess). In fact, unlike any other age group, excess mortality for those aged 0-44 years in the COVID-19 pandemic was the greatest in early 2021 (January to March). In contrast, during this time (January to March 2021), excess mortality for those aged 85 years or older was lower than at any other time in the COVID-19 pandemic (Table 2).

Beyond simply highlighting excess mortality, cremation data demonstrated that, as early as June 2020, deaths due to COVID-19 could not explain all of the excess mortality observed during the pandemic (Figure 4; Multimedia Appendix 3). With vital statistics data available for the first wave, it is clear that cremation data accurately estimated the magnitude of excess non-COVID-19 mortality (Figure 4).

During the period from March 23, 2020, to July 5, 2020, cremation data captured 55.6% (1638/2945) of the provincially reported COVID-19 deaths, which is consistent with the percent cremated among those aged 85 years or older (the age group with the majority of COVID-19 deaths during this time). When these COVID-19 deaths were removed, there was still significant excess mortality during the last 3 quarters of 2020. Specifically, there was a 7.9% (95% CI 5.7%-10.1%) increase in April to June 2020, 7.5% (95% CI 5.3%-7.8%) increase in July to August 2020, and 5.9% (95% CI 3.9%-8.1%) increase in September to December 2020. Elevated non-COVID-19 mortality was not detected in the first quarter of 2021 (Multimedia Appendix 3).

**Figure 4.** Annual trends of the weekly number of deaths with and without confirmed COVID-19 deaths for the cremation records and vital statistics data for Ontario, Canada (released May 2021). The pandemic waves in Ontario, Canada captured in this graph are as follows: wave 1 (April 2020 to June 2020), wave 2 (September 2020 to February 2021), and wave 3 (April 2021 onwards). Both trends have been smoothed using an exponential model with a fixed smoothing slope ( $\beta=.2$ ) and smoothing level ( $\alpha=.6$ ).



## Discussion

Real-time mortality information is an essential indicator to monitor during the COVID-19 pandemic. Many jurisdictions experience delays in official vital statistics data due to verification processes, and as a result, it can be difficult to obtain timely mortality information. This study aimed to provide evidence of the utility of cremation data as an early indicator of mortality trends during the COVID-19 pandemic. Within 3 weeks, cremation data captured approximately 70% of overall mortality, anticipating vital statistics mortality records by approximately 5 months. Timeliness was of critical value given that during this time, several public health measures and policies related to the pandemic response were being decided in the absence of mortality data in many jurisdictions, including in Canada, where official vital statistics reports lag by several months.

Additionally, this analysis demonstrated that the percentage of the population cremated remained stable during the COVID-19 pandemic. This was a key finding given that, at the start of the pandemic, there was a concern as to whether a shift in burial practices drove the increase in cremations. This was, in fact, not the case and aligned with the fact that there was no change in the guidance of embalming. These findings provide further confidence for the use of cremation data to obtain a robust and real-time estimate of mortality during public health emergencies in jurisdictions where the death investigation system is equipped with digital reporting tools.

In demonstrating the ability of cremation data to capture population-level mortality trends, the analysis captured excess mortality, that is, mortality beyond what is normally anticipated relative to the average of the 3 prior years. COVID-19

cremations could not account for all of this excess. Excess mortality due to COVID-19 has been demonstrated in several countries, including the United States [19-21], Portugal [9], Sweden [22], England [23], and many others [6-8,10,24,25]. While overall excess mortality (identified by this study) during the first wave in Canada (April to June 2020) is lower than in some countries, such as Spain [8] and England [23], it is substantially higher than in other industrialized countries, including Germany [26], Sweden [22], and Norway. The age structure of excess mortality that we report here is similar to that found by the European monitoring of excess mortality for public health action (EuroMOMO) network, with the majority of the excess occurring in older age groups [7]. Other provinces in Canada, such as Alberta and British Columbia, also detected a similar magnitude of excess mortality in 2020 [11,27]. Consequently, cremation data can provide an understanding of the most recent mortality trends (the fourth and fifth waves), for which vital statistics data are still provisional.

In addition, our results confirming that COVID-19 deaths did not account for the entirety of the excess mortality in Ontario are supported by similar findings in several other countries [8,9]. Excess mortality that cannot be accounted for by confirmed COVID-19 is attributable to several factors, and these factors likely changed throughout the pandemic, reflecting the various patterns seen in Figure 4. The initial large increase in non-SARS-CoV-2 mortality during March to April 2020 was likely due to underdiagnosis and underreporting of COVID-19 on death certificates, delays in care for conditions other than COVID-19, including cancer and cardiovascular care, worsening mental health and substance use, and hesitancy or fear that deterred patients from seeking emergency care for cardiac events, all of which have been observed as causes of non-COVID-19 excess in other jurisdictions [28-35]. The

decline in emergency care-seeking is supported by data showing that Canadian emergency department volumes dropped during the first wave of the COVID-19 pandemic [17,36]. Subsequent excess mortality, when COVID-19 testing was more widely employed, is likely due to the indirect effects of the pandemic, including acute drug toxicity, violence, and the economic and social disruptions of the pandemic [30,32,37]. Specifically, in Ontario, opioid-related deaths have increased by 79%, with the majority of these deaths occurring among those aged 25-44 years [37]. These increases in all-cause mortality highlight the value and necessity of real-time surveillance of population mortality rates, especially during public health emergencies. Given their robust and real-time nature, cremation data offer a critical source of mortality data that can provide these insights in the interim period before vital statistics data are available.

An important limitation of the study is the assumption of minimal growth in the population cremated. The quantifications of excess should be interpreted with these limitations in mind, particularly given that they contribute to the discrepancies in the reported values of excess mortality following the release of provisional vital statistics mortality data by Statistics Canada after several months, using the Farrington method. However, small discrepancies in the magnitude of excess mortality reported do not invalidate the use of cremation data, given that the value of a surveillance tool is in its ability to be both timely and accurate, and cremation data display strong congruency to vital statistics data, as seen in [Figure 4](#). Likewise, with other data sources that can be used to create robust modes, the long reporting delays introduce significant biases. Thus, the methodology that is employed and able to communicate these results in a timely manner clearly shows the value in using Ontario's cremation records to provide the earliest indication of excess all-cause mortality in a region with delays in vital statistics reporting.

There are some additional limitations to this analysis that should be considered. First, there was no advice against embalming during the COVID-19 pandemic; thus, in a public health emergency wherein embalming may be advised against, the utility of cremation data would need to be further investigated. Second, this study looked at all-cause mortality, and therefore,

it is not possible to conclude whether COVID-19 affected the likelihood of being cremated. However, we found no evidence of this, given that there was no relationship between the cause of death and the likelihood of cremation. The next steps of this research include determining the methodology for using cremation data to assess the effect of the COVID-19 pandemic on other specific causes (ie, cardiac events) and manners (ie, suicide) of death. Finally, it is important to note that while cremation data represent most deaths in Ontario, they do not represent all deaths. Certain segments of the population may have an intrinsically higher or lower cremation rate (ie, people of Jewish faith). This is an important consideration, and if some segments of the population have different rates of cremation and different rates of mortality, the estimates would be biased. However, this is not an invalidating limitation for a surveillance model. Likewise, owing to the structure of the cremation data set (ie, absence of ethnic data), we were unable to assess whether the magnitude of excess mortality differed among different subgroups of the population. Despite these limitations, the results compellingly demonstrate the utility of cremation data as an important surveillance data source when timelier estimates of mortality are needed, such as during a public health emergency. The purpose of this analysis was to demonstrate cremation data as a robust estimator of all-cause mortality in public health emergencies in the interim of vital statistics reporting, and not as a replacement of vital statistics data or as a predictor of a pandemic's effects on subpopulation mortality.

In conclusion, the COVID-19 pandemic emphasized the importance and need for real-time mortality information. Cremation data can be used for monitoring all-cause mortality in conjunction with vital statistics data. In addition, other jurisdictions with a lack of real-time mortality surveillance and high cremation rates will likely benefit from leveraging cremation data to estimate the complete impact of a public health emergency on all-cause mortality. This is an important finding, given that the utility of cremation data to estimate all-cause mortality in a public health emergency was previously unknown. This study demonstrates that cremation records can provide robust and timely indicators of all-cause mortality and should be used as interim mortality data during a public health emergency where more timely data can support the response.

---

## Acknowledgments

LR receives funding from the Canadian Institutes of Health Research and the Canada Research Chairs Program. MJD is the SHARCNET Research Chair in biocomputing. The authors acknowledge Dr Dirk Huyer, Chief Coroner for Ontario for the vision and support to analyze cremation certificates as a timely source of mortality data.

---

## Authors' Contributions

All authors contributed to the conception and design of the work; acquisition, analysis, and interpretation of data; and drafting the work and revising it critically. RM developed the concept for the study and coordinated the collaboration. GP conducted all analyses. GP and MJD had access to the data. All authors gave final approval to the version to be published and agreed to be held accountable for all aspects of the work.

---

## Conflicts of Interest

LR serves on the steering committee of the CDL-Rapid Screening Consortium, a government-funded not-for-profit endeavor to build a scalable workplace COVID screening system in workplaces. The other authors have no conflicts.

---

## Multimedia Appendix 1

Further analysis on the percent cremated by age and sex.

[[DOCX File , 41 KB - publichealth\\_v8i2e32426\\_app1.docx](#) ]

## Multimedia Appendix 2

Maturation of both cremation data and vital statistics data.

[[XLSX File \(Microsoft Excel File\), 46 KB - publichealth\\_v8i2e32426\\_app2.xlsx](#) ]

## Multimedia Appendix 3

Calculations for non-COVID-19 excess.

[[DOCX File , 17 KB - publichealth\\_v8i2e32426\\_app3.docx](#) ]

**References**

1. Mortality rates, by age group. Statistics Canada. URL: <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310071001> [accessed 2022-02-01]
2. Vital Statistics Act, R.S.O. 1990. Government of Ontario. URL: <https://www.ontario.ca/laws/statute/90v04> [accessed 2022-02-01]
3. Provisional weekly death counts, by age group and sex. Statistics Canada. URL: <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310076801> [accessed 2022-02-01]
4. Postill G, Murray R, Wilton A, Wells R, Sirbu R, Daley M, et al. Excess Mortality in Ontario During the COVID-19 Pandemic. Ontario COVID-19 Science Advisory Table. URL: <https://covid19-sciencetable.ca/sciencebrief/excess-mortality-in-ontario-during-the-covid-19-pandemic/> [accessed 2022-02-01]
5. Postill G, Murray R, Wilton A, Wells R, Sirbu R, Daley M, et al. An analysis of mortality in Ontario using cremation data: Rise in cremations during the COVID-19 pandemic. medRxiv. URL: <https://www.medrxiv.org/content/10.1101/2020.07.22.20159913v3> [accessed 2022-02-01]
6. Docherty K, Butt J, de Boer R, Dewan P, Køber L, Maggioni A, et al. Excess deaths during the Covid-19 pandemic: An international comparison. medRxiv. URL: <https://www.medrxiv.org/content/10.1101/2020.04.21.20073114v3> [accessed 2022-02-01]
7. Vestergaard L, Nielsen J, Richter L, Schmid D, Bustos N, Braeye T, ECDC Public Health Emergency Team for COVID-19, et al. Excess all-cause mortality during the COVID-19 pandemic in Europe - preliminary pooled estimates from the EuroMOMO network, March to April 2020. Euro Surveill 2020 Jul;25(26):2 [FREE Full text] [doi: [10.2807/1560-7917.ES.2020.25.26.2001214](https://doi.org/10.2807/1560-7917.ES.2020.25.26.2001214)] [Medline: [32643601](https://pubmed.ncbi.nlm.nih.gov/32643601/)]
8. Measuring the direct and indirect impact of COVID-19. OECD iLibrary. URL: [https://www.oecd-ilibrary.org/social-issues-migration-health/excess-mortality\\_e5dc0c50-en](https://www.oecd-ilibrary.org/social-issues-migration-health/excess-mortality_e5dc0c50-en) [accessed 2022-02-01]
9. Vieira A, Peixoto VR, Aguiar P, Abrantes A. Rapid Estimation of Excess Mortality during the COVID-19 Pandemic in Portugal -Beyond Reported Deaths. J Epidemiol Glob Health 2020 Sep;10(3):209-213 [FREE Full text] [doi: [10.2991/jegh.k.200628.001](https://doi.org/10.2991/jegh.k.200628.001)] [Medline: [32954711](https://pubmed.ncbi.nlm.nih.gov/32954711/)]
10. Karlinsky A, Kobak D. The World Mortality Dataset: Tracking excess mortality across countries during the COVID-19 pandemic. medRxiv. URL: <https://www.medrxiv.org/content/10.1101/2021.01.27.21250604v3> [accessed 2022-02-01]
11. Previous releases and revisions to provisional weekly death counts. Statistics Canada. URL: <https://www150.statcan.gc.ca/n1/en/catalogue/13100783> [accessed 2022-02-01]
12. German RR, Lee LM, Horan JM, Milstein RL, Pertowski CA, Waller MN, Guidelines Working Group Centers for Disease Control/Prevention (CDC). Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. MMWR Recomm Rep 2001 Jul 27;50(RR-13):1-35; quiz CE1. [Medline: [18634202](https://pubmed.ncbi.nlm.nih.gov/18634202/)]
13. Request a certificate to cremate a body or ship a body out of province. Government of Ontario. URL: <https://www.ontario.ca/page/request-certificate-cremate-body-or-ship-body-out-province> [accessed 2022-02-01]
14. Statistics Canada, Canadian Vital Statistics - Death database (CVSD). Statistics Canada. URL: <https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3233> [accessed 2022-02-01]
15. Austin PC. Using the Standardized Difference to Compare the Prevalence of a Binary Variable Between Two Groups in Observational Research. Communications in Statistics - Simulation and Computation 2009 Apr 09;38(6):1228-1234. [doi: [10.1080/03610910902859574](https://doi.org/10.1080/03610910902859574)]
16. Population estimates, quarterly. Statistics Canada. URL: <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710000901> [accessed 2022-02-01]
17. Gomez D, Simpson AN, Sue-Chue-Lam C, de Mestral C, Dossa F, Nantais J, et al. A population-based analysis of the impact of the COVID-19 pandemic on common abdominal and gynecological emergency department visits. CMAJ 2021 May 25;193(21):E753-E760 [FREE Full text] [doi: [10.1503/cmaj.202821](https://doi.org/10.1503/cmaj.202821)] [Medline: [34035055](https://pubmed.ncbi.nlm.nih.gov/34035055/)]
18. Statsmodels. URL: <https://www.statsmodels.org/stable/generated/statsmodels.tsa.holtwinters.Holt.html> [accessed 2022-02-01]



19. New York City Department of Health Mental Hygiene (DOHMH) COVID-19 Response Team. Preliminary Estimate of Excess Mortality During the COVID-19 Outbreak - New York City, March 11-May 2, 2020. *MMWR Morb Mortal Wkly Rep* 2020 May 15;69(19):603-605 [FREE Full text] [doi: [10.15585/mmwr.mm6919e5](https://doi.org/10.15585/mmwr.mm6919e5)] [Medline: [32407306](https://pubmed.ncbi.nlm.nih.gov/32407306/)]
20. Rivera R, Rosenbaum JE, Quispe W. Excess mortality in the United States during the first three months of the COVID-19 pandemic. *Epidemiol. Infect* 2020 Oct 29;148:e264. [doi: [10.1017/s0950268820002617](https://doi.org/10.1017/s0950268820002617)]
21. Woolf SH, Chapman DA, Sabo RT, Zimmerman EB. Excess Deaths From COVID-19 and Other Causes in the US, March 1, 2020, to January 2, 2021. *JAMA* 2021 Apr 02;325(17):1786 [FREE Full text] [doi: [10.1001/jama.2021.5199](https://doi.org/10.1001/jama.2021.5199)] [Medline: [33797550](https://pubmed.ncbi.nlm.nih.gov/33797550/)]
22. Modig K, Ahlbom A, Ebeling M. Excess mortality from COVID-19: weekly excess death rates by age and sex for Sweden and its most affected region. *Eur J Public Health* 2021 Feb 01;31(1):17-22 [FREE Full text] [doi: [10.1093/eurpub/ckaa218](https://doi.org/10.1093/eurpub/ckaa218)] [Medline: [33169145](https://pubmed.ncbi.nlm.nih.gov/33169145/)]
23. Sinnathamby M, Whitaker H, Coughlan L, Lopez Bernal J, Ramsay M, Andrews N. All-cause excess mortality observed by age group and regions in the first wave of the COVID-19 pandemic in England. *Euro Surveill* 2020 Jul;25(28):16 [FREE Full text] [doi: [10.2807/1560-7917.ES.2020.25.28.2001239](https://doi.org/10.2807/1560-7917.ES.2020.25.28.2001239)] [Medline: [32700669](https://pubmed.ncbi.nlm.nih.gov/32700669/)]
24. Ghislandi S, Muttarak R, Sauerberg M, Scotti B. News from the front: Estimation of excess mortality and life expectancy in the major epicenters of the COVID-19 pandemic in Italy. medRxiv. URL: <https://www.medrxiv.org/content/10.1101/2020.04.29.20084335v3> [accessed 2022-02-01]
25. Kontis V, Bennett JE, Rashid T, Parks RM, Pearson-Stuttard J, Guillot M, et al. Magnitude, demographics and dynamics of the effect of the first wave of the COVID-19 pandemic on all-cause mortality in 21 industrialized countries. *Nat Med* 2020 Dec 14;26(12):1919-1928. [doi: [10.1038/s41591-020-1112-0](https://doi.org/10.1038/s41591-020-1112-0)] [Medline: [33057181](https://pubmed.ncbi.nlm.nih.gov/33057181/)]
26. Stang A, Standl F, Kowall B, Brune B, Böttcher J, Brinkmann M, et al. Excess mortality due to COVID-19 in Germany. *J Infect* 2020 Nov;81(5):797-801 [FREE Full text] [doi: [10.1016/j.jinf.2020.09.012](https://doi.org/10.1016/j.jinf.2020.09.012)] [Medline: [32956730](https://pubmed.ncbi.nlm.nih.gov/32956730/)]
27. Moriarty T, Boczula A, Thind E, Loreto N, McElhaney J. Excess All-Cause Mortality During the COVID-19 Epidemic in Canada. Royal Society of Canada. URL: [https://rsc-src.ca/sites/default/files/EM%20PB\\_EN.pdf](https://rsc-src.ca/sites/default/files/EM%20PB_EN.pdf) [accessed 2022-02-01]
28. Dinmohamed AG, Visser O, Verhoeven RHA, Louwman MWJ, van Nederveen FH, Willems SM, et al. Fewer cancer diagnoses during the COVID-19 epidemic in the Netherlands. *The Lancet Oncology* 2020 Jun;21(6):750-751 [FREE Full text] [doi: [10.1016/S1470-2045\(20\)30265-5](https://doi.org/10.1016/S1470-2045(20)30265-5)] [Medline: [32359403](https://pubmed.ncbi.nlm.nih.gov/32359403/)]
29. Rosenbaum L. The Untold Toll — The Pandemic’s Effects on Patients without Covid-19. *N Engl J Med* 2020 Jun 11;382(24):2368-2371. [doi: [10.1056/nejmms2009984](https://doi.org/10.1056/nejmms2009984)]
30. Baumgartner J, Radley D. The Spike in Drug Overdose Deaths During the COVID-19 Pandemic and Policy Options to Move Forward. Commonwealth Fund. URL: <https://www.commonwealthfund.org/blog/2021/spike-drug-overdose-deaths-during-covid-19-pandemic-and-policy-options-move-forward> [accessed 2022-02-01]
31. Appleby L, Kapur N, Turnbull P, Richards N. Suicide in England since the COVID-19 pandemic - early figures from real-time surveillance. The University of Manchester. URL: <https://documents.manchester.ac.uk/display.aspx?DocID=51861> [accessed 2022-02-01]
32. Gunnell D, Appleby L, Arensman E, Hawton K, John A, Kapur N, et al. Suicide risk and prevention during the COVID-19 pandemic. *The Lancet Psychiatry* 2020 Jun;7(6):468-471. [doi: [10.1016/s2215-0366\(20\)30171-1](https://doi.org/10.1016/s2215-0366(20)30171-1)]
33. Hammad TA, Parikh M, Tashtish N, Lowry CM, Gorbey D, Forouzandeh F, et al. Impact of COVID-19 pandemic on ST-elevation myocardial infarction in a non-COVID-19 epicenter. *Catheter Cardiovasc Interv* 2021 Feb 01;97(2):208-214 [FREE Full text] [doi: [10.1002/ccd.28997](https://doi.org/10.1002/ccd.28997)] [Medline: [32478961](https://pubmed.ncbi.nlm.nih.gov/32478961/)]
34. Rattka M, Baumhardt M, Dreyhaupt J, Rothenbacher D, Thiessen K, Markovic S, et al. 31 days of COVID-19-cardiac events during restriction of public life-a comparative study. *Clin Res Cardiol* 2020 Dec 03;109(12):1476-1482 [FREE Full text] [doi: [10.1007/s00392-020-01681-2](https://doi.org/10.1007/s00392-020-01681-2)] [Medline: [32494921](https://pubmed.ncbi.nlm.nih.gov/32494921/)]
35. Clifford CR, Le May M, Chow A, Boudreau R, Fu AY, Barry Q, et al. Delays in ST-Elevation Myocardial Infarction Care During the COVID-19 Lockdown: An Observational Study. *CJC Open* 2020 Dec 15;3(5):565-573 [FREE Full text] [doi: [10.1016/j.cjco.2020.12.009](https://doi.org/10.1016/j.cjco.2020.12.009)] [Medline: [33521615](https://pubmed.ncbi.nlm.nih.gov/33521615/)]
36. COVID-19’s impact on emergency departments. Canadian Institute for Health Information. URL: <https://www.cihi.ca/en/covid-19-resources/impact-of-covid-19-on-canadas-health-care-systems/how-covid-19-affected> [accessed 2022-02-01]
37. Gomes T, Murray R, Kolla G, Leece P, Bansal S, Besharah J, et al. Changing Circumstances Surrounding Opioid-Related Deaths in Ontario during the COVID-19 Pandemic. The Ontario Drug Policy Research Network. URL: <https://odprn.ca/wp-content/uploads/2021/05/Changing-Circumstances-Surrounding-Opioid-Related-Deaths.pdf> [accessed 2022-02-01]

## Abbreviations

**RR:** risk ratio

*Edited by T Sanchez, G Eysenbach; submitted 27.07.21; peer-reviewed by T Kosatsky, P Dion; comments to author 14.12.21; revised version received 02.01.22; accepted 06.01.22; published 21.02.22.*

*Please cite as:*

*Postill G, Murray R, Wilton AS, Wells RA, Sirbu R, Daley MJ, Rosella L*

*The Use of Cremation Data for Timely Mortality Surveillance During the COVID-19 Pandemic in Ontario, Canada: Validation Study*  
*JMIR Public Health Surveill* 2022;8(2):e32426

URL: <https://publichealth.jmir.org/2022/2/e32426>

doi: [10.2196/32426](https://doi.org/10.2196/32426)

PMID: [35038302](https://pubmed.ncbi.nlm.nih.gov/35038302/)

©Gemma Postill, Regan Murray, Andrew S Wilton, Richard A Wells, Renee Sirbu, Mark J Daley, Laura Rosella. Originally published in JMIR Public Health and Surveillance (<https://publichealth.jmir.org>), 21.02.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

# Use of a Smartphone Self-assessment App for a Tobacco-Induced Disease (COPD, Cardiovascular Diseases, Cancer) Screening Strategy and to Encourage Smoking Cessation: Observational Study

Edouard Stavaux<sup>1</sup>, MD; François Goupil<sup>2</sup>, MD; Guillaume Barreau<sup>1</sup>, MSc; Anne Lise Septans<sup>3</sup>, PhD; Bertrand Dautzenberg<sup>4</sup>, MD, PhD; Armelle Foulet-Rogé<sup>5</sup>, MD; Norbert Padilla<sup>6</sup>, MD; Thierry Urban<sup>7</sup>, MD, PhD; Fabrice Denis<sup>8</sup>, MD, PhD

<sup>1</sup>Centre hospitalier universitaire de Nantes, Nantes, France

<sup>2</sup>Service de Pneumologie, Centre Hospitalier du Mans, Le Mans, France

<sup>3</sup>Weprom, Angers, France

<sup>4</sup>Service de Pneumologie, Hôpital de la Salpêtrière, Paris, France

<sup>5</sup>Service d'Anatomo-Pathologie, Centre Hospitalier du Mans, Le Mans, France

<sup>6</sup>Centre de pathologie Maine Normandie, Le Mans, France

<sup>7</sup>Service de Pneumologie, Centre Hospitalier Universitaire d'Angers, Angers, France

<sup>8</sup>Institut Inter-Regional de Cancérologie Jean Bernard, Le Mans, France

**Corresponding Author:**

Fabrice Denis, MD, PhD

Institut Inter-Regional de Cancérologie Jean Bernard

9 rue Beauverger

Le Mans, 72000

France

Phone: 33 684190480

Email: [f.denis@ilcgroupe.fr](mailto:f.denis@ilcgroupe.fr)

## Abstract

**Background:** Patient self-assessment via a mobile app detects actionable symptoms and has been shown to detect lung cancer relapses early, thereby lengthening survival.

**Objective:** The purpose of this study was to assess the incidence of chief symptoms associated with the main tobacco-induced pathologies in both current and ex-smokers through a self-assessment smartphone app and to evaluate the app's capacity to encourage users to quit smoking or reduce consumption, as well as its impact on early lung cancer stages at the time of diagnosis.

**Methods:** Current and ex-smokers were recruited through an advertising campaign in Sarthe county (France) proposing the free download of a smartphone app. App users were asked to answer 13 questions related to symptoms associated with tobacco-induced diseases (chronic obstructive pulmonary disease [COPD], cardiovascular diseases, cancer). In the event of any positive answer, a message was displayed recommending the user to consult a physician. In addition, they were asked about smoking cessation intention before and after answering these 13 questions. Finally, incidence of stage 1 or 2 lung cancers diagnosed during the launch period of our application was evaluated by comparing data from various sources to those from the same period during the previous year.

**Results:** Of the 5671 users who were eligible for evaluation, an alert was sent to the majority (4118/5671, 72.6%), with a higher incidence for current smokers (2833/3679, 77.0% vs 1298/1992, 65.2%;  $P < .001$ ). The most frequent symptoms triggering the notifications were fatigue (2023/5671, 35.7%), cough (1658/5671, 29.2%), dyspnea (1502/5671, 26.5%), and persistent chest pain (1286/5671, 22.7%). Of the current smokers, 14.0% (515/3679) showed symptoms suggesting COPD, 15.5% (571/3679) showed symptoms suggesting stable angina, 12.4% (455/3679) probably had lower extremity artery disease, and 6.8% (249/3679) had possible cancer. Of the users, 36.5% (1343/3679) claimed that they thought about quitting smoking, and 48.7% (1795/3679) had thought about reducing their consumption. Surgery-eligible stage 1 and 2 lung cancer incidence was 24% (14/58) during the

study period versus 9% (5/54) during the previous year in Sarthe county ( $P=.04$ ), whereas it remained unchanged in the neighboring county of Maine-et-Loire.

**Conclusions:** A majority of current and ex-smokers showed worrying symptoms, and the use of a self-assessment smartphone app may drive a majority of smokers toward the intention of smoking cessation or decreasing consumption. A randomized study should be performed to confirm this intention and to support the potential increase of symptomatic lung cancer detection at early, surgery-accessible stages.

**Trial Registration:** ClinicalTrials.gov NCT04048954; <https://www.clinicaltrials.gov/ct2/show/NCT04048954>

(*JMIR Public Health Surveill* 2022;8(2):e19877) doi:[10.2196/19877](https://doi.org/10.2196/19877)

## KEYWORDS

smoking cessation; mobile health; self-assessment, lung cancer; early detection; tobacco-induced pathologies

## Introduction

Lung cancer is the most frequent cancer worldwide (2.09 million new cases in 2018) and the leading cause of death by cancer (1.76 million deaths in 2018) [1], whereas it is linked to an identifiable and avoidable risk factor in 90% of cases: tobacco. Moreover, 70% to 80% of lung cancers are diagnosed at a late and untreatable stage [2].

In addition, tobacco consumption may also result in other tumoral diseases (22% of deaths by cancer), cardiovascular diseases (CVD), and chronic pulmonary diseases such as chronic obstructive pulmonary disease (COPD), extremely frequent yet underevaluated pathologies (80% of the concerned patients ignore their status) related to tobacco consumption in 85% of cases. Nonspecific symptoms are commonly disregarded by patients [3] even though COPD was expected to be the third cause of morbidity and mortality worldwide in 2020. Evolution of COPD includes exacerbations, notably reducing patients' quality of life, and a higher risk of lung cancer.

Yet, despite tobacco-cessation strategies ranging from awareness-raising campaigns to nicotine-replacement therapies, and—more recently—electronic cigarettes, long-term tobacco-cessation rates remain low [4].

New information and communication technology services in the form of mobile health (mHealth) have become a cornerstone in oncology. Various clinical studies have demonstrated longer survival in patients treated by chemotherapy [5] and earlier detection of lung cancer recurrence [6]. mHealth offers numerous advantages: Patients tend to better understand the stakes of their treatments and gain knowledge on sentinel symptoms, and doctors may react faster. In addition, patients treated for cancer do not feel more anxious with new access to medical information and, quite contrarily, show increased satisfaction regarding their care [7].

A study conducted in England in 2015 showed that reaching smokers with persistent cough through a simple poster advertising campaign led to the diagnosis of 9% more lung cancers over the campaign period and an increase in good-prognosis, surgery-eligible stage I cancer diagnosis by 3 percentage points. The tagline was “been coughing for 3 weeks or more? Tell your doctor” [8]. Current and ex-smokers often lack information regarding tobacco-related diseases [9] and

frequently wait several months with symptoms before consulting a physician.

A dedicated self-assessment app provides regular analysis of symptoms in relation to tobacco consumption reported by patients on their smartphones and notifies them to consult their general practitioner (GP) in the event of suspicious symptoms. Knowledge about symptoms possibly indicating COPD, CVD, or lung diseases may help patients to reduce tobacco consumption or even quit smoking [10].

The purpose of this prospective study was to assess the incidence of the main symptoms associated with the leading tobacco-related diseases, evaluate the capacity of the app to encourage smokers and ex-smokers to quit smoking or reduce tobacco use, and evaluate its impact on the incidence of symptomatic surgery-eligible lung cancer.

## Methods

### Overview

Current and ex-smokers who quit smoking during the past 5 years were recruited via an advertising campaign on social media, newspapers, and public posters in the city of Le Mans and in Sarthe county (France). The campaign was held between June 3, 2019 and June 20, 2019. County GPs also received information on the app and were encouraged to propose it to their patients.

The self-assessment smartphone app was available for free download on Android and Apple app stores. After entering personal anonymous data (sex, age, tobacco consumption status, length and frequency of tobacco consumption), current smokers were asked whether they considered quitting or did not wish to at all. Both current and ex-smokers were then requested to answer 13 simple questions linked to symptoms (chest pain, chest tightness, cough, dyspnea, unusual tiredness, lower limb claudication, hematuria, hemoptysis, unintentional weight loss, dysphagia, dysphonia, palpable subcutaneous nodule, need to expectorate) possibly corresponding with the following diseases: cancer (lung, head and neck, esophagus, kidney, bladder), COPD, lower extremity arterial disease (LEAD), and angina. In the event of a positive answer to any of the mentioned points, a notification was sent recommending the user to consult his or her GP about the suspicious symptoms.

Symptoms were selected by a board of experts according to specific semiology of each disease [6,11].

In addition, once the questionnaire was filled out, current smokers were asked whether they considered reducing tobacco consumption or quitting. In addition, a toll-free telephone number was provided for patient counseling on smoking cessation therapy.

### Inclusion and Exclusion Criteria

For the purpose of this study, participants were included based on the following criteria: aged 16 years or older, being a current or former smoker (cessation within the past 5 years), informed consent on data use.

Exclusion criteria were a lack of consent and nonuse of the app.

### Ethical Approval

The study was approved by the French National Health Data Institute, which reviews ethical issues in human research, data confidentiality, and safety.

We focused on the diagnosis of cancer in the 2 counties as data were available through the registries of thoracic surgical oncology, unlike information regarding COPD or arteritis. In addition, surgery remains the main curative treatment for lung cancer.

### Statistical Analysis

We first assessed the number of users presenting with symptoms. We then evaluated the impact on smoking cessation intention after using the app and compared the rate of symptomatic lung cancer eligible for surgery in Sarthe county and neighboring Maine-et-Loire during the app roll-out period with the same period in 2018. Analysis of new cases started 12 weeks after launching the app to allow time for required examinations (eg, endoscopy, computed tomography [CT] scan, positron emission tomography scan). This evaluation was conducted on the basis

of private and public pathology laboratories' records in both Sarthe and Maine-et-Loire counties, thus providing 100% of documented lung cancer cases. Records from multidisciplinary team thoracic oncology meetings in both counties were analyzed to obtain the number of symptomatic lung cancers eligible for surgery.

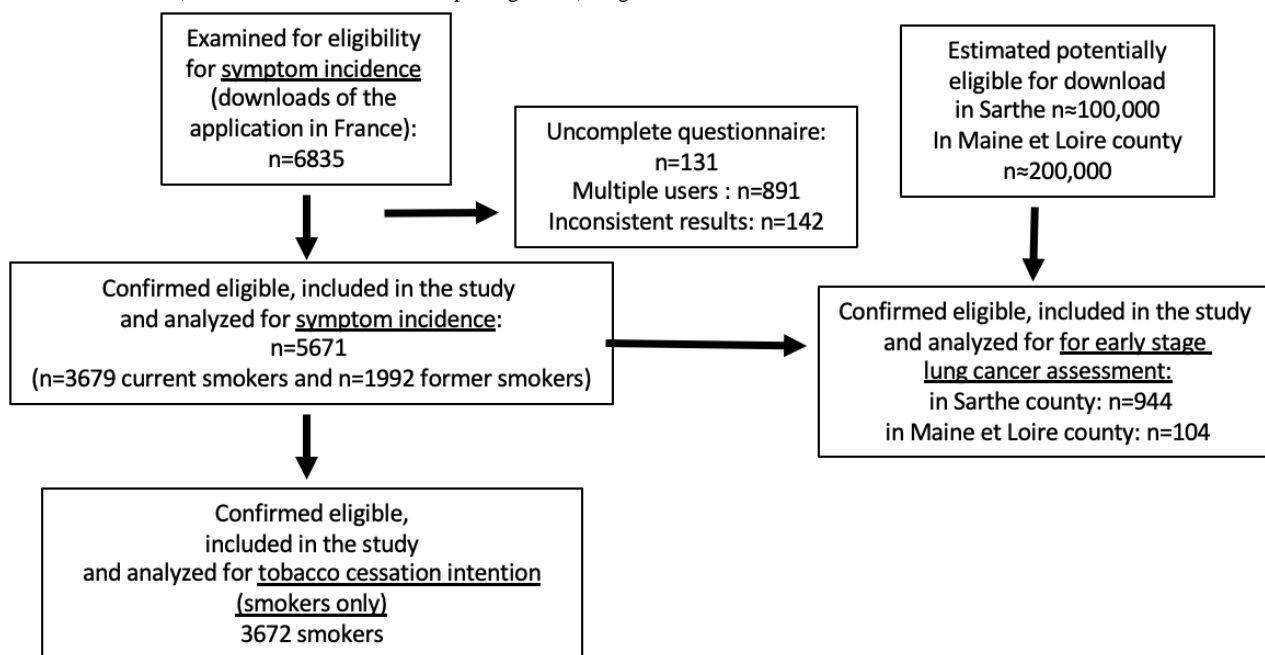
Fisher and Mac-Neymar tests were used for descriptive analyses, with  $P < .05$  considered to be statistically significant.

## Results

### Recruited Population

The app was downloaded 6835 times between June 3, 2019 and December 31, 2019, and 5671 users were eligible. Among them, 3679 (64.9%) were current smokers, and 1992 (35.1%) were ex-smokers. Median age was 39 years, 57.8% (3276/5671) were men, and 42.2% (2395/5671) were women. Regarding smoking history, 3425 (3425/5671, 60.4%) had a 20 pack-year smoking history, and 4450 (4450/5671, 78.5%) had been smoking for at least 20 years. The app was downloaded by 944 users in Sarthe county and 104 in neighboring Maine-et-Loire (Figure 1). In step 1, the baseline questions were asked of current smokers only ( $n=3679$ ). These questions asked about smoking cessation (answers = Yes/No), to which 627 answered "Yes" to "You don't want to stop smoking?" and 3052 answered "Yes" to "Or are you hesitating to stop?" In step 2, 13 questions about symptoms were asked of the entire study population ( $n=5671$ ). In Step 3, after completing the symptom questionnaire, the users were asked "At the end of the first use of Smokecheck, have you considered continuing your current consumption, without a reduction in consumption (Yes: 531/3672, 14.4%), reducing your consumption (Yes: 1796/3672, 49.0%), or total cessation of tobacco use (Yes: 1345/3672, 36.6%)?"

**Figure 1.** CONSORT (Consolidated Standards of Reporting Trials) diagram.



## Notifications

A notification was sent to a majority of users (4118/5671, 72.6%). Current smokers were notified significantly more often than ex-smokers (2833/3679, 77.0% vs 1298/1992, 65.2%;

$P < .001$ ). The most frequently detected symptoms were unusual tiredness (2023/5671, 35.7%), persistent cough (1658/5671, 29.2%), dyspnea (1502/5671, 26.5%), and persistent chest pain (1286/5671, 22.7%; [Table 1](#)).

**Table 1.** Symptoms self-assessed on a smartphone app by current or former smokers who stopped smoking within the past 5 years.

Symptoms that triggered a notification to the user <sup>a</sup>	Current smokers (n=3679), n (%)	Former smokers (n=1992), n (%)	P value
Have you noticed persistent, unusual tiredness that has lasted for 3 weeks?	1401 (38.1)	622 (31.2)	<.001
Have you been coughing for 3 weeks or more?	1269 (34.5)	389 (19.5)	<.001
Have you had, for 3 weeks or more, unusual shortness of breath when walking on flat terrain compared with someone of the same age?	1092 (29.7)	410 (20.6)	<.001
Do you have unusual shoulder pain or chest pain that has been persistent for 3 weeks or more?	870 (23.6)	416 (20.9)	.01
In the last 3 past weeks, have you had any pain or a squeezing sensation in your chest during effort?	571 (16)	222 (11.1)	<.001
In the previous few weeks, have you had pain in 1 or both legs while walking that has caused you to stop walking?	455 (12.4)	205 (10.3)	.02
Do you currently have a change in your voice that has lasted 3 weeks or more?	380 (10.3)	161 (8.1)	.01
Have you unintentionally lost 3 kg or more in the last 3 months?	325 (8.8)	123 (6.2)	<.001
Have you recently noticed an unusual lump under your skin?	303 (8.2)	155 (7.8)	.57
Have you had persistent difficulties swallowing food for 3 weeks or more?	126 (3.4)	77 (3.9)	.41
In the last 3 weeks, have you observed the presence of blood in your sputum?	82 (2.2)	49 (2.5)	.52
In the past 3 weeks, have you observed blood in your urine?	43 (1.2)	18 (0.9)	.42

<sup>a</sup>In the event of a positive answer to any of the questions, a notification was sent recommending the user to consult his or her general practitioner about the suspicious symptoms.

## Symptoms and Smoking Duration

We observed that smokers with a longer smoking history were more frequently notified.

Indeed, 78.0% (3213/4118) of the patients who received a notification had a smoking history of over 20 years, while 22.0% (905/4118) had less than 20 years of smoking history ( $P < .001$ ).

## Number of Symptoms

On average, users who received a notification had 1.93 symptoms. Current smokers had 2.13 symptoms, whereas former smokers had 1.57 symptoms ( $P < .001$ ).

## Tobacco-Induced Diseases

Several symptoms, isolated or in association, may suggest tobacco-induced diseases accessible to specific screening.

Associated symptoms that may be indicators of COPD, such as a persistent cough for 3 weeks with dyspnea or a need to expectorate, was reported by 14% (515/3679) of current smokers as well as 7.5% (150/1992) of ex-smokers ( $P < .001$ ).

Lower-limb pain suggestive of LEAD was found in 12.4% (455/3679) of notified current smokers and 10.3% (205/1992) of notified ex-smokers ( $P = .02$ ).

Of current smokers who received a notification, 15.5% (571/3679) presented with symptoms suspicious of stable angina, compared with 11.1% (222/1992) of ex-smokers ( $P < .001$ ).

Of current smokers, 6.8% (249/3679) were notified for symptoms suggesting pulmonary, urinary, or head-and-neck cancer, such as hemoptysis, hematuria, and unintentional weight loss associated with chest pain or dysphagia, as were 5.9% (117/1992) of former smokers ( $P = .20$ ).

## More Specific Symptoms

Asthenia associated with another symptom may suggest tobacco-induced complications, but it cannot be considered as specifically linked to tobacco consumption when isolated. Provided that the app also recorded notifications to patients who had only checked this symptom in the app, we removed these data and then noticed a lower number of notifications, amounting to 3784, representing 66.7% (3784/5671 vs 4118/5671, 72.6%), with 71.1% (2615/3679) of current smokers and 58.0% (1157/1992) of ex-smokers notified.

## Cessation Intention

After using the app, 36.5% (1343/3679) of users considered quitting smoking, 48.7% (1795/3679) considered reducing their consumption, and 14.4% (530/3679) did not plan any change.

Before answering the 13-item questionnaire, 627 of 3679 current smokers (17.0%) had declared that they did not wish to quit smoking. However, 5.7% (36/627) changed their mind and declared they wanted to quit after completing the questionnaire ( $P < .001$ ; [Table 2](#)).

Among the 3052 smokers who hesitated to quit before answering the 13 questions, 42.8% (1306/3052) then declared their

intention to quit, and 48.5% (1480/3052) declared an intent to reduce their consumption.

After completing the questionnaire, 1490 of the 1949 smokers who had declared that they wanted to reduce their consumption had received a notification (76.4%) and so did 1368 of the 1772 (77.2%) who declared that they wanted to quit.

**Table 2.** Consideration of smoking cessation as reported by 3679 current smokers before the symptom questionnaire intervention and subsequent intention for smoking reduction or cessation after completing the symptom questionnaire in the app.

Status of current smokers at baseline	Smoking intention after app use, n (%)			P value <sup>a</sup>
	No smoking change	Smoking reduction	Smoking cessation	
Would not consider cessation (n=627)	269 (42.9)	316 (50.4)	36 (5.7)	<.001
<b>Alert status for those who would not consider cessation</b>				
Alert received (n=470)	201 (42.8)	235 (50.0)	29 (6.2)	<sub>b</sub>
No alert received (n=157)	68 (43.3)	81 (51.6)	7 (4.5)	-
Would consider cessation (n=3052)	261 (8.6)	1479 (48.5)	1307 (42.8)	<.001
<b>Alert status for those who would consider cessation</b>				
Alert received (n=2361)	202 (8.6)	1138 (48.2)	1019 (43.2)	-
No alert received (n=689)	59 (8.6)	341 (49.5)	288 (41.8)	-

<sup>a</sup>Assessed using the Mac-Neymar test.

<sup>b</sup>Comparisons not conducted.

### Incidence of Symptomatic Lung Cancers Eligible for Surgery

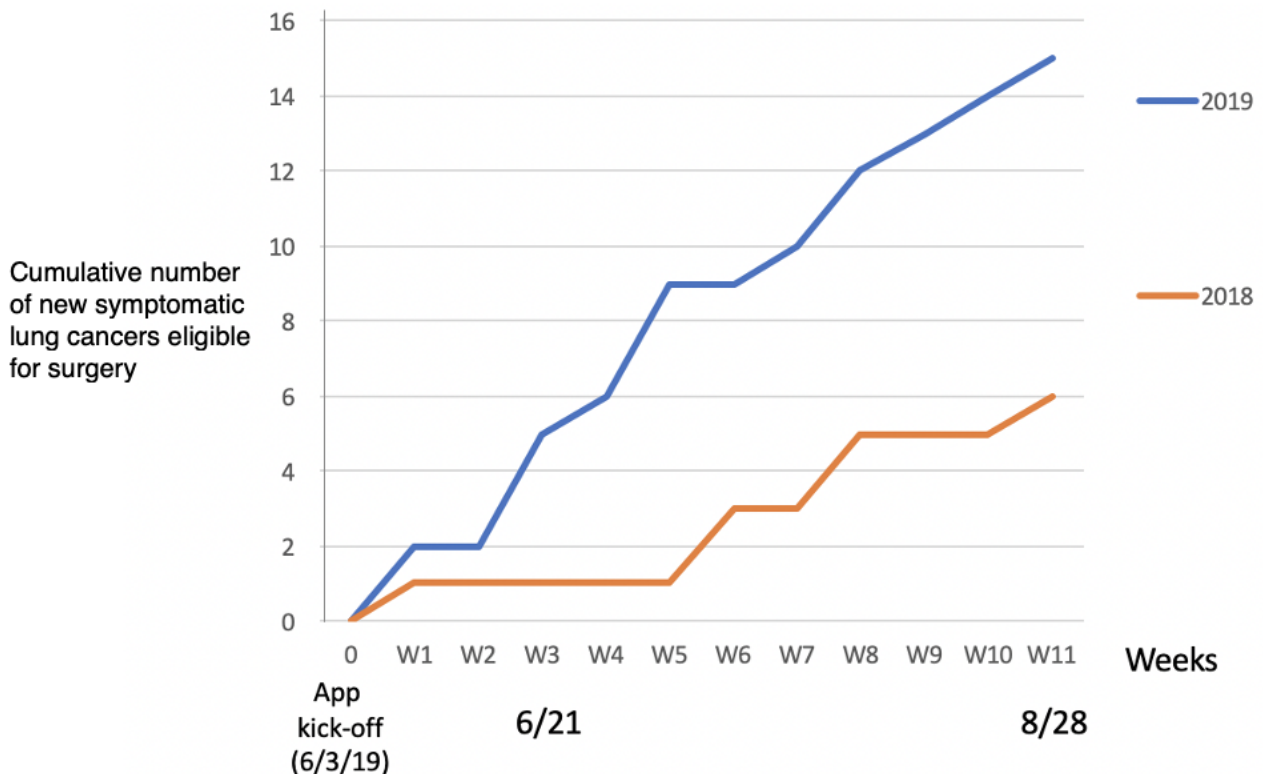
#### In Sarthe County (944 Downloads)

Between June 21, 2018 and August 21, 2018, 54 new cases of lung cancer were diagnosed, whereas 58 new cases were

diagnosed over the same period in 2019 (ie, 12 weeks after launching the app in Sarthe county).

The number of symptomatic lung cancers eligible for surgery increased from 9% (5/54) during that period in 2018 to 24% (14/58) over the same period in 2019 ( $P=.004$ ; [Figure 2](#), [Table 3](#)).

**Figure 2.** Cumulative number of new symptomatic lung cancers eligible for surgery during the 12 weeks after the application kick-off in 2019 (assessed between June 21, 2019 and August 28, 2019) and during the same period without application availability in Sarthe in 2018.



**Table 3.** Patients with symptomatic lung cancer characteristics between June 21, 2018 and August 28, 2018 and after the app kick-off in Sarthe county during the same period in 2019 (kick-off on June 3, 2019).

Year and patient number	Age (years)	Symptom(s)	TNM <sup>a</sup>	Stage	Histology	Surgery	Performed	pTNM <sup>b</sup>
<b>2018</b>								
Patient 1	64	Persistent cough	T2N1M0	2	ADK <sup>c</sup>	Yes	Yes	pT2pN1
Patient 2	70	Chest pain	T2N0M0	1	ADK	Yes	Yes	pT1pN0
Patient 3	64	Dyspnea	T1N0M0	1	ADK	Yes	Yes	pT1pN0
Patient 4	64	Cough	T2N0M0	1	ADK	Yes	Yes	pT3pN2
Patient 5	69	Cough asthenia	T2N1M0	2	SCC <sup>d</sup>	Yes	Yes	pT3pN1
<b>2019</b>								
Patient 1	59	Persistent cough	T2N0M0	1	ADK	Yes	Yes	pT2pN2
Patient 2	60	Chest pain	T1N0M0	1	K sarc <sup>e</sup>	Yes	Yes	pT3pN2
Patient 3	88	Persistent cough	T1N0M0	1	ADK	Yes	Yes	pT1pN0
Patient 4	51	Cough and hemoptysis	T1N0M0	1	Carcinoid	Yes	Yes	pT1pN2
Patient 5	51	Persistent cough and asthenia	T3N0M0	2	SCC	Yes	Yes	pT2pN1
Patient 6	54	Persistent cough and asthenia	T2N0M0	1	SCC	Yes	No	No surgery
Patient 7	64	Dyspnea	T2N0M0	1	ADK	Yes	Yes	pT2pN1
Patient 8	53	Lump under skin	T1N0M0	1	ADK	Yes	Yes	pT3pN0
Patient 9	59	Persistent cough	T1N0M0	1	ADK	Yes	Yes	pT1pN1
Patient 10	29	Persistent cough	T1N0M0	1	Carcinoid	Yes	Yes	pT1pN0
Patient 11	66	Chest pain and weight loss	T3N0M0	2	SCC	Yes	Yes	pT4pN1
Patient 12	75	Persistent cough and weight loss	T1N0M0	1	ADK	Yes	Yes	pT2pN0
Patient 13 <sup>f</sup>	46	Pain in one leg	T1N0M0	1	ADK	Yes	Yes	pT1pN0
Patient 14 <sup>f</sup>	58	Pain in one leg	T1N0M0	1	ADK	Yes	Yes	pT1pN0

<sup>a</sup>TNM: staging classification that describes the tumor (T), node (N), and metastasis (M) categories.

<sup>b</sup>pTNM: pathological tumor-node-metastasis.

<sup>c</sup>ADK: adenocarcinoma.

<sup>d</sup>SCC: squamous cell carcinoma.

<sup>e</sup>K sarc: sarcomatoid carcinoma.

<sup>f</sup>Experienced leg claudication that led to a visit.

### ***In Neighboring Maine-et-Loire County (104 Downloads)***

The number of symptomatic lung cancers eligible for surgery was stable between 2018 (10/74, 14%) and 2019 (12/78, 15%) in this county, where the app was not deployed ( $P=.82$ ).

## ***Discussion***

### **Principal Findings**

#### ***Representative Population and Important Use of the App Despite an Advertising Campaign Limited in Time and Space***

To the extent of our knowledge, this is the first prospective real-life study with a self-assessment app aiming to detect symptoms of smokers and ex-smokers in order to encourage them to quit smoking and to consult their physician for early detection of tobacco-induced diseases, such as lung cancer.

Despite a local campaign in a single county (Sarthe) over a limited 2-week period, the app was downloaded nearly 7000 times, 80% of which occurred in other counties. Social networks and word of mouth among users' families certainly spread the use of the app further. This strong participation allowed it to quickly reach a high number of users with clinical features similar to those of smokers in other studies. Indeed, the median age of smokers and ex-smokers, who had quit within the past 5 years, was 39 years, with 42.2% (2395/5671) of our study population as women. Hajek et al [12] randomized e-cigarettes and nicotine replacement therapy, with a median age of 41 years and 48% women.

#### ***Frequency of Symptoms and Diseases***

A majority of users received a notification (72.6%), especially among current smokers (77.0% vs 64.6% of ex-smokers).



A significant number of smokers presented with symptoms or associations of symptoms suggesting tobacco-induced diseases: 14.0% with COPD, 15.5% with stable angina, 12.4% with potential LEAD, and 6.8% with potential cancer. Number and incidence of symptoms indicating these diseases appeared to be lower after recent cessation (<5 years), which contributes to showing that smoking cessation can help to prevent them.

### **Smoking Cessation Intention**

After using the app, a significant number of current smokers declared their intention to reduce their tobacco consumption (1795/3679, 48.7%), and 36.5% (1343/3679) declared their intention to quit altogether.

Among those who declared an intention to reduce consumption after filling out the questionnaire, 76.4% received a notification through the app, and 77.2% of those who stated that they wanted to quit were notified.

Even though the questionnaire only collected declarations of intent, which cannot be easily verified in practice, triggered notifications brought to light some concerns, which seemed to prompt smokers to reduce consumption or even quit smoking. This behavior change elicited by raised awareness about tobacco-induced symptoms is a new personalized approach, as smokers are directly and objectively facing the abnormality of such symptoms. These can lead to smoking cessation, as has been observed among 42.8% of the smokers who claimed to be hesitant about quitting (representing 83% of all smokers in this study).

### **Cancer Screening**

In the United States, the US Preventive Services Task Force established guidelines for lung cancer screening. Patients have to be over the age of 55 years, have a 30-pack year history of smoking, and be a current or former smoker [13], but our app is not a “lung cancer screening” tool as per US specificities but actually aims to refer users to their physician in case of suspicious symptoms. It therefore returns an alert, independent of the user’s age or smoking history. The launch of the app in June 2019 was linked to a significant increase in the number of symptomatic surgery-eligible lung cancers within the 12 weeks following the start of the advertising campaign and the no-cost availability of the app in Sarthe county where it was downloaded 944 times. Over the same period in 2018, only 9% of lung cancers diagnosed based on symptoms were eligible for surgery. In the neighboring Maine-et-Loire county, where the app was only downloaded 104 times, the number of symptomatic cancers diagnosed at an operable stage had not increased significantly (14% between June 21, 2018 and August 28, 2018; 15% between the same dates in 2019). These figures suggest a positive correlation between the availability of the app (as well as its download) and the increase in detections of symptomatic lung cancers eligible for surgery.

### **Limitations**

This study had several limitations. First, there was no control group. Second, the app is a self-assessment app in which

information is self-reported, rather than a patient-reported outcomes app that reports data to physicians. Thus, data were left unverified, whether regarding symptoms or smoking cessation intentions. Several symptoms triggering notifications are not very specific of tobacco-induced diseases when isolated and thus raise the number of user alerts. By removing isolated asthenia from all notifications (n=334), we reached a rate of notification of 66.7% (vs 72.6%), and by removing isolated subcutaneous nodules (n=79), we reached 65.3%. These 2 symptoms, when recorded as isolated symptoms, could be removed from further versions of the app, in order to be more specific in detecting tobacco-induced complications. This may also prevent the difficult absorption by physicians of the numerous additional consultations expected after nationwide coverage.

In order to evaluate the impact of the app on the detection of tobacco-induced diseases, a questionnaire could have been sent to users a few weeks after receiving the alert encouraging to consult a doctor. The app was not designed with this characteristic, which could be grounds for further study. We decided to focus on lung cancer and chose the eligibility for curative surgery as our primary outcome.

The increase in the incidence of surgery-eligible cancers was not assessed through direct contact with users, as this study was based on health data requiring full anonymity. Evaluation was indirectly conducted, by triangulating oncology and pathology records, which gather nearly 100% of cases in the concerned counties (Sarthe and Maine-et-Loire), as well as data from app downloads. This result would have been more relevant if a direct link between the use of the app and cancer diagnosis had been established, but the chosen methodology could not allow it. GPs’ awareness-raising activities and the advertising campaign itself also may have contributed to the early detection of these cancers.

### **Generalizability**

Among smokers and ex-smokers, symptoms related to tobacco-induced diseases are frequent. The app is a self-assessment mHealth tool that could encourage GP consults whenever suspicious symptoms appear, setting up strong prevention dynamics in total safety as opposed to some use of e-cigarettes. Physicians could also implement earlier treatments against diseases such as COPD, which is a significant risk factor for cancer and the third cause of mortality in France, as well as coronary and other arterial diseases, early treatment of which also provide major benefits. The early detection and management of these diseases may lead to reduced cancer incidence and support the preventive role of this app against lung cancer. Finally, this app may be a relevant tool to prompt symptomatic smokers to schedule a thoracic CT, with the hope to foster participation in this modality of cancer screening, whose benefit on survival rates was proven despite weak participation in real life [14,15].

This exploratory study warrants a follow-up randomized study to evaluate the impact of this tool on early lung cancer screening.

## Acknowledgments

We thank the users for their participation in this study as well as Magali Balavoine, MSc, from Weprom in Angers, France. No one received compensation for their contributions. The sponsor (Weprom) designed and conducted the study; collected, managed, analyzed, and interpreted the data; prepared, reviewed, or approved the manuscript; and made the decision to submit the manuscript for publication.

Funding sources for this study include “Le Mans Métropole,” “Conseil Régional des Pays de la Loire,” “Conseil Départemental de la Sarthe,” “ELSAN,” and “Caisse Primaire d’Assurance Maladie of Sarthe.” None of these had any role in the design of the study; the collection, analysis, or interpretation of the data; the writing of the manuscript; or the decision to submit the manuscript for publication.

## Conflicts of Interest

TU reports receiving personal fees from AstraZeneca and Novartis. FD reports receiving personal fees from AstraZeneca, Ipsen, Kelindi, Pfizer, Chugai, and Roche.

## References

1. Cancer. World Health Organization. 2018. URL: <https://www.who.int/news-room/fact-sheets/detail/cancer> [accessed 2022-02-05]
2. Tumeur maligne, affection maligne du tissu lymphatique ou hématopoïétique. Haute Autorité de Santé. 2013. URL: <https://tinyurl.com/33brhxtc> [accessed 2019-09-06]
3. Connaissance de la bronchopneumopathie chronique obstructive (BPCO) en France : Baromètre santé 2017. Santé publique France. 2019 Sep 6. URL: <https://tinyurl.com/4v6f9epb> [accessed 2022-02-05]
4. Gomajee R, El-Khoury F, Goldberg M, Zins M, Lemogne C, Wiernik E, et al. Association between electronic cigarette use and smoking reduction in France. *JAMA Intern Med* 2019 Sep 01;179(9):1193-1200 [FREE Full text] [doi: [10.1001/jamainternmed.2019.1483](https://doi.org/10.1001/jamainternmed.2019.1483)] [Medline: [31305860](https://pubmed.ncbi.nlm.nih.gov/31305860/)]
5. Basch E, Deal AM, Dueck AC, Scher HI, Kris MG, Hudis C, et al. Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. *JAMA* 2017 Jul 11;318(2):197-198 [FREE Full text] [doi: [10.1001/jama.2017.7156](https://doi.org/10.1001/jama.2017.7156)] [Medline: [28586821](https://pubmed.ncbi.nlm.nih.gov/28586821/)]
6. Denis F, Basch E, Septans A, Bennouna J, Urban T, Dueck AC, et al. Two-year survival comparing web-based symptom monitoring vs routine surveillance following treatment for lung cancer. *JAMA* 2019 Jan 22;321(3):306-307 [FREE Full text] [doi: [10.1001/jama.2018.18085](https://doi.org/10.1001/jama.2018.18085)] [Medline: [30667494](https://pubmed.ncbi.nlm.nih.gov/30667494/)]
7. Gravis G, Protière C, Eisinger F, Boher JM, Tarpin C, Coso D, et al. Full access to medical records does not modify anxiety in cancer patients: results of a randomized study. *Cancer* 2011 Oct 15;117(20):4796-4804 [FREE Full text] [doi: [10.1002/cncr.26083](https://doi.org/10.1002/cncr.26083)] [Medline: [21607939](https://pubmed.ncbi.nlm.nih.gov/21607939/)]
8. Moffat J, Bentley A, Ironmonger L, Boughey A, Radford G, Duffy S. The impact of national cancer awareness campaigns for bowel and lung cancer symptoms on sociodemographic inequalities in immediate key symptom awareness and GP attendances. *Br J Cancer* 2015 Mar 31;112 Suppl 1:S14-S21 [FREE Full text] [doi: [10.1038/bjc.2015.31](https://doi.org/10.1038/bjc.2015.31)] [Medline: [25734383](https://pubmed.ncbi.nlm.nih.gov/25734383/)]
9. Global briefing: symptom awareness and attitudes to lung cancer Findings from a global study. Global Lung Cancer Coalition. 2021 Mar. URL: <http://www.lungcancercoalition.org/uploads/docs/Global%20briefing%20-%20FINAL.pdf> [accessed 2022-02-05]
10. Ironmonger L, Ohuma E, Ormiston-Smith N, Gildea C, Thomson CS, Peake MD. An evaluation of the impact of large-scale interventions to raise public awareness of a lung cancer symptom. *Br J Cancer* 2015 Jan 06;112(1):207-216 [FREE Full text] [doi: [10.1038/bjc.2014.596](https://doi.org/10.1038/bjc.2014.596)] [Medline: [25461805](https://pubmed.ncbi.nlm.nih.gov/25461805/)]
11. Guillemin L. *Sémiologie médicale*, 2ème édition. Cachan, Ile de France: Lavoisier MSP; 2011.
12. Hajek P, Phillips-Waller A, Przulj D, Pesola F, Myers Smith K, Bisal N, et al. A randomized trial of e-cigarettes versus nicotine-replacement therapy. *N Engl J Med* 2019 Feb 14;380(7):629-637. [doi: [10.1056/nejmoa1808779](https://doi.org/10.1056/nejmoa1808779)]
13. Final Recommendation Statement: Lung Cancer: Screening. U.S. Preventive Services Task Force. 2021 Mar 09. URL: <https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/lung-cancer-screening> [accessed 2022-02-05]
14. Tanoue L. Reduced lung-cancer mortality with low-dose computed tomographic screening. *Yearbook of Medicine* 2012 Jan;2012:259-261 [FREE Full text] [doi: [10.1016/s0084-3873\(12\)00222-2](https://doi.org/10.1016/s0084-3873(12)00222-2)]
15. de Koning HJ, van der Aalst CM, de Jong PA, Scholten ET, Nackaerts K, Heuvelmans MA, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial. *N Engl J Med* 2020 Feb 06;382(6):503-513. [doi: [10.1056/nejmoa1911793](https://doi.org/10.1056/nejmoa1911793)]

## Abbreviations

**COPD:** chronic obstructive pulmonary disease

**CT:** computed tomography

**CVD:** cardiovascular disease  
**GP:** general practitioner  
**LEAD:** lower extremity artery disease  
**mHealth:** mobile health

*Edited by G Eysenbach, R Kukafka; submitted 05.05.20; peer-reviewed by A Rojewski, P Zarogoulidis; comments to author 29.06.20; revised version received 21.11.20; accepted 02.08.21; published 23.02.22.*

*Please cite as:*

*Stavaux E, Goupil F, Barreau G, Septans AL, Dautzenberg B, Foulet-Rogé A, Padilla N, Urban T, Denis F  
Use of a Smartphone Self-assessment App for a Tobacco-Induced Disease (COPD, Cardiovascular Diseases, Cancer) Screening  
Strategy and to Encourage Smoking Cessation: Observational Study*

*JMIR Public Health Surveill 2022;8(2):e19877*

*URL: <https://publichealth.jmir.org/2022/2/e19877>*

*doi: [10.2196/19877](https://doi.org/10.2196/19877)*

*PMID: [35195530](https://pubmed.ncbi.nlm.nih.gov/35195530/)*

©Edouard Stavaux, François Goupil, Guillaume Barreau, Anne Lise Septans, Bertrand Dautzenberg, Armelle Foulet-Rogé, Norbert Padilla, Thierry Urban, Fabrice Denis. Originally published in JMIR Public Health and Surveillance (<https://publichealth.jmir.org>), 23.02.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

# Public Reactions to the New York State Policy on Flavored Electronic Cigarettes on Twitter: Observational Study

Li Sun<sup>1</sup>, BSc, BA; Xinyi Lu<sup>1</sup>, BSc, BA; Zidian Xie<sup>2</sup>, PhD; Dongmei Li<sup>2</sup>, PhD

<sup>1</sup>Goergen Institute for Data Science, University of Rochester, Rochester, NY, United States

<sup>2</sup>Department of Clinical & Translational Research, University of Rochester Medical Center, Rochester, NY, United States

**Corresponding Author:**

Dongmei Li, PhD

Department of Clinical & Translational Research

University of Rochester Medical Center

Saunders Research Building 1.303J

265 Crittenden Boulevard CU 420708

Rochester, NY, 14642-0708

United States

Phone: 1 5852767285

Email: [Dongmei\\_Li@urmc.rochester.edu](mailto:Dongmei_Li@urmc.rochester.edu)

## Abstract

**Background:** Flavored electronic cigarettes (e-cigarettes) have become popular in recent years, especially among youth and young adults. To address the epidemic of e-cigarettes, New York State approved a ban on sales of most flavored vaping products other than tobacco and menthol flavors on September 17, 2019.

**Objective:** This study aims to examine the attitude of Twitter users to the policy on flavored e-cigarettes in New York State and the impact of this policy on public perceptions of e-cigarettes. This study also compares the attitudes and topics between New York Twitter users and Twitter users from other states who were not directly affected by this policy.

**Methods:** Tweets related to e-cigarettes and the New York State policy on flavored e-cigarettes were collected using the Twitter streaming application programming interface from June 2019 to December 2019. Tweets from New York State and those from other states that did not have a flavored e-cigarette policy were extracted. Sentiment analysis was applied to analyze the proportion of negative and positive tweets about e-cigarettes or the flavor policy. Topic modeling was applied to e-cigarette-related data sets and New York flavor policy-related data sets to identify the most frequent topics before and after the announcement of the New York State policy.

**Results:** We found that the average number of tweets related to e-cigarettes and the New York State policy on flavored e-cigarettes increased in both New York State and other states after the flavor policy announcement. Sentiment analysis revealed that after the announcement of the New York State flavor policy, in both New York State and other states, the proportion of negative tweets on e-cigarettes increased from 34.07% (4531/13,299) to 44.58% (18,451/41,390) and from 32.48% (14,320/44,090) to 44.40% (64,262/144,734), respectively, while positive tweets decreased significantly from 39.03% (5191/13,299) to 32.86% (13,601/41,390) and from 42.78% (18,863/44,090) to 33.93% (49,105/144,734), respectively. The majority of tweets related to the New York State flavor policy were negative both before and after the announcement of this policy in both New York (87/98, 89% and 3810/4565, 83.46%, respectively) and other states (200/255, 78.4% and 12,695/15,569, 81.54%, respectively), while New York State had a higher proportion of negative tweets than other states. Topic modeling results demonstrated that teenage vaping and health problems were the most discussed topics associated with e-cigarettes.

**Conclusions:** Public attitudes toward e-cigarettes became more negative on Twitter after New York State announced the policy on flavored e-cigarettes. Twitter users in other states that did not have such a policy on flavored e-cigarettes paid close attention to the New York State flavor policy. This study provides some valuable information about the potential impact of the flavored e-cigarettes policy in New York State on public attitudes toward flavored e-cigarettes.

(*JMIR Public Health Surveill* 2022;8(2):e25216) doi:[10.2196/25216](https://doi.org/10.2196/25216)

**KEYWORDS**

New York State policy; flavored e-cigarettes; Twitter; social media; vaping; e-cigarette

## Introduction

Tobacco smoking is a well-known risk factor for many diseases such as heart disease, cancer, and pulmonary disease [1]. Since flavored tobacco products attract people by hiding the natural harshness and taste of tobacco, the US Food and Drug Administration banned the sale of candy- and fruit-flavored cigarettes in 2009 [1,2]. Cigarette smoking among youth has declined in recent years, but the usage of electronic cigarettes (e-cigarettes) especially among youth has increased dramatically in recent several years [3]. Since 2014, e-cigarettes have become the most commonly used tobacco product among the youth [4]. The Centers for Disease Control and Prevention has reported that between 2011 and 2015, the usage of e-cigarettes has increased by more than 800% among middle school and high school students [5]. The 2019 National Youth Tobacco Survey data has shown that 27.5% of high school students and 10.5% of middle school students are currently using e-cigarettes [4].

Similar to flavored cigarette, flavored e-cigarette attracts people by its affordability, accessibility, convenience, and more importantly, a variety of flavors [6]. A study based on Population Assessment of Tobacco and Health Study (PATH) Wave 3 data showed that the most popular e-cigarette flavors are fruit and candy [7]. However, recent studies showed that flavored e-cigarettes could be harmful to lung tissues by imposing oxidative stress and inflammatory responses [8]. It is well known that e-cigarettes release volatile carbonyls, furans, nickels, lead, and chromium, which may be poisonous to the lungs [9]. In addition, e-cigarettes could harm endothelial cells that line the interior of human blood vessels and may increase the risk of heart disease [10]. The number of reported cases of e-cigarette or vaping use-associated lung injury (EVALI) increased rapidly in the United States in 2019 [11]. As of October 8, 2019, there were 1299 cases reported to the Centers for Disease Control and Prevention [12], and as of January 14, 2020, the number increased to 2668 [11]. Among patients with EVALI, 76% are younger than 35 years [11].

Owing to the potential negative health effects of flavored e-cigarettes, starting from June 2019, many states and cities in the United States have announced the ban on flavored e-cigarettes. On June 25, 2019, San Francisco became the first US city to ban the sale and distribution of e-cigarettes in the city [13]. Michigan (starting from September 4, 2019) and New York (starting from September 17, 2019) announced the policy regulating the sales of most flavored vaping products [14,15]. Following New York State, Rhode Island, Los Angeles County, Oregon State, Montana State, Washington State, New Jersey, and Massachusetts passed the ban on the sale of flavored vaping products [15-22].

Many recent e-cigarettes studies have utilized social media data to identify topics related to e-cigarettes. For example, Kavuluru and Sabbir [23] developed a supervised predictive model to identify e-cigarette proponents on Twitter. Zhou et al [24] investigated the influence of flavors on the propagation of e-cigarette-related information on Facebook. As one of the most popular social media platforms in the United States, Twitter contains many e-cigarette-related posts (tweets), which provides

us an ideal avenue to investigate the public opinion on the policies regulating flavored e-cigarettes. In addition, messages from social media can influence people's attitudes and behaviors [25]. Pew Research Center found that approximately 20% of social media users might change their opinions after they view related messages on social media [25]. According to Pew Research Center data in 2019, compared to the whole population, Twitter users are younger, which corresponds to the potential users of e-cigarettes [26]. Compared to national surveys such as the PATH studies that have been used to study public opinions on tobacco products, social media studies could provide immediate reactions to policy shifts, larger sample size, much less data collection cost, and less biased responses [27].

Previous studies have investigated public attitudes toward e-cigarette regulation and policy by using sentiment analysis. The results showed that regulation was considered as a fundamental requirement for public health protection [28]. Instead of supporting a blanket ban on public vaping due to the perceptions of insufficient evidence on the harm of vaping, the study participants supported the right for individuals and organizations to restrict vaping [28]. The public attitudes toward the health policy and especially e-cigarettes will potentially affect user behavior, which is the primary goal of these health policies. Since flavored e-cigarettes policy was announced in the United States, it was important to understand the public attitudes toward the policy and how the public might react to the policy.

In this study, we aimed to investigate public responses toward the New York State flavor policy on Twitter by applying sentiment analysis and topic modeling to related tweets before and after the announcement of the policy. Furthermore, we compared the sentiments and topics from New York State and other states that did not have a flavored e-cigarette policy to examine the potential impact of the flavor policy in New York State on public attitudes toward e-cigarettes. We hypothesized that the public attitude toward flavored e-cigarettes would become more negative after the announcement of the flavor policy in both New York State and other states.

## Methods

### Data Collection and Preprocessing

e-Cigarette-related Twitter posts between June 2019 and December 2019 were downloaded from the Twitter streaming application programming interface by using e-cigarette-related keywords. The e-cigarette-related keywords include e-cig, e-cigs, ecig, ecigs, electroniccigarette, ecigarette, ecigarettes, vape, vapers, vaping, vapes, e-liquid, ejuice, eliquid, e-juice, vapercon, vapeon, vafefam, vaperation, and juul [29-31]. To avoid the potential impact of related flavor ban information posted on Twitter right before the announcement of the New York State flavor policy on September 17, 2019, we excluded the tweets posted from September 1 to September 16, 2019. Meanwhile, to avoid the potential overreaction of e-cigarette users to the New York State flavor policy immediately after its announcement and to examine more logical responses to the New York State flavor policy, we excluded tweets posted from September 17 to September 30, 2019 in our study.

To remove the promotion tweets, we filtered out Twitter IDs that contained promotion-related keywords (such as dealer, deal, store, supply, e-cig, store, promo, and promotion) [31]. In addition, we filtered out Twitter posts that contained promotion-related keywords (dealer, deal, customer, promotion, promo, promos, discount, sale, free shipping, sell, \$, %, dollar, offer, percent off, store, save, price, wholesale) [31]. After these 2 filtering steps, 2 data sets with e-cigarette-related tweets were created based on the posted date. One data set included e-cigarette tweets between June 13, 2019 and August 22, 2019, which is considered as before New York State announced the policy on flavored vaping products. The other data set was between October 1, 2019 and December 31, 2019, which is regarded as after New York State announced the policy. The first data set contained 724,345 e-cigarette tweets, and the second one contained 2,130,748 e-cigarette tweets. The number of unique users was identified based on the Twitter user ID. The first data set contained 599,146 unique users, and the second one contained 680,967 unique users.

We further extracted tweets related to the New York State policy on flavored e-cigarettes by filtering with keywords “ban” and “bans.” To ensure that these tweets were about the policy in New York State, we eliminated the tweets that mentioned other states with a ban on flavored e-cigarettes but not mentioned New York State. In total, we collected 68,318 New York State flavor policy-related tweets from June 2019 to December 2019, which included 353 before the policy and 20,134 tweets after the announcement of the New York flavor policy.

For both e-cigarette-related and New York flavor policy-related data sets, we applied 2 state-filtering processes on the geotagged tweets or the users who included the location information in the profile metadata to derive a New York State subset and other states subset (without a ban on flavored e-cigarettes) as the control group. First, we filtered the data sets by keywords “ny” and “new york” on the location of the user and the place of the tweets, which is the New York State subset. Second, we used the same procedure to filter the data sets with keywords “usa,” “united states,” and “us” to extract US tweets, and then, we eliminated the tweets from San Francisco, Michigan, New York State, Rhode Island, Los Angeles County, Oregon State, Montana State, Washington State, New Jersey, and Massachusetts that have policies on flavored e-cigarettes, which is the data set for other states.

### Sentiment Analysis

The Valence Aware Dictionary and sEntiment Reasoner was used as the sentiment analyzer to extract Twitter users' opinions on e-cigarettes and New York State flavor policy [32]. For both e-cigarette-related data sets and New York State flavor policy data set, we calculated the sentiment scores for each tweet.

Tweets with sentiment scores between  $-1.00$  and  $-0.05$  were classified as negative tweets, tweets with scores between  $-0.05$  and  $+0.05$  (not including  $-0.05$  and  $0.05$ ) were classified as neutral tweets, and tweets with scores between  $+0.05$  and  $+1.00$  were classified as positive tweets. To better compare the differences between different states and periods, we normalized the number of negative, neutral, and positive tweets by the total number of tweets in different states in each period. Finally, we conducted 2-sided 2-proportion z-tests to test whether the proportions of negative and positive tweets between New York State and other states were significantly different.

### Topic Modeling

The Latent Dirichlet Allocation model was conducted on e-cigarette-related data sets to extract the most frequently discussed topics. Latent Dirichlet Allocation, a type of topic modeling algorithm, is an unsupervised learning model that gives the number of topics, assigns each word in the document to a specific topic, and calculates a weight for each word in every topic [33]. First, after removing all punctuations and converting all texts to lowercase, we tokenized every sentence. Second, we applied the Natural Language Toolkit package to remove stop words (eg, the, a, in). Third, we used the Genism package to convert some frequent bigrams and trigrams to a single term. At last, we lemmatized all texts by implementing spaCy by changing all tenses to present tense and keeping only nouns, adjectives, verbs, and adverbs. To identify the optimal number of topics, we calculated the coherence scores that measure the relative distance between words within a topic. The number of topics was chosen based on the coherence scores. We selected the number of topics based on the maximum coherence score.

## Results

### Tweets Related to e-Cigarettes and the New York Flavor Policy on e-Cigarettes

We observed that in New York State or other states without any flavor policy, there was a significant increase in the percentage of daily tweets related to the flavor policy after the announcement of the New York flavor policy (Table 1). For the tweets related to the flavor policy, 0.74% (98/13,299) of all e-cigarette tweets were from New York State and 0.58% (256/44,090) of all e-cigarette tweets were from other states before the New York State flavor policy was announced. After the New York State flavor policy was announced, 10.93% (4565/41,764) of all e-cigarette tweets were related to the flavor policy from New York State and 11.11% (16,083/144,734) of all e-cigarette tweets were related to the flavor policy from other states.

**Table 1.** Proportion of tweets related to the New York flavor policy before and after its announcement in New York State and other states.

Time, state	Tweets, n (%)
<b>Before the New York flavor policy</b>	
In New York State (n=13,299)	98 (0.74)
In other states (n=44,090)	256 (0.58)
<b>After the New York flavor policy</b>	
In New York State (n=41,764)	4565 (10.93)
In other states (n=144,734)	16,083 (11.11)

### Public Attitudes Toward e-Cigarettes on Twitter

To examine whether there was any change in the public opinions toward e-cigarettes with the announcement of the New York flavor policy on e-cigarettes, we compared the proportions of negative, positive, and neutral tweets between before and after the announcement of the New York flavor policy in New York State and other states (Table 2 and Table 3, respectively). We observed that in both New York State and other states, compared to the period before the announcement of the New York flavor policy, the proportion of positive tweets on e-cigarettes significantly decreased ( $P<.001$ ) after the New York flavor policy from 39.03% (5191/13,299) to 32.86% (13,601/41,390) and from 42.78% (18,863/44,090) to 33.93% (49,105/144,734) respectively. In contrast, the proportion of negative tweets related to e-cigarettes significantly increased ( $P<.001$ ) after the

New York State flavor policy announcement in New York State (from 4531/13,299, 34.07% to 18,451/41,390, 44.58%) and in other states (from 14,320/44,090, 32.48% to 64,262/144,734, 44.40%).

In both periods, the proportion of positive tweets in other states was significantly higher than that in New York State (18,863/44,090, 42.78% vs 5191/13,299, 39.03% before the policy; 49,105/144,734, 33.93% vs 13,601/41,390, 32.86% after the policy). Comparing the proportion of negative tweets in other states, before the New York State flavor policy was announced, New York State had a significantly higher ( $P<.001$ ) proportion of negative posts (4531/13,299, 34.07% vs 14,320/44,090, 32.48%). However, after the announcement of the New York State flavor policy, there was no significant difference (18,451/41,390, 44.58% vs 64,262/144,734, 44.40%).

**Table 2.** Proportion of electronic cigarette-related tweets with different sentiments before and after the New York flavor policy announcement in New York State.

Time, sentiments	Tweets, n (%)
<b>Before the New York flavor policy (n=13,299)</b>	
Negative	4531 (34.07)
Positive	5191 (39.03)
Neutral	3577 (26.90)
<b>After the New York flavor policy (n=41,390)</b>	
Negative	18,451 (44.58)
Positive	13,601 (32.86)
Neutral	9338 (22.56)

**Table 3.** Proportion of electronic cigarette-related tweets with different sentiments before and after the New York flavor policy announcement in other states.

Time, sentiments	Tweets, n (%)
<b>Before the New York flavor policy (n=44,090)</b>	
Negative	14,320 (32.48)
Positive	18,863 (42.78)
Neutral	10,907 (24.74)
<b>After the New York flavor policy (n=144,734)</b>	
Negative	64,262 (44.40)
Positive	49,105 (33.93)
Neutral	31,367 (21.67)

## Public Attitudes Toward the New York State Flavor Policy on Twitter

To examine public attitudes toward the flavor policy on e-cigarettes, we conducted a sentiment analysis on the tweets related to the flavored e-cigarettes policy. As shown in [Table 4](#) and [Table 5](#), we observed that the majority of tweets related to the New York flavor policy were negative in both New York State (from 87/98, 89% to 3810/4565, 83.46%) and other states (200/255, 78.4% to 12,695/15,569, 81.54%). There was no significant change in the proportion of either positive or negative tweets between before and after the New York flavor policy in either New York State or other states.

**Table 4.** Proportion of tweets with different sentiments toward the flavor policy before and after the New York flavor policy announcement in New York State.

Time, sentiments	Tweets, n (%)
<b>Before the New York flavor policy (n=98)</b>	
Negative	87 (88.78)
Positive	10 (10.20)
Neutral	1 (1.02)
<b>After the New York flavor policy (n=4565)</b>	
Negative	3810 (83.46)
Positive	650 (14.24)
Neutral	105 (2.30)

**Table 5.** Proportion of tweets with different sentiments toward the flavor policy before and after the New York flavor policy announcement in other states.

Time, sentiments	Tweets, n (%)
<b>Before the New York flavor policy (n=255)</b>	
Negative	200 (78.43)
Positive	39 (15.29)
Neutral	16 (6.27)
<b>After the New York flavor policy (n=15,569)</b>	
Negative	12,695 (81.54)
Positive	2551 (16.39)
Neutral	323 (2.07)

## Top Topics Discussed on e-Cigarettes

To further understand how the New York State flavor policy affected the public attitudes toward e-cigarettes, top topics related to e-cigarettes were generated before and after the announcement of the New York State flavor policy in New York ([Table 6](#)) and other states ([Table 7](#)). We observed that before the announcement of the New York State flavor policy, the topics about e-cigarettes between New York State and other states were similar. The majority of tweets focused on health or teenager vaping-related topics. For example, in both New York State and other states, a typical tweet is “Juul has created nicotine addiction in a whole generation of people who were statistically unlikely to start smoking cigarettes in the first place.”

We conducted 2 proportion z-tests to compare the proportion of tweets with different sentiments toward the flavor policy between the New York State and other states. In both time periods, New York State had a significantly higher proportion of negative tweets than other states ( $P=.03$  before the policy,  $P=.003$  after the policy). There was no significant difference in the proportion of positive posts between the New York State and other states before the announcement of the New York State flavor policy ( $P=.21$ ). However, after the announcement of the policy, the proportion of positive tweets in other states was significantly higher than that in New York State ( $P<.001$ ).

After the announcement of the New York State flavor policy, we observed that while people kept discussing teenage vaping and smoking-related topics, the proportion of topics related to the policy increased, while the proportion of topics related to health decreased in both New York and other states. In addition, other states without a flavor policy had a higher proportion of topics related to the flavor policy compared to the New York State. The keyword “ban” appeared in the third topic in New York State while it showed up in the first and third topics in other states.

To examine whether there were some changes in the discussion about the New York flavor policy, top topics that related to New York flavor policy were generated before and after the announcement of the flavor policy in New York ([Table 8](#)) and other states ([Table 9](#)). Before the announcement of the New



York flavor policy, the majority of the tweets focused on discussions about banning flavored e-cigarette products in both New York State and other states. However, after the New York flavor policy was announced, the topics were more diverse. Besides the discussions on banning flavored e-cigarettes, there

were some discussions about teenager vaping and health problems, especially in New York State. Comparing to the topics from New York State after the announcement of the policy, the topics in other states tended to focus more on the policy of the flavored vaping products.

**Table 6.** Top topics related to electronic cigarettes discussed in New York State.

Topics	Token, n (%)	Keywords
<b>Before the New York flavor policy (n=13,299)</b>		
Vaping leads to nicotine addiction in those who are unlikely to smoke	3524 (26.50)	cigarette, create, generation, addiction, smoking, first, start, whole, unlikely, statistically
Lung diseases linked to vaping	2248 (16.90)	vape, stare, stupid_face, say, people, link, lung, case, teen, tell
Quit smoking and vaping	2061 (15.50)	vape, go, make, think, smoking, quit, thing, vaping, year, smoker
Juul gets good kids ill	2021 (15.20)	juul, hit, pod, look, get, good, kid, buy, illness, day
Vaping leads to diseases and hospitalization	1942 (14.60)	use, friend, vaping, vape, beer, dear, disease, hospitalize, almost, call
Ban cigarettes	1503 (11.30)	find, help, product, state, cigarette, would, report, add, ban, cig
<b>After the New York flavor policy (n=41,390)</b>		
Teenager vaping	12,127 (29.30)	vape, vaping, say, go, want, kid, think, single, teen, see
Smoking and vaping	11,920 (28.80)	vape, cigarette, vaping, get, smoking, people, consider, product, tobacco, age
Ban flavored electronic cigarettes	7285 (17.60)	ban, flavor, let, vape, next, public, cig, pick, homeless, cup
A joke about juul is cool like a refrigerator	5091 (12.30)	smoke, juul, year, thank, day, refrigerator, beaesg, mad, man, easy
Discussion about flavored electronic cigarette policy	4925 (11.90)	juul, pod, government, impeach, formal_warn, flavor, look, hit, take, guy

**Table 7.** Top topics related to electronic cigarettes discussed in other states.

Topics	Token, n (%)	Keywords
<b>Before the New York flavor policy (n=44,090)</b>		
Vaping leads to nicotine addiction in those who are unlikely to smoke	11,463 (26)	smoking, cigarette, generation, whole, addiction, start, first, unlikely, statistically, find
Teenage vaping juul	7628 (17.30)	juul, vape, hit, be, get, people, pod, kid, say, think
Lung diseases linked to vaping	7495 (17)	vape, go, stupid, face, lung, material, year, would, bad, cbd
Vaping among friends	7231 (16.40)	vape, new, use, friend, level, baby, stare, dear, stop, state
Health problems associated with teenager vaping	5820 (13.20)	create, vaping, smoke, link, cigarette, teen, damage, add, cig, health
Vaping is bad	4409 (10)	juul, case, meanwhile, chad, look, black, vaper, fuck, almost, rip
<b>After the New York flavor policy (n=144,734)</b>		
Discussions about the policy on banning flavored electronic cigarettes	54,999 (38)	vape, flavor, government, smoking, age, start, thank, product, warning, ban
Death associated with vaping	35,315 (24.40)	vape, vaping, beaesg, change, look, kill, fast, seem, teen, industry
Discussion on banning vaping	32,131 (22.20)	ban, next, single, consider, maybe, public, coffee, water, bottle, pick
Stop vaping juul	22,144 (15.30)	juul, smoke, pod, let, go, formal, fuck, ask, stop, bring

**Table 8.** Top topics related to the New York flavor policy discussed in New York State.

Topic	Token, n (%)	Keywords
<b>Before the New York flavor policy (n=98)</b>		
Ban flavored electronic cigarettes	52 (53)	ban, vaping, cigarette, vaping, ecig, product, smoking, smoker, would, mormon
Ban flavored electronic cigarettes, teenager vaping	46 (47)	ban, flavor, vape, kid, vaping, people, tobacco, cigarette, lead, young
<b>After the New York flavor policy (n=4565)</b>		
Ban flavored electronic cigarettes, teenager vaping	1278 (28)	ban, vape, vaping, flavor, product, cigarette, kid, nicotine, people, tobacco
Ban flavored electronic cigarettes, teenager vaping, health-related issue	1,173 (25.70)	ban, vape, flavor, teen, vaping, vaper, cig, health, people, tobacco
Discussion about the policy on banning flavored electronic cigarettes	1,137 (24.90)	ban, flavor, vape, trump, product, vaping, vote, say, go, back
Discussion about the policy on banning flavored electronic cigarettes	977 (21.40)	flavor, government, juul let, pod, impeach, formal, warning, information, false

**Table 9.** Top topics related to the New York flavor policy discussed in other states.

Topics	Token, n (%)	Keywords
<b>Before the New York flavor policy (n=255)</b>		
Ban flavored electronic cigarettes	147 (57.60)	ban, cigarette, smoker, hire, vape, smoke, quit, much, vaping, cig
Discussions about banning vaping	108 (42.40)	ban, vape, vaping, cig, juul, smoking, cigarette, smoke, mormon, public
<b>After the New York flavor policy (n=15,569)</b>		
Discussions about banning flavored vaping product	6071 (39)	ban, vape, product, flavor, vaping, trump, shop, industry, people, business
Discussions about banning flavored vaping product, death associated with vaping	4063 (26.10)	flavor, die, ban, vaping, vape, year, illegal, lead, hand, crisis
Discussions about the policy on banning flavored electronic cigarettes	3316 (21.30)	flavor, government, juul, let, pod, warning, impeach, formal, public, nicotine
Discussions about banning disguising vaping product	2119 (13.61)	vape, disguise, ban, see, maga_meh, away, show, school, listen, walk

## Discussion

### Principal Findings

With the epidemic of e-cigarettes in the United States especially among youth and young adults, all tobacco regulatory policies aim to prevent the initiation of e-cigarettes use in youth. The New York State flavor policy was announced with the intention to protect youth, as flavors are one of the major reasons for the dramatic increase in youth vaping initiation and maintenance. Meanwhile, flavors are also key marketing strategies of vaping retailers and companies to attract youth to vape. Therefore, it is of utmost importance to evaluate the public perception of such a flavor policy, and more importantly, how the flavor policy affected the public perception of e-cigarettes, which might potentially affect user behavior to further protect public health, especially of the youth. We hypothesize that the New York State flavor policy will be supported by parents, health educators, and public health professionals and be opposed by current vapers or e-cigarette retailers or companies. To better test our hypothesis, we could distinguish individuals from

organizations in future studies and examine the differences in sentiments and topics between individuals and organizations. This could help us explore and compare the attitudes of different groups of people.

In this study, we showed that after the announcement of the New York State flavor policy, the public attitudes to e-cigarettes became more negative in New York State and other states. In both New York State and other states, before the announcement of the New York State flavor policy, the greatest proportion of e-cigarette-related tweets was positive tweets, but after the policy was announced, the greatest proportion was negative tweets. One possible reason for more negative attitudes toward e-cigarettes could be the increased exposure of the public to the potential harm of vaping. Meanwhile, although the keyword “ban” was not included in these tweets, it is possible that some tweets might be critical of the New York flavor policy, which could partially contribute to more negative attitude. Our results showed that the public attitudes toward the flavor policy on flavored e-cigarettes remained negative and did not change much between before and after the New York flavor policy in

both New York State and other states. One possible explanation is that these Twitter users might be more likely to be e-cigarette users who want to continue vaping.

Although not statistically significant, we observed an increase in negative tweets related to the New York flavor policy in other states, which contrasted the decrease in negative tweets related to the New York flavor policy in New York after the New York flavor policy was announced. This might be because Twitter users in New York State might accept the policy after the New York flavor policy was announced while Twitter users in other states might worry that they would have a similar flavor policy in their states. As neither the changes in New York State nor other states were significant, the observed differences could also be due to random noise.

By applying topic modeling to examine the main topics related to e-cigarettes and the New York flavor policy after this policy announcement on flavored e-cigarettes, we showed that besides the discussion about the flavor ban, the main topics were teenage vaping and health-related, which might cause the increase in the proportion of negative tweets. In addition, these 2 topics were also mentioned frequently by Twitter users in New York State and other states before the New York State flavor policy was announced, which could be due to the occurrence of EVALI in 2019. These topics showed public awareness of e-cigarettes' harmfulness. In addition, we showed that other states had a higher proportion of tweets discussing the flavor ban after the policy in New York State was announced. These results suggest that Twitter users in the states that did not have a ban on flavored e-cigarettes had a significant concern about the potential regulatory policy on flavored e-cigarettes in their states.

### Comparison With Prior Work

Compared with that in a previous study analyzing e-cigarette tweets between October 2015 and February 2016 [34], the proportion of positive tweets toward e-cigarettes in our study decreased significantly. The percentage of positive tweets about e-cigarettes decreased from previously reported 66.4% (589/887) to 39.03% (5191/13,299) in New York State and 42.78% (18,863/44,090) in other states before the New York State policy on flavored e-cigarettes was announced, which might result from the epidemic of EVALI in 2019. After the New York State flavor policy was announced, the proportion of positive tweets on e-cigarettes was even lower. One previous study showed that although the prevalent topics were about the stigma and the harmfulness of e-cigarettes, most tweets denied that e-cigarettes were health hazards [34]. However, in our study, people were more concerned about the health problems and teenage vaping. Therefore, the public attitudes toward e-cigarettes became less positive over time, which might be due to the wide awareness of the potential health effects of e-cigarette use. There have been few studies on flavored e-cigarettes policy on social media. One study showed that although the flavored cigarette ban could be considered as successful in controlling adolescent tobacco use, there was a high probability that they would switch to other flavored tobacco products [3]. In addition, another study showed that after New York City banned flavored cigarettes, the sale of nonflavored tobacco products increased [35]. In our study, we showed that the proportion of negative e-cigarette tweets

increased in both New York and other states, which might be due to the public awareness of the negative health effects of e-cigarettes or the potential effects of the New York State flavor policy.

### Limitations

In this study, we used Twitter data to analyze users' attitudes toward e-cigarettes and the New York State policy on flavored e-cigarettes. Although Twitter is one of the most popular social media platforms in the United States, Twitter users might not represent the whole population as the demographic composition of Twitter users is different from that of the whole population. According to Pew Research Center data in 2018, approximately 24% of US adults used Twitter and 45% of the younger Americans (18-24 years old) were Twitter users [36]. Among adult Twitter users, only 15% regularly use Twitter, and young adults and minorities tend to be more highly represented on Twitter than in the general population. Meanwhile, very active and passive users are more prevalent than moderate users on Twitter. Thus, the results of this study were from a nonuniform subsample of tweets posted by a nonrepresentative portion of the US populations.

Other demographic information (including age, gender) were not included in our study owing to the limitation of Twitter data. In addition, the geographical location of users can be collected only if they are willing to share. Gore et al [37] mentioned that 95% of the Twitter users preferred not to share the location for a single tweet, and 1% of the users were willing to share the locations for the majority of the tweets they posted. However, in our data set, there were 68.10% (301,4419/4,426,290) of tweets containing the location of either tweets or users. Some tweets without geolocation information were not included in our study, which might introduce some biases. In addition, Padilla et al [38] showed that both temporal and spatial measures could bias the sentiment of an individual's tweet. We did not examine the effect of temporal and spatial measures on the sentiment of the tweets, which might bias the sentiment results. In this study, we did not examine how the policy on flavored e-cigarettes affects the users' behavior patterns such as switching to different flavored e-cigarettes, quitting vaping, or switching to cigarette smoking, which require further investigation. In addition, considering the co-occurrence of EVALI during our study period, there could be some biases with the potential impact of the New York flavor policy on public attitudes toward e-cigarettes.

In our study, we focused on analyzing the differences between public response to New York flavor policy in New York State and other states. Thus, we categorized the states that did not have a flavor policy as 1 group. However, this may cause limitations because these states may have their unique characteristics that might impact their attitudes toward New York flavor policy, such as their previous policies on tobacco products and government's attitudes toward e-cigarettes. In addition, since San Francisco announced the flavor e-cigarette policy as a city, we only excluded the posts from that city. However, this policy might influence other cities in California, which could not be measured in this study.

Our results were insufficient for capturing the nuance of the conversation about flavored e-cigarette use. An increase in negative sentiments among policy-related tweets in New York State could reflect resistance toward additional regulation or concerns that banning flavored e-cigarettes might lead to increased cigarette usage among new and existing smokers. A machine learning classifier could be used in future studies to differentiate between individuals' reasons for positive or negative perceptions toward the ban. Although the flavor policy on e-cigarettes in New York State was announced on September 17, 2019, it was never actually implemented in New York State during the study period. Therefore, the impact of the New York flavor policy might be underestimated. This might be one of the reasons that the changes in public attitudes toward e-cigarettes between New York State and other states were similar. The recent US Food and Drug Administration flavor enforcement policy implemented on February 6, 2020 and the New York State law on flavored vapor products, implemented

on May 18, 2020, might have more obvious impact on public attitudes toward e-cigarettes, which will be explored in our future studies.

## Conclusions

Using social media data from Twitter, our study showed that after the policy on flavored e-cigarettes in New York State was announced, the discussions about e-cigarettes and the flavor policy increased significantly. Twitter users in the states that did not have a flavored e-cigarette policy have similar concerns about the flavor policy as those in New York State. Sentiment analysis revealed that after the New York flavor policy was announced, the public tended to have a more negative attitude toward e-cigarettes in New York State and other states. Together, our study provides an initial investigation about the potential impact of the New York State policy of flavored e-cigarettes on the public attitudes toward e-cigarettes, which might subsequently affect user behavior.

## Acknowledgments

The research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health and the Food and Drug Administration Center for Tobacco Products under award U54CA228110. This content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the US Food and Drug Administration. The data and scripts used for analysis and creating figures are available on request from the corresponding author DL.

## Authors' Contributions

LS, ZX, and DL conceived and designed the study. LS analyzed the data and wrote the manuscript. LS, XL, ZX, and DL assisted with interpretation of analyses and edited the manuscript.

## Conflicts of Interest

None declared.

## References

1. King BA, Tynan MA, Dube SR, Arrazola R. Flavored-little-cigar and flavored-cigarette use among U.S. middle and high school students. *J Adolesc Health* 2014 Jan;54(1):40-46 [FREE Full text] [doi: [10.1016/j.jadohealth.2013.07.033](https://doi.org/10.1016/j.jadohealth.2013.07.033)] [Medline: [24161587](https://pubmed.ncbi.nlm.nih.gov/24161587/)]
2. Villanti AC, Richardson A, Vallone DM, Rath JM. Flavored tobacco product use among U.S. young adults. *Am J Prev Med* 2013 Apr;44(4):388-391. [doi: [10.1016/j.amepre.2012.11.031](https://doi.org/10.1016/j.amepre.2012.11.031)] [Medline: [23498105](https://pubmed.ncbi.nlm.nih.gov/23498105/)]
3. Courtemanche CJ, Palmer MK, Pesko MF. Influence of the Flavored Cigarette Ban on Adolescent Tobacco Use. *Am J Prev Med* 2017 May;52(5):e139-e146 [FREE Full text] [doi: [10.1016/j.amepre.2016.11.019](https://doi.org/10.1016/j.amepre.2016.11.019)] [Medline: [28081999](https://pubmed.ncbi.nlm.nih.gov/28081999/)]
4. Gentzke AS, Creamer M, Cullen KA, Ambrose BK, Willis G, Jamal A, et al. Vital Signs: Tobacco Product Use Among Middle and High School Students - United States, 2011-2018. *MMWR Morb Mortal Wkly Rep* 2019 Feb 15;68(6):157-164 [FREE Full text] [doi: [10.15585/mmwr.mm6806e1](https://doi.org/10.15585/mmwr.mm6806e1)] [Medline: [30763302](https://pubmed.ncbi.nlm.nih.gov/30763302/)]
5. Banning the sale of flavored tobacco products. City and County of San Francisco Ordinance. URL: <https://sfbos.org/sites/default/files/o0140-17.pdf> [accessed 2021-05-17]
6. Grana R, Benowitz N, Glantz SA. E-Cigarettes. *Circulation* 2014 May 13;129(19):1972-1986. [doi: [10.1161/circulationaha.114.007667](https://doi.org/10.1161/circulationaha.114.007667)]
7. Schneller LM, Bansal-Travers M, Goniewicz ML, McIntosh S, Ossip D, O'Connor RJ. Use of Flavored E-Cigarettes and the Type of E-Cigarette Devices Used among Adults and Youth in the US-Results from Wave 3 of the Population Assessment of Tobacco and Health Study (2015-2016). *Int J Environ Res Public Health* 2019 Aug 20;16(16):2991 [FREE Full text] [doi: [10.3390/ijerph16162991](https://doi.org/10.3390/ijerph16162991)] [Medline: [31434229](https://pubmed.ncbi.nlm.nih.gov/31434229/)]
8. Gerloff J, Sundar IK, Freter R, Sekera ER, Friedman AE, Robinson R, et al. Inflammatory Response and Barrier Dysfunction by Different e-Cigarette Flavoring Chemicals Identified by Gas Chromatography-Mass Spectrometry in e-Liquids and e-Vapors on Human Lung Epithelial Cells and Fibroblasts. *Appl In Vitro Toxicol* 2017 Mar 01;3(1):28-40 [FREE Full text] [doi: [10.1089/avt.2016.0030](https://doi.org/10.1089/avt.2016.0030)] [Medline: [28337465](https://pubmed.ncbi.nlm.nih.gov/28337465/)]

9. Gotts JE, Jordt S, McConnell R, Tarran R. What are the respiratory effects of e-cigarettes? *BMJ* 2019 Sep 30;366:15275 [FREE Full text] [doi: [10.1136/bmj.15275](https://doi.org/10.1136/bmj.15275)] [Medline: [31570493](https://pubmed.ncbi.nlm.nih.gov/31570493/)]
10. Lee WH, Ong S, Zhou Y, Tian L, Bae HR, Baker N, et al. Modeling Cardiovascular Risks of E-Cigarettes With Human-Induced Pluripotent Stem Cell-Derived Endothelial Cells. *J Am Coll Cardiol* 2019 Jun 04;73(21):2722-2737 [FREE Full text] [doi: [10.1016/j.jacc.2019.03.476](https://doi.org/10.1016/j.jacc.2019.03.476)] [Medline: [31146818](https://pubmed.ncbi.nlm.nih.gov/31146818/)]
11. Krishnasamy VP, Hallowell BD, Ko J, Board A, Hartnett KP, Salvatore PP, Lung Injury Response Epidemiology/Surveillance Task Force. Update: Characteristics of a Nationwide Outbreak of E-cigarette, or Vaping, Product Use-Associated Lung Injury - United States, August 2019-January 2020. *MMWR Morb Mortal Wkly Rep* 2020 Jan 24;69(3):90-94 [FREE Full text] [doi: [10.15585/mmwr.mm6903e2](https://doi.org/10.15585/mmwr.mm6903e2)] [Medline: [31971931](https://pubmed.ncbi.nlm.nih.gov/31971931/)]
12. Siegel D. Update: Interim Guidance for Health Care Providers Evaluating and Caring for Patients with Suspected E-cigarette, or Vaping, Product Use Associated Lung Injury. *Morbidity and Mortality Weekly Report* 2019;68(41):919-927. [doi: [10.3410/f.736757259.793575010](https://doi.org/10.3410/f.736757259.793575010)]
13. Restricting the sale, manufacture, and distribution of tobacco products, including electronic cigarettes. City and County of San Francisco. URL: <https://sf.gov/legistar.com/View.ashx?M=F&ID=7111897&GUID=7C3912E3-BB56-420D-896E-6FEA1391287D> [accessed 2021-05-17]
14. Michigan's ban of flavored nicotine vaping products to protect children's health effective immediately with rules filing. MDHHS. URL: <https://www.michigan.gov/mdhhs/0,5885,7-339--507760--,00.html> [accessed 2021-05-17]
15. Governor Cuomo announces New York State implements first-in-the-nation ban on flavored e-cigarettes. CBS News. URL: <https://www.cbsnews.com/news/governor-andrew-cuomo-announces-flavored-vaping-e-cigarette-ban-in-new-york-state-amid-lung-disease-deaths/> [accessed 2021-05-17]
16. Emergency notice of changes in regulations that may affect your business. Rhode Island Department of Health. URL: [https://health.ri.gov/publications/letters/20191004\\_EREcigFlavorBan\\_LetterToRetailers\\_EnglishAndSpanish.pdf](https://health.ri.gov/publications/letters/20191004_EREcigFlavorBan_LetterToRetailers_EnglishAndSpanish.pdf) [accessed 2021-05-17]
17. Tobacco control and prevention program. County of Los Angeles Public Health. URL: <http://publichealth.lacounty.gov/tob/> [accessed 2021-05-17]
18. OHA, OLCC file rules banning flavored vaping sales, including online. Oregon.gov. URL: <https://www.oregon.gov/oha/ERD/Pages/OLCCFileRulesBanningFlavoredVapingSalesIncludingOnline.aspx> [accessed 2021-05-17]
19. Governor Bullock directs ban on flavored e-cigarettes to address public health emergency. Great Falls Tribune. URL: <https://www.greatfallstribune.com/story/news/2019/10/08/montana/3910301002/> [accessed 2021-05-17]
20. Washington State bans vapor products containing vitamin E acetate. Washington State Department of Health. URL: <https://www.doh.wa.gov/Newsroom/Articles/ID/1038/Washington-state-bans-vapor-products-containing-vitamin-E-acetate> [accessed 2021-05-17]
21. Blumenfeld K. Electronic smoking devices. GASP. 2015. URL: [https://www.njgasp.org/wp-content/uploads/2014/05/E-Cigs\\_White\\_Paper.pdf](https://www.njgasp.org/wp-content/uploads/2014/05/E-Cigs_White_Paper.pdf) [accessed 2021-05-17]
22. Minimum standards for retail sale of tobacco and electronic nicotine delivery systems. Department of Public Health. URL: <https://www.mass.gov/doc/105-cmr-665000/download> [accessed 2021-05-17]
23. Kavuluru R, Sabbir A. Toward automated e-cigarette surveillance: Spotting e-cigarette proponents on Twitter. *J Biomed Inform* 2016 Jun;61:19-26 [FREE Full text] [doi: [10.1016/j.jbi.2016.03.006](https://doi.org/10.1016/j.jbi.2016.03.006)] [Medline: [26975599](https://pubmed.ncbi.nlm.nih.gov/26975599/)]
24. Zhou J, Zhang Q, Zeng DD, Tsui KL. Influence of Flavors on the Propagation of E-Cigarette-Related Information: Social Media Study. *JMIR Public Health Surveill* 2018 Mar 23;4(1):e27 [FREE Full text] [doi: [10.2196/publichealth.7998](https://doi.org/10.2196/publichealth.7998)] [Medline: [29572202](https://pubmed.ncbi.nlm.nih.gov/29572202/)]
25. Chu K, Allem J, Unger JB, Cruz TB, Akbarpour M, Kirkpatrick MG. Strategies to find audience segments on Twitter for e-cigarette education campaigns. *Addict Behav* 2019 Apr;91:222-226 [FREE Full text] [doi: [10.1016/j.addbeh.2018.11.015](https://doi.org/10.1016/j.addbeh.2018.11.015)] [Medline: [30497815](https://pubmed.ncbi.nlm.nih.gov/30497815/)]
26. Wojcik S, Hughes A. Sizing up Twitter users. Pew Research Center. URL: <https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users/> [accessed 2021-05-17]
27. Schober MF, Pasek J, Guggenheim L, Lampe C, Conrad FG. Social Media Analyses for Social Measurement. *Public Opin Q* 2016;80(1):180-211 [FREE Full text] [doi: [10.1093/poq/nfv048](https://doi.org/10.1093/poq/nfv048)] [Medline: [27257310](https://pubmed.ncbi.nlm.nih.gov/27257310/)]
28. Farrimond H. E-cigarette regulation and policy: UK vapers' perspectives. *Addiction* 2016 Jun;111(6):1077-1083. [doi: [10.1111/add.13322](https://doi.org/10.1111/add.13322)] [Medline: [26802864](https://pubmed.ncbi.nlm.nih.gov/26802864/)]
29. Lienemann BA, Unger JB, Cruz TB, Chu K. Methods for Coding Tobacco-Related Twitter Data: A Systematic Review. *J Med Internet Res* 2017 Mar 31;19(3):e91 [FREE Full text] [doi: [10.2196/jmir.7022](https://doi.org/10.2196/jmir.7022)] [Medline: [28363883](https://pubmed.ncbi.nlm.nih.gov/28363883/)]
30. Chen L, Lu X, Yuan J, Luo J, Luo J, Xie Z, et al. A Social Media Study on the Associations of Flavored Electronic Cigarettes With Health Symptoms: Observational Study. *J Med Internet Res* 2020 Jun 22;22(6):e17496 [FREE Full text] [doi: [10.2196/17496](https://doi.org/10.2196/17496)] [Medline: [32568093](https://pubmed.ncbi.nlm.nih.gov/32568093/)]
31. Lu X, Chen L, Yuan J, Luo J, Luo J, Xie Z, et al. User Perceptions of Different Electronic Cigarette Flavors on Social Media: Observational Study. *J Med Internet Res* 2020 Jun 24;22(6):e17280 [FREE Full text] [doi: [10.2196/17280](https://doi.org/10.2196/17280)] [Medline: [32579123](https://pubmed.ncbi.nlm.nih.gov/32579123/)]

32. Hutto C, Gilbert E. VADER: A parsimonious rule-based model for sentiment analysis of social media text. Proceedings of the International AAAI Conference on Web and Social Media. 2015. URL: <https://ojs.aaai.org/index.php/ICWSM/article/view/14550> [accessed 2022-01-27]
33. Blei D, Ng AY, Jordan MI. Latent dirichlet allocation. Journal of Machine Learning Research. 2003. URL: <https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf> [accessed 2022-01-27]
34. Martinez LS, Hughes S, Walsh-Buhi ER, Tsou M. "Okay, We Get It. You Vape": An Analysis of Geocoded Content, Context, and Sentiment regarding E-Cigarettes on Twitter. J Health Commun 2018;23(6):550-562. [doi: [10.1080/10810730.2018.1493057](https://doi.org/10.1080/10810730.2018.1493057)] [Medline: [29979920](https://pubmed.ncbi.nlm.nih.gov/29979920/)]
35. Farley SM, Johns M. New York City flavoured tobacco product sales ban evaluation. Tob Control 2017 Jan;26(1):78-84. [doi: [10.1136/tobaccocontrol-2015-052418](https://doi.org/10.1136/tobaccocontrol-2015-052418)] [Medline: [26872486](https://pubmed.ncbi.nlm.nih.gov/26872486/)]
36. Smith A, Anderson M. Social media use in 2018. Pew Research Center. URL: <https://www.pewresearch.org/internet/2018/03/01/social-media-use-in-2018/> [accessed 2021-05-17]
37. Gore RJ, Diallo S, Padilla J. You Are What You Tweet: Connecting the Geographic Variation in America's Obesity Rate to Twitter Content. PLoS One 2015;10(9):e0133505 [FREE Full text] [doi: [10.1371/journal.pone.0133505](https://doi.org/10.1371/journal.pone.0133505)] [Medline: [26332588](https://pubmed.ncbi.nlm.nih.gov/26332588/)]
38. Padilla JJ, Kavak H, Lynch CJ, Gore RJ, Diallo SY. Temporal and spatiotemporal investigation of tourist attraction visit sentiment on Twitter. PLoS One 2018;13(6):e0198857 [FREE Full text] [doi: [10.1371/journal.pone.0198857](https://doi.org/10.1371/journal.pone.0198857)] [Medline: [29902270](https://pubmed.ncbi.nlm.nih.gov/29902270/)]

## Abbreviations

**e-Cigarette:** electronic cigarette

**EVALI:** e-cigarette or vaping use-associated lung injury

**PATH:** Population Assessment of Tobacco and Health

*Edited by Y Khader; submitted 22.10.20; peer-reviewed by R Gore, N Cesare, R Halkes; comments to author 11.01.21; revised version received 07.03.21; accepted 20.11.21; published 03.02.22.*

*Please cite as:*

*Sun L, Lu X, Xie Z, Li D*

*Public Reactions to the New York State Policy on Flavored Electronic Cigarettes on Twitter: Observational Study*

*JMIR Public Health Surveill 2022;8(2):e25216*

*URL: <https://publichealth.jmir.org/2022/2/e25216>*

*doi: [10.2196/25216](https://doi.org/10.2196/25216)*

*PMID: [35113035](https://pubmed.ncbi.nlm.nih.gov/35113035/)*

©Li Sun, Xinyi Lu, Zidian Xie, Dongmei Li. Originally published in JMIR Public Health and Surveillance (<https://publichealth.jmir.org>), 03.02.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

# America's HIV Epidemic Analysis Dashboard: Protocol for a Data Resource to Support Ending the HIV Epidemic in the United States

Patrick Sean Sullivan<sup>1</sup>, DVM, PhD; Cory R Woodyatt<sup>2</sup>, BSN, RN; Oskian Kouzouian<sup>3</sup>, JD; Kristen J Parrish<sup>4</sup>, MAT, PMP; Jennifer Taussig<sup>1</sup>, MPH; Chris Conlan<sup>5</sup>, BA; Harold Phillips<sup>3</sup>, MRP

<sup>1</sup>Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, United States

<sup>2</sup>Oregon Health & Science University, Portland, OR, United States

<sup>3</sup>Office of Infectious Disease and HIV/AIDS Policy, US Department of Health and Human Services, Washington, DC, United States

<sup>4</sup>Insignia Federal Group, Falls Church, VA, United States

<sup>5</sup>Conlan Scientific, Bethesda, MD, United States

**Corresponding Author:**

Patrick Sean Sullivan, DVM, PhD

Department of Epidemiology

Rollins School of Public Health

Emory University

1518 Clifton Road

Atlanta, GA, 30329

United States

Phone: 1 4047272038

Email: [pssulli@emory.edu](mailto:pssulli@emory.edu)

## Abstract

**Background:** The Ending the HIV Epidemic (EHE) plan aims to end the HIV epidemic in the United States by 2030. Having timely and accessible data to assess progress toward EHE goals at the local level is a critical resource to achieve this goal.

**Objective:** The aim of this paper was to introduce America's HIV Epidemic Analysis Dashboard (AHEAD), a data visualization tool that displays relevant data on the 6 HIV indicators provided by the Centers for Disease Control and Prevention. AHEAD can be used to monitor progress toward ending the HIV epidemic in local communities across the United States. Its objective is to make data available to stakeholders, which can be used to measure national and local progress toward 2025 and 2030 EHE goals and to help jurisdictions make local decisions that are grounded in high-quality data.

**Methods:** AHEAD displays data from public health data systems (eg, surveillance systems and census data), organized around the 6 EHE indicators (HIV incidence, knowledge of HIV status, HIV diagnoses, linkage to HIV medical care, viral HIV suppression, and preexposure prophylaxis coverage). Data are displayed for each of the EHE priority areas (48 counties in Washington, District of Columbia, and San Juan, Puerto Rico) which accounted for more than 50% of all US HIV diagnoses in 2016 and 2017 and 7 primarily southern states with high rates of HIV in rural communities. AHEAD also displays data for the 43 remaining states for which data are available. Data features prioritize interactive data visualization tools that allow users to compare indicator data stratified by sex at birth, race or ethnicity, age, and transmission category within a jurisdiction (when available) or compare data on EHE indicators between jurisdictions.

**Results:** AHEAD was launched on August 14, 2020. In the 11 months since its launch, the Dashboard has been visited 26,591 times by 17,600 unique users. About one-quarter of all users returned to the Dashboard at least once. On average, users engaged with 2.4 pages during their visit to the Dashboard, indicating that the average user goes beyond the informational landing page to engage with 1 or more pages of data and content. The most frequently visited content pages are the jurisdiction webpages.

**Conclusions:** The Ending the HIV Epidemic plan is described as a "whole of society" effort. Societal public health initiatives require objective indicators and require that all societal stakeholders have transparent access to indicator data at the level of the health jurisdictions responsible for meeting the goals of the plan. Data transparency empowers local stakeholders to track movement toward EHE goals, identify areas with needs for improvement, and make data-informed adjustments to deploy the expertise and resources required to locally tailor and implement strategies to end the HIV epidemic in their jurisdiction.

(*JMIR Public Health Surveill* 2022;8(2):e33522) doi:[10.2196/33522](https://doi.org/10.2196/33522)

**KEYWORDS**

HIV; dashboard; data; data dashboard; infectious disease; infodemiology; surveillance; public health; United States; monitoring

## *Introduction*

The HIV epidemic remains a critical public health priority in the United States. Despite welcome advances in treatment, survival of people living with HIV [1], and HIV prevention [2], HIV remained the 9th leading cause of death among Americans aged 25-34 years in 2019 [3]. Realizing the full potential of the available powerful tools for HIV diagnosis, treatment and prevention will require coordination, bringing interventions to scale, and monitoring progress in HIV prevention, treatment, and care. To facilitate this process, in 2019, the US Department of Health and Human Services initiated the Ending the HIV Epidemic (EHE) initiative and operational plan, which provided a targeted infusion of new resources and support to local communities working together and with the federal government to end the HIV epidemic in America [4]. The goal of the EHE initiative is to accelerate HIV prevention progress and reduce the number of new HIV infections in the United States by 75% by 2025 and 90% by 2030 [4]. The initiative seeks to achieve this goal by providing those areas most in need with the additional expertise, technology, and resources required to scale up 4 key strategies (diagnose, treat, prevent, and respond) needed to end the HIV epidemic in their communities. To achieve maximum impact, EHE focuses its efforts in 48 priority counties, Washington, District of Columbia, and San Juan, Puerto Rico, where more than 50 percent of new HIV diagnoses occurred in 2016 and 2017, and an additional 7 priority states with a substantial rate of HIV diagnoses in rural areas, bringing the total number of priority jurisdictions to 57 [4].

Health statistics are recognized globally as critical to the development, implementation, monitoring, and evaluation of public health programs [5]. Accordingly, the EHE initiative requires the effective use of data to identify baseline levels of new HIV infections, establish 5-year and 10-year goals for reductions in transmissions, monitor interim progress toward these goals, and provide the jurisdictions responsible for public health with tools to easily illustrate progress and areas of opportunities to local prevention providers and stakeholders. Although existing online tools provide access to HIV surveillance and related data in graphical and tabular formats [6,7], the existing tools are not specific to the EHE initiative and are not organized around either the explicit goals of EHE, the EHE jurisdictions, or the 4 EHE key strategies. Further, experience with other dashboard tools suggests that the mechanisms by which dashboards can promote changes in public health programs include the opportunity for a comparative

assessment of indicators among jurisdictions [8]. For these reasons, we developed and implemented a data visualization dashboard focused on the EHE jurisdictions, indicators, and program goals, which incorporates best practices from the existing tools. We then launched the dashboard and collected data on the usage of the dashboard during its first year of use. In this manuscript, we describe the rationale, stakeholder input process, and scope and sources of the data elements and the use of America's HIV Epidemic Analysis Dashboard (AHEAD) in its first year of public availability.

## *Methods*

### **Overview of AHEAD**

AHEAD was launched on August 14, 2020, as a data visualization tool and primary source of tracking progress on the presidential initiative to end the HIV epidemic in America, the EHE. AHEAD was developed by the US Department of Health and Human Services through its Office of Infectious Disease and HIV/AIDS Policy in collaboration with the Office of the Assistant Secretary for Health, Centers for Disease Control and Prevention (CDC), Health Resources and Services Administration, Indian Health Service, National Institutes of Health, Substance Abuse and Mental Health Services, and Department of Housing and Urban Development.

AHEAD displays baseline and progress data for the nation and the 57 jurisdictions (Figure 1) for 6 indicators: HIV incidence, knowledge of HIV status, HIV diagnoses, linkage to HIV medical care, viral suppression, and pre-exposure prophylaxis (PrEP) coverage. AHEAD also displays indicator data for the remaining 43 US states, when available. Each indicator was chosen with specific public health goals in mind and in line with the 4 key strategies of the initiative: diagnose, prevent, treat, and respond [4]. Incidence measures the overarching goal of reducing new transmissions by 90% by 2030. Influencing just one indicator will not achieve the overall goals of the initiative; these indicators are contingent and progress in one will influence others. HIV diagnoses and knowledge of HIV status are both key to linking people to care and represent important steps on the HIV care continuum. Data have shown that, upon diagnosis, immediate linkage to care and treatment results in improved HIV outcomes [9], so it is important to track how these indicators change over time. Viral suppression and PrEP coverage will have the greatest impact on reducing new transmissions if they are scaled up and optimized for use among priority populations [10].



**Figure 1.** Ending the HIV Epidemic (EHE) priority areas. Asterisks indicate nonpriority states with one or more EHE priority counties.

<b>Alabama</b>	<b>Illinois*</b>	<b>New York*</b>
<b>Arizona*</b>	Cook County	Bronx County
Maricopa County	<b>Indiana*</b>	Kings County
<b>Arkansas</b>	Marion County	New York County
<b>California*</b>	<b>Kentucky</b>	Queens County
Alameda County	<b>Louisiana*</b>	<b>North Carolina*</b>
Los Angeles County	East Baton Rouge Parish	Mecklenburg County
Orange County	Orleans Parish	<b>Ohio*</b>
Riverside County	<b>Maryland*</b>	Cuyahoga County
Sacramento County	Baltimore City	Franklin County
San Diego County	Montgomery County	Hamilton County
San Francisco County	Prince George's County	<b>Oklahoma</b>
<b>District of Columbia</b>	<b>Massachusetts*</b>	<b>Pennsylvania*</b>
<b>Florida*</b>	Suffolk County	Philadelphia County
Broward County	<b>Michigan*</b>	<b>Puerto Rico*</b>
Duval County	Wayne County	San Juan Municipio
Hillsborough County	<b>Mississippi</b>	<b>South Carolina</b>
Miami-Dade County	<b>Missouri</b>	<b>Tennessee*</b>
Orange County	<b>Nevada*</b>	Shelby County
Palm Beach County	Clark County	<b>Texas*</b>
Pinellas County	<b>New Jersey*</b>	Bexar County
<b>Georgia*</b>	Essex County	Dallas County
Cobb County	Hudson County	Harris County
Dekalb County		Tarrant County
Fulton County		Travis County
Gwinnett County		<b>Washington*</b>
		King County

AHEAD displays both baseline (2017) and subsequent annual data for the 6 EHE indicators. 2017 was chosen as the baseline year to create the 2025 and 2030 goals because 2017 was the most recent year for which complete annual data were available when the initiative was announced and included stability of estimates for county-level incidence and knowledge of HIV status. Goals were calculated by applying programmatic goals (75% and 90% reductions in new HIV infections), as indicated by CDC data on HIV incidence. Goals are displayed for each of the Phase I EHE jurisdictions.

### Stakeholder Engagement

The development of electronic resources to promote public health requires input from diverse stakeholders, including data creators, intended users, representatives of public health organizations, and people living with HIV [8,11]. To guide the development of the content and functionality of AHEAD, we conducted a series of interviews with representatives of these groups to help identify the types of data and tailor the formats for display that would be most aligned with the needs of intended users. For all stakeholder interactions, online interview sessions were recorded with the permission of the participants, and field notes were recorded by dedicated notetakers during the discussion. Notes and review of recordings were used to develop summaries of individual stakeholder input and to identify recurring themes across stakeholder data.

### Prelaunch Stakeholder Interviews

From July to August 2020, 13 interviews were conducted with diverse stakeholders from health departments and other organizations supporting public health and advocacy responses to the HIV epidemic. Stakeholders were selected without regard to their previous knowledge of the project. The participants were selected to represent multiple EHE jurisdictions in diverse geographic regions of the United States. The participants were provided with information about the purpose of AHEAD and planned content, and feedback was elicited regarding the overall project, including the features, functions, usability, and content value.

### Limited User Assessment Interviews

Following the initial launch of AHEAD, we elicited further feedback from October and December 2020 via a series of limited user assessments. These were conducted to gather insight into stakeholder flow or types of use, intuitiveness, data scannability, usability of interactive functionalities, design aesthetics, topics for educational tools, strengths, weaknesses, opportunities, and possible threats. The feedback was used to update and refine the website further, ensuring the AHEAD project continued to be in tune and responsive to user needs.







### Data Sources

Of the 6 EHE indicators, 5 (HIV incidence, knowledge of HIV status, HIV diagnoses, linkage to care, and viral HIV suppression) reported on the site are from the National HIV Surveillance System (NHSS) [12]. NHSS is the primary source for monitoring HIV trends in the United States and CDC funds and assists state and local health departments to collect the information [13]. States and local jurisdictions, including the District of Columbia and Puerto Rico, have laws or regulations that require confidential reporting by name for adults and adolescents (ie, persons aged 13 and older) with a confirmed HIV diagnosis. After the removal of personally identifiable information, data from these reports are submitted to CDC. Health departments report these de-identified data to CDC so that information from around the country can be analyzed to determine who is being affected and why, which can inform where HIV prevention and treatment resources are most needed. The 6th indicator, PrEP coverage, uses 4 different data sources that are derived from population-based data sources and the published estimates of the size of groups with PrEP indications [14].

The demographic information used to create stratifications for data is based on information included in the NHSS records for people with diagnosed HIV infection [12]. For PrEP coverage, information about the age distribution and sex of PrEP users is derived from commercial pharmacy data sources using algorithms to identify likely PrEP users [15,16].

The extent of stratification of indicators is based on the level of geography. At smaller geographic levels, stratification is more limited to prevent the possible indirect identification of individuals in small intersectional strata. Accordingly, stratification is more extensive at the national level on AHEAD; the indicators are presented overall, and stratified by age, race or ethnicity, sex at birth (for incidence, knowledge of HIV status and PrEP coverage) or gender (for diagnoses, linkage to HIV medical care, and HIV viral suppression), and transmission category. EHE goals are overall population goals, and demographic categories do not have separate 2025 or 2030 goals. The stratification criteria and categories are provided in detail in [Multimedia Appendix 1](#) and are summarized in [Figure 2](#).

**Figure 2.** Ending the HIV Epidemic indicators and categories for demographic stratification for AHEAD (America’s HIV Epidemic Analysis Dashboard). PrEP: pre-exposure prophylaxis coverage.







Indicators	Age (years)	Race/Ethnicity	Sex at Birth	Transmission Category
 <b>Incidence</b>  <b>Knowledge of Status</b>	<ul style="list-style-type: none"> <li>• 13-24</li> <li>• 25-34</li> <li>• 35-44</li> <li>• 45-54</li> <li>• Over 55</li> </ul>	<ul style="list-style-type: none"> <li>• American Indian/Alaska Native</li> <li>• Asian</li> <li>• Black/African American</li> <li>• Hispanic/Latinx</li> <li>• Multiple races</li> <li>• Native Hawaiian/Other Pacific Islander</li> <li>• White</li> </ul>	<ul style="list-style-type: none"> <li>• Female</li> <li>• Male</li> </ul>	<ul style="list-style-type: none"> <li>• Heterosexual contact (female)</li> <li>• Heterosexual contact (male)</li> <li>• Injection drug use (female)</li> <li>• Injection drug use (male)</li> <li>• Male-to-male sexual contact</li> <li>• Male-to-male sexual contact and injection drug use</li> </ul>
 <b>Diagnoses</b>  <b>Linkage to HIV Medical Care</b>  <b>Viral Suppression</b>	<ul style="list-style-type: none"> <li>• 13-24</li> <li>• 25-34</li> <li>• 35-44</li> <li>• 45-54</li> <li>• Over 55</li> </ul>	<ul style="list-style-type: none"> <li>• American Indian/Alaska Native</li> <li>• Asian</li> <li>• Black/African American</li> <li>• Hispanic/Latinx</li> <li>• Multiple races</li> <li>• Native Hawaiian/Other Pacific Islander</li> <li>• White</li> </ul>	<ul style="list-style-type: none"> <li>• Additional gender identity</li> <li>• Female (gender)</li> <li>• Male (gender)</li> <li>• Transgender female-to-male</li> <li>• Transgender male-to-female</li> </ul>	<ul style="list-style-type: none"> <li>• Heterosexual contact (female)</li> <li>• Heterosexual contact (male)</li> <li>• Injection drug use (female)</li> <li>• Injection drug use (male)</li> <li>• Male-to-male sexual contact</li> <li>• Male-to-male sexual contact and injection drug use</li> <li>• Other (female)</li> <li>• Other (male)</li> </ul>
 <b>PrEP Coverage</b>	<ul style="list-style-type: none"> <li>• 13-24</li> <li>• 25-34</li> <li>• 35-44</li> <li>• 45-54</li> <li>• Over 55</li> </ul>	<ul style="list-style-type: none"> <li>• Asian/other (PrEP only)</li> <li>• Black/African American</li> <li>• Hispanic/Latinx</li> <li>• White</li> </ul>	<ul style="list-style-type: none"> <li>• Female</li> <li>• Male</li> </ul>	

## Available Data

The data available through AHEAD (Figure 3) are organized around the EHE indicators [4]. Data are not available for all indicators, years, and jurisdictions for several reasons. For some indicators, data analyses are not initiated until sufficient time has passed to allow for substantially complete reporting of

HIV-related laboratory results. Preliminary data are the earliest data released and are in essence a sneak preview to the complete annual data set. They do not have 12 months of reporting delay and can apply to both annual and cumulative quarterly data. Provisional data are data that do have at least 12 months of reporting delay and apply only to annual data.

**Figure 3.** EHE indicators, available data as of Fall 2021, and the number of EHE states, territories, and counties with indicator data. EHE: Ending the HIV Epidemic; PrEP: pre-exposure prophylaxis coverage. \*Data are preliminary as of March 2021; †Data are preliminary as of December 2020; ‡Data for years 2017 and 2018 are provided for 42 jurisdictions.

Indicator	Available Data	National	No. of EHE States/Territories	No. of EHE Counties/Areas
 <b>Incidence</b>	2017, 2018, 2019	X	38	48 + Washington, DC
 <b>Knowledge of Status</b>	2017, 2018, 2019	X	51	48 + Washington, DC + San Juan, PR
 <b>Diagnoses</b>	2017, 2018, 2019, 2020*, 2021*	X	51	48 + Washington, DC + San Juan, PR
 <b>Linkage to HIV Medical Care</b>	2017, 2018, 2019, 2020*	X	45†	45 + Washington, DC
 <b>Viral Suppression</b>	2017, 2018, 2019	X	45‡	45 + Washington, DC
 <b>PrEP Coverage</b>	2017, 2018, 2019, 2020†	X	51	48 + Washington, DC + San Juan, PR

## Functionalities for Data Display and Insights

Functionalities are provided in AHEAD to allow users to interact with the data in ways that promote a nuanced understanding of the data and suggest areas for possible programmatic focus, or for the collection of further contextual data to understand programmatic successes or needs for improvement. Depending on the levels of stratification available within the jurisdiction, such insights might allow the identification of specific subgroups (eg, Black people assigned female sex at birth) in which progress toward indicators is exemplary or limited compared to other subgroups. The illustration of goals allows users to assess whether the current progress toward indicators is sufficient to achieve eventual EHE goals, or whether the programs and efforts addressing specific indicators need to be modified or intensified to meet EHE goals.

Data can be viewed as either a chart or a table and by national, state, or county levels for each of the 6 indicators. Additional filters can be applied by age, race or ethnicity, sex or gender, and transmission category. Examples of data charts and tables are illustrated in the context of AHEAD tools for facilitating insights into the data.









## Similar Jurisdiction Functionality

The “Show Similar Jurisdiction” functionality is based on a “nearest neighbor” data analysis; in this case, distance is not geographical distance, but rather based on a distance matrix computer science algorithm [17], which identifies the 3 most similar EHE counties or areas based on a given jurisdiction’s EHE indicator data points. The purpose of using a nearest neighbor analysis is to examine progress across all 6 indicators and identify other jurisdictions that are quantitatively most similar to the selected jurisdiction. These similar jurisdictions may be interested in communicating with and exploring similar strategies to better achieve EHE goals.

## Implementing EHE Goals at the Jurisdictional Level

Although EHE goals are overall national goals [4], achieving national goals will not be possible unless most or all of the EHE jurisdictions meet these same goals at the local or state level. The focus on jurisdictional goals is also important from the perspective of health equity [18]. Therefore, it is useful to apply the EHE goals to the jurisdictional level. For HIV incidence, the jurisdiction-specific goals are derived based on the baseline levels of incidence in the jurisdiction and are therefore specific to each jurisdiction. Goals around care continuum indicators are proportional in both national and subnational jurisdictions (Figure 4).

**Figure 4.** Ending the HIV Epidemic goals and indicators at the national and jurisdictional levels, United States, through 2030.

Goal (decrease or increase)	Indicator	By 2030...	National	State/Territory/County/Area
	 <b>Incidence</b>	<b>By 90%</b>	<b>3000 people</b>	<b>Numeric goal is specific to the jurisdiction</b>
	 <b>Knowledge of Status</b>	<b>By 90%</b>	<b>3000 people</b>	<b>Numeric goal is specific to the jurisdiction</b>
	 <b>Diagnoses</b>	<b>To 95%</b>	<b>95%</b>	<b>95%</b>
	 <b>Linkage to HIV Medical Care</b>	<b>To 95%</b>	<b>95%</b>	<b>95%</b>
	 <b>Viral Suppression</b>	<b>To 95%</b>	<b>95%</b>	<b>95%</b>
	 <b>PrEP Coverage</b>	<b>To 50%</b>	<b>50%</b>	<b>50%</b>

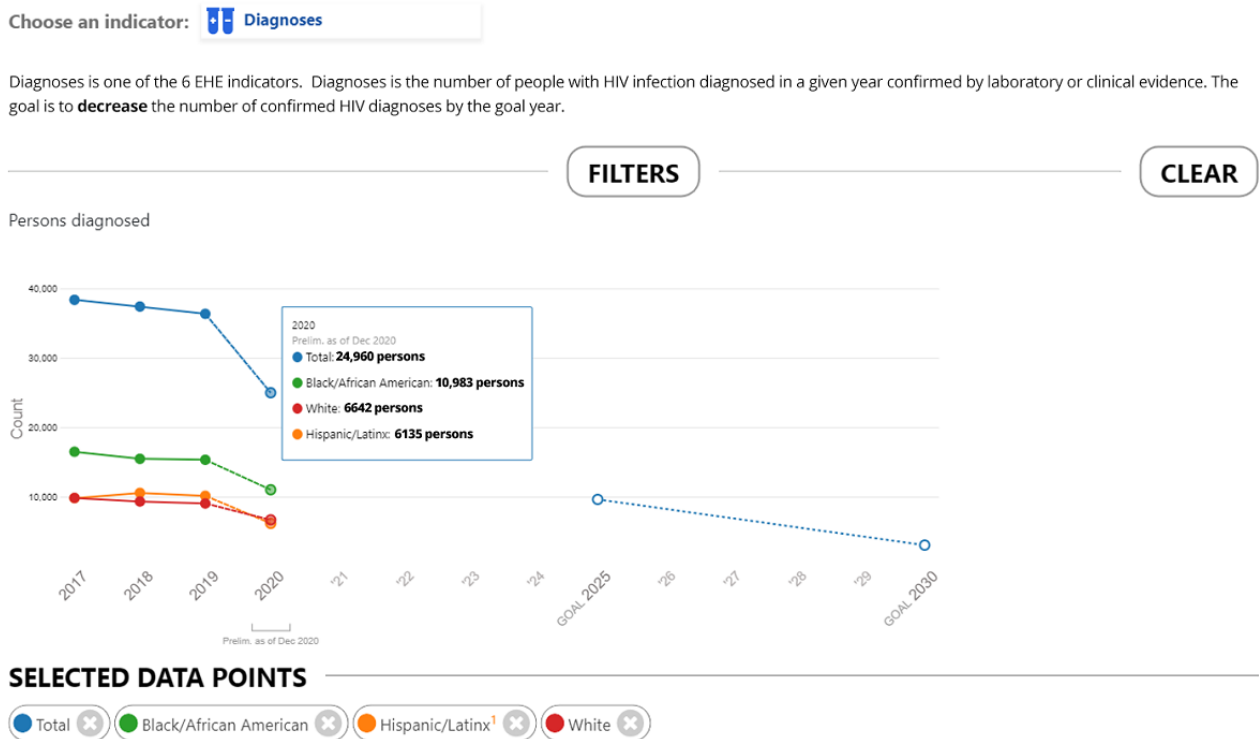
The methods for deriving individual indicators from national data sources have been described previously, and our implementation of those methods for AHEAD are summarized in [Multimedia Appendix 2](#). AHEAD adds value to the existing data sources by illustrating the annual trends for each indicator by jurisdiction and by displaying the historical trends in relation

to the 2025 and 2030 goals ([Figure 5](#)). At some geographic levels and for some indicators, stratified data are available to show progress in important subgroups (eg, groups by race or ethnicity; [Figure 6](#)). In some cases, not all indicators are available in all jurisdictions ([Figure 7](#)).

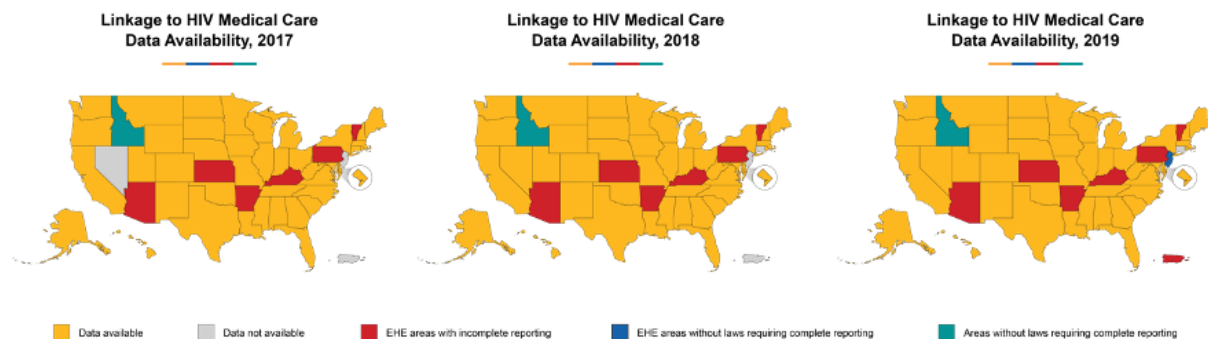
**Figure 5.** HIV incidence by year, United States, 2017-2019, and Ending the HIV Epidemic goals for 2025 and 2030.



**Figure 6.** HIV Diagnoses overall and among Black, White, and Hispanic people, by year of diagnosis, United States, 2017-2020, and 2025 and 2030 EHE goals; the 2020 data displayed in this figure are preliminary as of December 2020. Interpret preliminary 2020 data with caution as the impact of COVID-19 on 2020 data year is not known to date. EHE: Ending the HIV Epidemic.



**Figure 7.** Linkage to HIV medical care data availability, 2017-2019. EHE: Ending the HIV Epidemic.



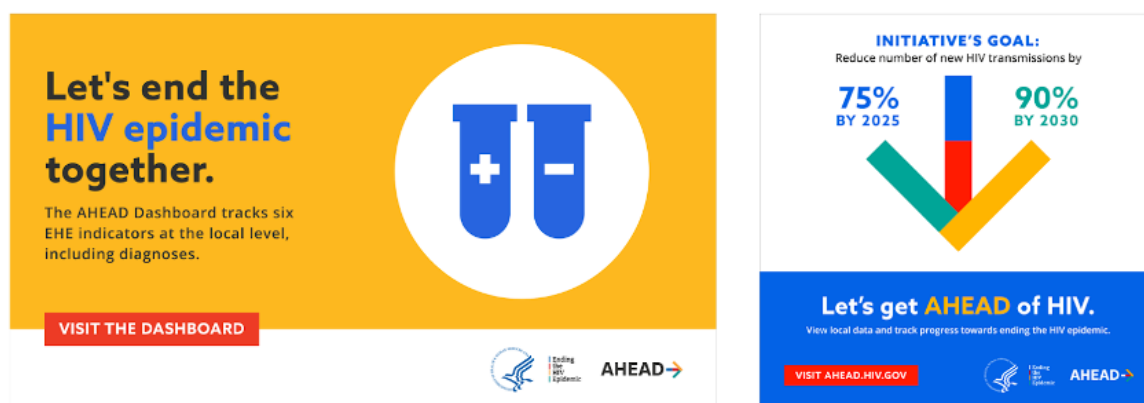
**Downloadable Materials**

Data are available for download in the following 3 formats: active data (CSV), all data (Excel spreadsheet), or by chart (PNG). Additional materials include a 1-page document outlining the features of AHEAD, an AHEAD user guide, as well as a comprehensive toolkit for jurisdictions, community-based organizations, advocacy organizations, and federal agencies (including sample tweets, Facebook posts, Instagram posts, e-blasts, and newsletter blurbs).

**Infographics**

Infographics (Figure 8) are graphical panels that can be designed to illustrate key messages of the EHE plan, emphasize EHE goals, or illustrate specific data elements in ways that might increase access to or comprehension of information. Infographics for AHEAD are optimized for dissemination through social media channels and are freely available for download on the site.

**Figure 8.** Examples of AHEAD infographics optimized for social media dissemination. AHEAD: America's HIV Epidemic Analysis Dashboard; EHE: Ending the HIV Epidemic.



### Other Resources

AHEAD provides users access to other valuable sites and resources including HIV.gov's site Ready, Set, PrEP [19], notice of funding opportunity announcements, and a locator tool that provides geo-located HIV testing, PrEP, and care services sites. In addition, blog posts are prepared to provide information about how to use and interpret the data on AHEAD. To date, HIV.gov has released the following series of blog posts related to AHEAD:

1. AHEAD Dashboard - EHE Indicators: PrEP Coverage
2. AHEAD Dashboard and Stakeholders
3. AHEAD Dashboard Enhances Functionality
4. AHEAD Dashboard Updated to Reflect 2020 Data on Two Key Indicators
5. AHEAD Dashboard: Launched
6. AHEAD Dashboard: Understanding the Sources of EHE Indicator Data
7. AHEAD Informational Webinar Materials Available
8. AHEAD: Monitoring Progress Toward Achieving the Nation's Viral Suppression Goals
9. AHEAD: Tracking Linkage to HIV Medical Care Data to End the HIV Epidemic
10. Coming Soon: The AHEAD Dashboard
11. Ending the HIV Epidemic with Data: Why AHEAD Data Matters
12. Ending the HIV Epidemic: 6 Indicators. 4 Strategies. 1 Goal - The AHEAD Dashboard
13. HOPWA Resource Tool Helps Jurisdictions Plan for and Evaluate Housing Needs
14. New HIV Indicator Data on AHEAD
15. Sign-Up for Free TA on AHEAD
16. The Impact of Incidence Data on Ending the HIV Epidemic
17. What You Need to Know About the 6 EHE Indicators: The AHEAD Dashboard

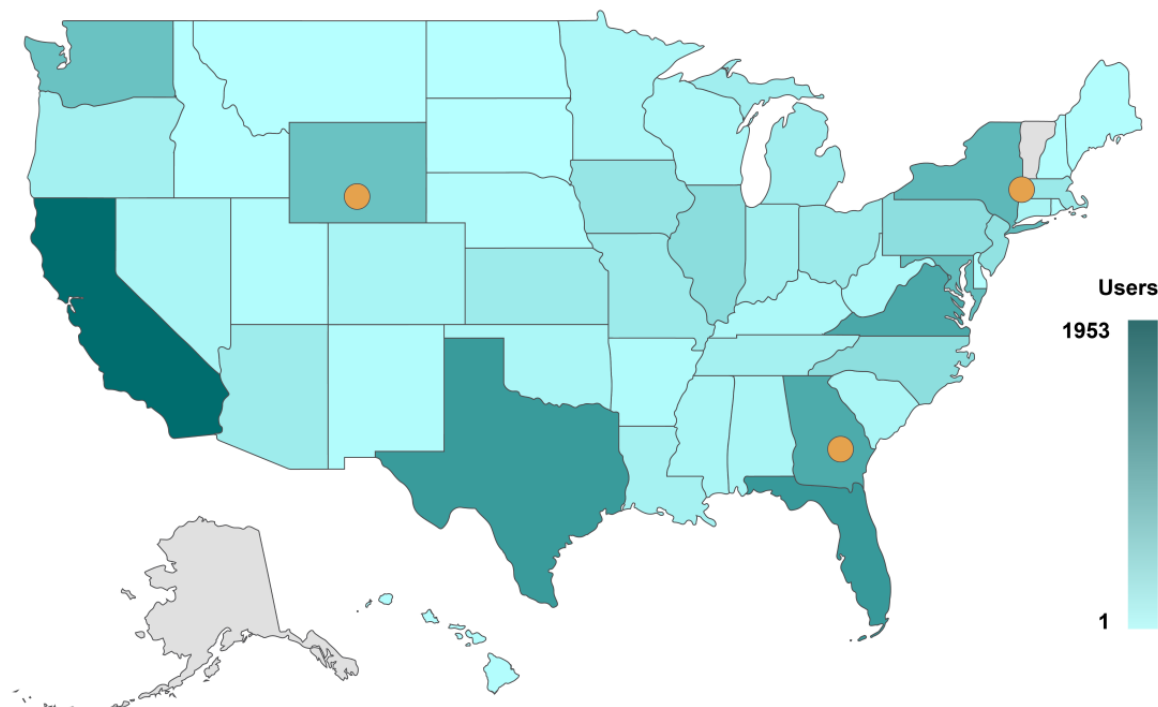
## Results

### Website Sessions and Users

From August 2020 to July 2021, on average, AHEAD garnered approximately 2646 sessions per month. In total, AHEAD has received 29,110 sessions since August 2020. The average session duration was 2 minutes and 18 seconds, and users are engaging with 2.35 pages per session, indicating the AHEAD audience is engaging with the content. During the same period, there have been approximately 1013 downloads from the AHEAD site of the materials available to users indicating AHEAD's audience is using the resources provided on the site. There were 6932 (23.81%) returning visitors; thus, about one-quarter of the users visited the site more than once. The top locations that are being selected within the data page are Alabama, California, Florida, Georgia, and Alabama. The states with the highest rates of engagement are California, Florida, and Georgia (Figure 9). By channel, most sessions (11,120/29,110, 38.2%) were referred from Google/organic. Moreover, 9636 (33.1%) out of 29,110 sessions were referred from Bing/organic; 5938 (20.4%) out of 29,110 sessions were referred from other search engines; and 2416 (8.3%) out of 29,110 sessions were accessed directly through the site URL.

**Figure 9.** AHEAD (America's HIV Epidemic Analysis Dashboard) users by location and number of users per jurisdiction, August 2020 to July 2021.

### Users by Location



### Uses by Media

AHEAD can also be used by media channels as a resource to increase the awareness of the HIV epidemic and the EHE plan for America. For example, in July 2021, Edge Media Network reported on AHEAD as a new site that tackles HIV at the local level [20]. A local network news affiliate used data from the site to support a study about the HIV epidemic in Florida [21]. Additional uses of the data include a story by a national online news site focusing on LGBT (lesbian, gay, bisexual, and transgender) health, which highlighted the more granular county-level data on AHEAD and described the site as an important new tool to tackle the epidemic at a local level.

### Discussion

Ambitious public health programs are unlikely to be successful without transparent, up-to-date, and accessible tools for monitoring public progress at each jurisdictional level responsible for public health response. In the case of the EHE plan, national goals for ending the epidemic will be difficult or impossible to achieve unless there is substantial progress toward the goals in each of the 57 EHE jurisdictions. For these reasons, AHEAD was created to supplement existing data visualization tools to (1) provide focus on the EHE jurisdictions; (2) tailor data elements to align with EHE indicators; (3) allow visualization of 2025 and 2030 goals along with data on progress

of individual jurisdictions; and (4) allow jurisdictions to see their progress and to compare their progress to peer jurisdictions. AHEAD contextualizes where we currently stand, the progress being made, and how close we are to achieving our goals.

A previous analysis of the uses of data from an HIV dashboard [7] suggested that the examination of local data can lead directly to public health responses that are informed by the geographic areas or populations at highest risks for poor outcomes [8]. Prior uses of comparative geographic data on morbidity and service locations have led to the development of programs that geographically focus on new resources for prevention or care services [8]. Thus, we propose 3 ways that jurisdictions and stakeholders can use AHEAD to inform their local planning and advocacy to drive toward meeting EHE goals: evaluation of progress toward EHE goals, examination of stratified analyses to identify subgroups of people who might need enhanced services, and comparison with peer jurisdictions to promote tailored peer-to-peer sharing of best practices and the development of peer networks.

Ideally, jurisdictions will use the data and features in AHEAD to evaluate their progress periodically and assess whether their pace of progress is sufficient to meet 2025 and 2030 goals. Even in the absence of more formal analysis or statistical inference, the inspection of Figure 5 and extrapolation of the trends in HIV incidence from 2017-2019 suggest that the current pace of reductions in incidence will not lead to sufficient reductions

to meet 2025 or 2030 goals. The examination of the continuum indicators might provide insight into more specific interim goals and priority populations for programmatic intervention.

Second, AHEAD provides online tools to evaluate progress toward individual EHE goals in stratified groups of people. For example, a visualization of PrEP coverage for persons with a PrEP indication ([Multimedia Appendix 3](#)) indicates that declines in PrEP coverage overall in 2020 compared to 2019 were larger for 16- to 24-year-olds than for older persons. This observation should lead to hypotheses about why PrEP coverage differentially decreased in 2020 for younger people (eg, COVID-19 impacts and lower autonomy in accessing health care providers) and consideration of programs to facilitate access to young people who might benefit from PrEP.

Third, AHEAD offers the opportunity for local public health officials to compare their jurisdiction to other similarly positioned jurisdictions. These comparisons can be made based on prior knowledge of jurisdictions with similar demographics or policy settings or can be identified empirically using built-in tools to identify the underlying similarities in factors that might shape risks and effectiveness of public health interventions. This allows for areas to identify “peer” jurisdictions, with whom these peers can discuss challenges, and what solutions have been effective for them. For example, a jurisdiction experiencing slow progress in a particular indicator could look for another county or area that has been successful in that indicator and contact that jurisdiction to brainstorm possible solutions together.

The data used in AHEAD are all population-based, and the underlying HIV surveillance system [12] is routinely evaluated according to standard criteria for completeness, timeliness, and other key performance metrics [22]. However, surveillance data on HIV diagnoses are likely minimal counts of people with HIV, because not everyone who acquires HIV is tested and diagnosed, and some states allow anonymous testing. People tested for HIV anonymously will not be reflected in the NHSS data until confidential confirmatory testing is completed and

the information reported to CDC. Similarly, PrEP use data that are developed using a commercial pharmacy data source, which comprises about 80% to 92% of US prescriptions filled in a commercial pharmacy; however, these data are subject to bias if the patrons of included pharmacies differ from those of excluded pharmacies [16,23]. Finally, users should be aware that the data displayed on AHEAD for specific jurisdictions might not be the same as the data reported in local surveillance reports. This is because CDC deduplicates case reports when one individual is reported by multiple jurisdictions [24], and because CDC data are statistically adjusted to assign transmission categories to people reported with missing risks for HIV acquisition [25].

The initial version of AHEAD displays baseline data for the nation and jurisdictions (where data are available) for each of the 6 indicators. In 2021, additional features and information about the EHE initiative will be added to the site. As AHEAD evolves, the site will become more interactive, integrate more data sources, and provide both national and jurisdictional leaders with the tools to make decisions on resources and interventions in their communities. This site will also serve as a primary site to share information about the EHE progress, interventions, and success stories from the field.

Surveillance data are meant to drive public health action [26]. HIV surveillance data have been described as “the conscience of the HIV epidemic” [27]. As we look to the initial assessment of EHE goals in 2025, we must develop, deploy, and promote tools that help make high-quality data accessible to stakeholders in all aspects of ending the HIV epidemic—providers, public health professionals, community organizations, policymakers, researchers, academics, and advocates. AHEAD occupies a unique space among online HIV data tools in its focus on the EHE jurisdictions and its benchmarking to EHE indicators and initiative goals. Data visualization tools democratize data [28], and increasing the number of people interacting with data on the HIV epidemic can only enhance our efforts toward ending the epidemic.

---

## Acknowledgments

We would like to acknowledge the following who participated in the prelaunch stakeholder interviews and/or the limited user assessment interviews: Heather Bradley; Amanda Castel; Denis Nash; Jessica Fridge; Lauren Ostrenga; Debbie Wendell; Kirk Shull; Antony Price; Samira Khan; Pascale Wortley; Alisha Barrett; Jen Opoku; Sylvia Odem; Paul Kawata; Natalie Cramer; Jen Kates; Lindsay Dawson; Lucy Slater; Raniyah Copeland; Greg Millett; and Guillermo Chacon. We also acknowledge the contributions of Shane Oleson, Anna Satcher Johnson, and Norma Harris.

---

## Conflicts of Interest

None declared.

---

### Multimedia Appendix 1

Details on demographic groups used for stratification in AHEAD (America’s HIV Epidemic Analysis Dashboard).

[[DOCX File, 14 KB - publichealth\\_v8i2e33522\\_app1.docx](#) ]

---

### Multimedia Appendix 2

Methods for calculation of indicators.

[[DOCX File, 18 KB - publichealth\\_v8i2e33522\\_app2.docx](#) ]



## Multimedia Appendix 3

Stratified analysis of PrEP (pre-exposure prophylaxis coverage) illustrating the importance of stratified data visualization.  
[PNG File , 172 KB - [publichealth\\_v8i2e33522\\_app3.png](#) ]

**References**

1. Siddiqi A, Hall HI, Hu X, Song R. Population-Based Estimates of Life Expectancy After HIV Diagnosis: United States 2008-2011. *J Acquir Immune Defic Syndr* 2016 Jun 01;72(2):230-236 [FREE Full text] [doi: [10.1097/QAI.0000000000000960](#)] [Medline: [26890283](#)]
2. Sullivan PS, Carballo-Diéguez A, Coates T, Goodreau SM, McGowan I, Sanders EJ, et al. Successes and challenges of HIV prevention in men who have sex with men. *Lancet* 2012 Jul 28;380(9839):388-399 [FREE Full text] [doi: [10.1016/S0140-6736\(12\)60955-6](#)] [Medline: [22819659](#)]
3. Health, United States. National Center for Health Statistics. 2019. URL: <https://www.cdc.gov/nchs/data/hus/19-508.pdf> [accessed 2022-01-18]
4. Fauci AS, Redfield RR, Sigounas G, Weahkee MD, Giroir BP. Ending the HIV Epidemic: A Plan for the United States. *JAMA* 2019 Mar 05;321(9):844-845. [doi: [10.1001/jama.2019.1343](#)] [Medline: [30730529](#)]
5. Murray CJ. Towards good practice for health statistics: lessons from the Millennium Development Goal health indicators. *Lancet* 2007 Mar 10;369(9564):862-873 [FREE Full text] [doi: [10.1016/S0140-6736\(07\)60415-2](#)] [Medline: [17350457](#)]
6. Atlas Plus. Centers for Disease Control and Prevention. 2019. URL: <https://www.cdc.gov/nchhstp/atlas/> [accessed 2020-11-12]
7. Sullivan PS, Woodyatt C, Koski C, Pembleton E, McGuinness P, Taussig J, et al. A Data Visualization and Dissemination Resource to Support HIV Prevention and Care at the Local Level: Analysis and Uses of the AIDSvu Public Data Resource. *J Med Internet Res* 2020 Oct 23;22(10):e23173 [FREE Full text] [doi: [10.2196/23173](#)] [Medline: [33095177](#)]
8. Valdiserri RO, Sullivan PS. Data Visualization Promotes Sound Public Health Practice: The AIDSvu Example. *AIDS Educ Prev* 2018 Feb;30(1):26-34. [doi: [10.1521/aeap.2018.30.1.26](#)] [Medline: [29481299](#)]
9. Kitahata MM, Gange SJ, Abraham AG, Merriman B, Saag MS, Justice AC, NA-ACCORD Investigators. Effect of early versus deferred antiretroviral therapy for HIV on survival. *N Engl J Med* 2009 Apr 30;360(18):1815-1826 [FREE Full text] [doi: [10.1056/NEJMoa0807252](#)] [Medline: [19339714](#)]
10. HIV National Strategic Plan for the United States: A Roadmap to End the Epidemic 2021-2025. US Department of Health and Human Services. 2021. URL: <https://files.hiv.gov/s3fs-public/HIV-National-Strategic-Plan-2021-2025.pdf> [accessed 2022-01-23]
11. Goldenberg T, McDougal SJ, Sullivan PS, Stekler JD, Stephenson R. Building a Mobile HIV Prevention App for Men Who Have Sex With Men: An Iterative and Community-Driven Process. *JMIR Public Health Surveill* 2015;1(2):e18 [FREE Full text] [doi: [10.2196/publichealth.4449](#)] [Medline: [27227136](#)]
12. Cohen SM, Gray KM, Ocfemia MCB, Johnson AS, Hall HI. The status of the National HIV Surveillance System, United States, 2013. *Public Health Rep* 2014;129(4):335-341 [FREE Full text] [doi: [10.1177/003335491412900408](#)] [Medline: [24982536](#)]
13. Nakashima AK, Fleming PL. HIV/AIDS Surveillance in the United States, 1981-2001. *JAIDS Journal of Acquired Immune Deficiency Syndromes* 2003;32:S68-S85. [doi: [10.1097/00126334-200302011-00011](#)]
14. Smith DK, Van Handel M, Grey J. Estimates of adults with indications for HIV pre-exposure prophylaxis by jurisdiction, transmission risk group, and race/ethnicity, United States, 2015. *Ann Epidemiol* 2018 Dec;28(12):850-857.e9. [doi: [10.1016/j.annepidem.2018.05.003](#)] [Medline: [29941379](#)]
15. Core Indicators for Monitoring EHE (Preliminary Data): HIV Diagnoses and Linkage to HIV Medical Care, 2019 (Reported through December 2019); and Preexposure Prophylaxis (PrEP)? 2018, Updated. Centers for Disease Control and Prevention. 2020. URL: <https://www.cdc.gov/hiv/library/reports/surveillance-data-tables/vol-1-no-2/index.html> [accessed 2021-08-27]
16. Sullivan P, Giler R, Mouhanna F, Pembleton E, Guest JL, Jones J, et al. Trends in the use of oral emtricitabine/tenofovir disoproxil fumarate for pre-exposure prophylaxis against HIV infection, United States, 2012-2017. *Ann Epidemiol* 2018 Dec;28(12):833-840 [FREE Full text] [doi: [10.1016/j.annepidem.2018.06.009](#)] [Medline: [30037634](#)]
17. Zapala MA, Schork NJ. Statistical properties of multivariate distance matrix regression for high-dimensional data analysis. *Front Genet* 2012;3:190 [FREE Full text] [doi: [10.3389/fgene.2012.00190](#)] [Medline: [23060897](#)]
18. Wood BR, Young JD, Abdel-Massih RC, McCurdy L, Vento TJ, Dhanireddy S, et al. Advancing Digital Health Equity: A Policy Paper of the Infectious Diseases Society of America and the HIV Medicine Association. *Clin Infect Dis* 2021 Mar 15;72(6):913-919 [FREE Full text] [doi: [10.1093/cid/ciaa1525](#)] [Medline: [33033829](#)]
19. Ready, Set, PrEP. HIV.gov. URL: <https://www.hiv.gov/federal-response/ending-the-hiv-epidemic/prep-program> [accessed 2022-01-26]
20. Melloy K. New government initiative tackles HIV at local level. *Edge Media Network*. 2021 Jul 27. URL: [https://www.edgemedianetwork.com/story.php?ch=news&sc=aids&id=307332&new\\_govt\\_initiative\\_tackles\\_hiv\\_at\\_local\\_level](https://www.edgemedianetwork.com/story.php?ch=news&sc=aids&id=307332&new_govt_initiative_tackles_hiv_at_local_level) [accessed 2021-08-20]
21. Axelbank E. Florida leads nation in new AIDS cases. *FOX 13 Tampa Bay*. 2021 Jun 14. URL: <https://www.fox13news.com/news/florida-leads-nation-in-new-aids-cases.amp> [accessed 2021-08-27]

22. German RR, Lee LM, Horan JM, Milstein RL, Pertowski CA, Waller MN, Guidelines Working Group Centers for Disease Control/Prevention (CDC). Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. *MMWR Recomm Rep* 2001 Jul 27;50(RR-13):1-35; quiz CE1. [Medline: [18634202](#)]
23. Lyons SJ, Gant Z, Johnson AS, Hu X, Yu C, Jin C, et al. Monitoring selected national HIV prevention and care objectives by using HIV surveillance data: United States and 6 dependent areas, 2018. Centers for Disease Control and Prevention. 2021. URL: <https://stacks.cdc.gov/view/cdc/107935> [accessed 2021-09-21]
24. Glynn MK, Ling Q, Phelps R, Li J, Lee LM. Accurate monitoring of the HIV epidemic in the United States: case duplication in the national HIV/AIDS surveillance system. *J Acquir Immune Defic Syndr* 2008 Mar 01;47(3):391-396. [doi: [10.1097/QAI.0b013e318160d52a](#)] [Medline: [18176325](#)]
25. Harrison KM, Kajese T, Hall HI, Song R. Risk factor redistribution of the national HIV/AIDS surveillance data: an alternative approach. *Public Health Rep* 2008;123(5):618-627 [FREE Full text] [doi: [10.1177/003335490812300512](#)] [Medline: [18828417](#)]
26. Lee LM, Teutsch SM, Thacker SB, St. Louis ME. *Principles and Practice of Public Health Surveillance*. New York, US: Oxford University Press; 2000.
27. Sullivan PS, Delpech V. HIV Surveillance: The conscience of the epidemic. *Open AIDS J* 2012;6:65-66 [FREE Full text] [doi: [10.2174/1874613601206010065](#)] [Medline: [23049654](#)]
28. Sullivan PS, Satcher Johnson A, Pembleton ES, Stephenson R, Justice AC, Althoff KN, et al. Epidemiology of HIV in the USA: epidemic burden, inequities, contexts, and responses. *Lancet* 2021 Mar 20;397(10279):1095-1106. [doi: [10.1016/S0140-6736\(21\)00395-0](#)] [Medline: [33617774](#)]

## Abbreviations

**AHEAD:** America's HIV Epidemic Analysis Dashboard  
**CDC:** Centers for Disease Control and Prevention  
**EHE:** Ending the HIV Epidemic  
**LGBT:** lesbian, gay, bisexual, and transgender  
**NHSS:** National HIV Surveillance System  
**PrEP:** pre-exposure prophylaxis coverage

*Edited by G Eysenbach; submitted 08.10.21; peer-reviewed by R Hart-Malloy; comments to author 01.11.21; accepted 21.11.21; published 10.02.22.*

*Please cite as:*

Sullivan PS, Woodyatt CR, Kouzouian O, Parrish KJ, Taussig J, Conlan C, Phillips H

America's HIV Epidemic Analysis Dashboard: Protocol for a Data Resource to Support Ending the HIV Epidemic in the United States  
*JMIR Public Health Surveill* 2022;8(2):e33522

URL: <https://publichealth.jmir.org/2022/2/e33522>

doi: [10.2196/33522](#)

PMID: [35142639](#)

©Patrick Sean Sullivan, Cory R Woodyatt, Oskian Kouzouian, Kristen J Parrish, Jennifer Taussig, Chris Conlan, Harold Phillips. Originally published in *JMIR Public Health and Surveillance* (<https://publichealth.jmir.org>), 10.02.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

[Open Source/Open Data](#)

# Physical Activity, Sedentary Behavior, and Sleep on Twitter: Multicountry and Fully Labeled Public Data Set for Digital Public Health Surveillance Research

Zahra Shakeri Hossein Abad<sup>1,2</sup>, PhD; Gregory P Butler<sup>3</sup>, MSc; Wendy Thompson<sup>3</sup>, MSc; Joon Lee<sup>2,4,5</sup>, PhD

<sup>1</sup>Department of Biomedical Informatics, Harvard Medical School, Harvard University, Boston, MA, United States

<sup>2</sup>Data Intelligence for Health Lab, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

<sup>3</sup>Centre for Surveillance and Applied Research, Public Health Agency of Canada, Ottawa, ON, Canada

<sup>4</sup>Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

<sup>5</sup>Department of Cardiac Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

## Corresponding Author:

Joon Lee, PhD

Data Intelligence for Health Lab

Cumming School of Medicine

University of Calgary

3280 Hospital Dr NW

Calgary, AB, T2N 4Z6

Canada

Phone: 1 403 220 2968

Email: [joonwu.lee@ucalgary.ca](mailto:joonwu.lee@ucalgary.ca)

## Abstract

**Background:** Advances in automated data processing and machine learning (ML) models, together with the unprecedented growth in the number of social media users who publicly share and discuss health-related information, have made public health surveillance (PHS) one of the long-lasting social media applications. However, the existing PHS systems feeding on social media data have not been widely deployed in national surveillance systems, which appears to stem from the lack of practitioners and the public's trust in social media data. More robust and reliable data sets over which supervised ML models can be trained and tested reliably is a significant step toward overcoming this hurdle. The health implications of daily behaviors (physical activity, sedentary behavior, and sleep [PASS]), as an evergreen topic in PHS, are widely studied through traditional data sources such as surveillance surveys and administrative databases, which are often several months out-of-date by the time they are used, costly to collect, and thus limited in quantity and coverage.

**Objective:** The main objective of this study is to present a large-scale, multicountry, longitudinal, and fully labeled data set to enable and support digital PASS surveillance research in PHS. To support high-quality surveillance research using our data set, we have conducted further analysis on the data set to supplement it with additional PHS-related metadata.

**Methods:** We collected the data of this study from Twitter using the Twitter livestream application programming interface between November 28, 2018, and June 19, 2020. To obtain PASS-related tweets for manual annotation, we iteratively used regular expressions, unsupervised natural language processing, domain-specific ontologies, and linguistic analysis. We used Amazon Mechanical Turk to label the collected data to self-reported PASS categories and implemented a quality control pipeline to monitor and manage the validity of crowd-generated labels. Moreover, we used ML, latent semantic analysis, linguistic analysis, and label inference analysis to validate the different components of the data set.

**Results:** LPHEADADA (Labelled Digital Public Health Dataset) contains 366,405 crowd-generated labels (3 labels per tweet) for 122,135 PASS-related tweets that originated in Australia, Canada, the United Kingdom, or the United States, labeled by 708 unique annotators on Amazon Mechanical Turk. In addition to crowd-generated labels, LPHEADADA provides details about the three critical components of any PHS system: place, time, and demographics (ie, gender and age range) associated with each tweet.

**Conclusions:** Publicly available data sets for digital PASS surveillance are usually isolated and only provide labels for small subsets of the data. We believe that the novelty and comprehensiveness of the data set provided in this study will help develop,

evaluate, and deploy digital PASS surveillance systems. LPHEADA will be an invaluable resource for both public health researchers and practitioners.

(*JMIR Public Health Surveill* 2022;8(2):e32355) doi:[10.2196/32355](https://doi.org/10.2196/32355)

## KEYWORDS

digital public health surveillance; social media analysis; physical activity; sedentary behavior; sleep; machine learning; online health information; infodemiology; public health database

## Introduction

### Digital Public Health Surveillance

Almost two-thirds of the world's population now uses the internet, taking the global total to 4.57 billion (59%) by July 2020 [1]. Overall, 87% of internet users and 65% (3.96 billion) of the world's total eligible population (ie, aged >13 years) now use social media. The combined time that these users spend on social media adds up to more than 1 million years every day [1], contributing to a large amount of user-generated data on different social media platforms. In 2020, Twitter alone reported 500 million tweets generated per day from 145 million daily active users. The low-cost data stream available on social media and other internet-based sources serves to make research advances on digital public health surveillance (DPHS) more accessible for public health officials, clinicians, patients, and the public. This helps disseminate insights into different aspects of public health and promote healthy lifestyles and health policies [2,3]. The open access to the public data about users and their opinions, the ease of use, and a large user base have made Twitter one of the most popular data sources for studying different aspects of public health [4,5], with Google Scholar indexing 1.32 million articles mentioning Twitter and public health. Moreover, more than 85% of Twitter users, with a wide breadth of demographic groups [4], also use Facebook, Instagram, and YouTube (this number for other platforms varies between 52% and 82%) [1], indicating that Twitter users reasonably represent active social media users in general.

Since 2011, Twitter has been the most popular form of social media used for public health communication [6,7]. A recent scoping review of 755 articles on DPHS shows that Twitter is the most studied of all platforms and most used platform to study communicable diseases, behavioral risk factors, mental health, drug use, and vaccine [7].

### Limitations of Digital Public Health Data

However, a number of limitations that mainly stem from the limitations associated with the data are still the major obstacles toward the adoption of digital data for public health surveillance (PHS) [4,7]. Given the main aims of any PHS system are to measure, monitor, and improve the overall health status of their target populations, the systematic incorporation of time, demographics (ie, age and gender), and place data into the surveillance process is critical to the reliability and generalizability of this process [8,9]. However, nearly one-third (32%) of the DPHS studies published between 2005 and 2020 (with the majority of them related to behavioral risk factors surveillance) did not capture age, gender, or place information for their analyses [7]. Moreover, most studies on DPHS do not

consider whether their findings are associated with the user's personal experience (self-reported or not), leading to content bias, incorrect results, and potentially flawed interpretations [7].

Considering the location-dependent nature of health policies, along with the essential role of place data in assessing the representativeness of a PHS system, the impact of a PHS system can vary considerably with geographical location [10-13]. However, the number of DPHS studies that have stratified their results by a more granular geographic region is small [7]. Because of a lack of annotated data sets for the development of automatic models, more than two-thirds (69%) of DPHS studies published before 2020 are limited by labor-intensive, manual, and abstract analysis methods (eg, manual coding, qualitative analysis, and rule-based natural language processing [NLP]), which makes these studies limited in terms of sample size, scope, and generalizability [7].

Given that all of these challenges are data-oriented, an increase in both data quality and quantity enriched with concrete demographics and location information can help deal with all these challenges. Moreover, to facilitate the development and evaluation of robust machine learning (ML) models to address the limited scope of manual data analysis techniques, annotated data sets for various PHS aspects are required. However, only a handful of annotated data sets are publicly available for research on DPHS [14-21]. Jimeno-Yepes et al [15] provided an annotated data set of 1300 tweets related to disease symptoms and pharmacologic substances. The open data set developed by Aphinyanaphongs et al [16] contains 13,146 labeled tweets resulting from hashtag filtering and covers a time span from January 2010 to January 2015. This data set is developed for training binary classifiers to detect tweets that indicate e-cigarette use for smoking cessation. Crowdbreaks [18], an open health tracking platform, crowdsources the labeling of vaccine sentiment and COVID-19-related tweets to the public. Although the data set provided by this system, compared with other open DPHS data sets, is in a better position in terms of size, it lacks demographics and geospatial data, and each tweet is labeled by only 1 annotator (without any control over their labeling quality).

### Objective

Given that in addition to physical inactivity, as the leading risk factor for noncommunicable diseases and premature death [22], prolonged sedentary behavior and inadequate sleep are also important risk factors for chronic diseases [23], this work presents a multicountry and fully labeled digital public health data set (LPHEADA [Labelled Public Health Dataset]) of tweets related to physical activity, sedentary behavior, and sleep

(PASS) that originated in Australia, Canada, the United Kingdom, or the United States. We selected these countries because they have some of the highest proportions of social media users in the world (Australia, 71%; Canada, 66%; the United Kingdom, 66%; the United States, 69%; and world, 51%) [1]. LPHEADA comprises 366,405 labels, labeled by 708 unique annotators on Amazon Mechanical Turk (AMT), for 122,135 unique tweets generated by 72,749 unique users between November 28, 2018, and June 19, 2020. AMT is a software service operated by Amazon that allows users (ie, requesters) to crowdsource work, broken into microtasks called human intelligence tasks (HITs), to a large number of workers who are compensated for each HIT completed [24]. As LPHEADA was collected and labeled in collaboration with the Public Health Agency of Canada to develop PASS indicators for the Canadian population, 80.83% (98,722/122,136) of the tweets included in our data set were collected from Canada. Tweets from the United States and the United Kingdom make up 8.35% (10,193/122,136) and 7.49% (9154/122,136) of the data set, respectively, and Australian tweets make up the remaining 3.33% (4067/122,136) of the data set.

Along with the labeled tweets, we provide detailed information on users' gender, age range, and geospatial information, whether the tweet was self-reported, and whether it was posted by an organization. We evaluated the quality of the data set and its labels using latent semantic analysis, linguistic analysis, ML models, and truth inference models. The data set we provide in this paper can be used to develop unsupervised or supervised ML models for digital PASS surveillance.

## Methods

### Collection and Preparation of the Data Set

We collected the data of this study from Twitter using the Twitter livestream application programming interface (API) between November 28, 2018, and June 19, 2020. The entire data set (ie, 1,902,980,841 tweets) was filtered for Canadian tweets potentially relevant to PASS. From 22,729,110 collected Canadian tweets, 0.46% (103,911/22,729,110) were selected using keywords and regular expressions related to PASS categories. To define the search strings and regular expressions, we used NLP techniques (eg, topic modeling, language modeling, and linguistic analysis) to detect latent word patterns relevant to PASS-related contexts. Moreover, we pilot-tested the labeling process first to validate the extracted keywords and iteratively updated the list of keywords for each category after manually reviewing the labels and the filtered tweets. [Multimedia Appendix 1](#) provides a complete list of the words used for generating regular expressions and filtering the data set. Each of these 103,911 tweets was labeled by 3 AMT workers, from which 95.01% (98,722/103,911) of tweets received 3 valid labels (ie, multiple or missing labels were invalid and rejected), with almost half of them related to physical activity. For the Canadian data set, 610 unique workers participated in our data labeling tasks and completed 103,911 HITs, from which 4.99% (5189/103,911) HITs were removed as they did not receive 3 valid answers. The majority of these workers (530/610, 86.9%) completed less than 100 HITs each,

among which 30.9% (164/530) completed only 1 HIT each. Among all 610 workers, 1 (0.2%) worker completed 21,801 HITs, and 3 (0.5%) workers completed between 5000 and 10,000 HITs.

In addition to the Canadian tweets, we filtered a random subset of the data set for tweets that originated in the United Kingdom, the United States, and Australia. This data set spans the same data collection period as the Canadian data set and contains 70,239 labels collected for 23,413 tweets (ie, 3 labels per tweet). Adding the data from these countries will provide an important epidemiological diversity that can be used for implementing comparative studies and evaluating the generalizability of the PASS surveillance models trained on the Canadian data set.

### Labeling Process

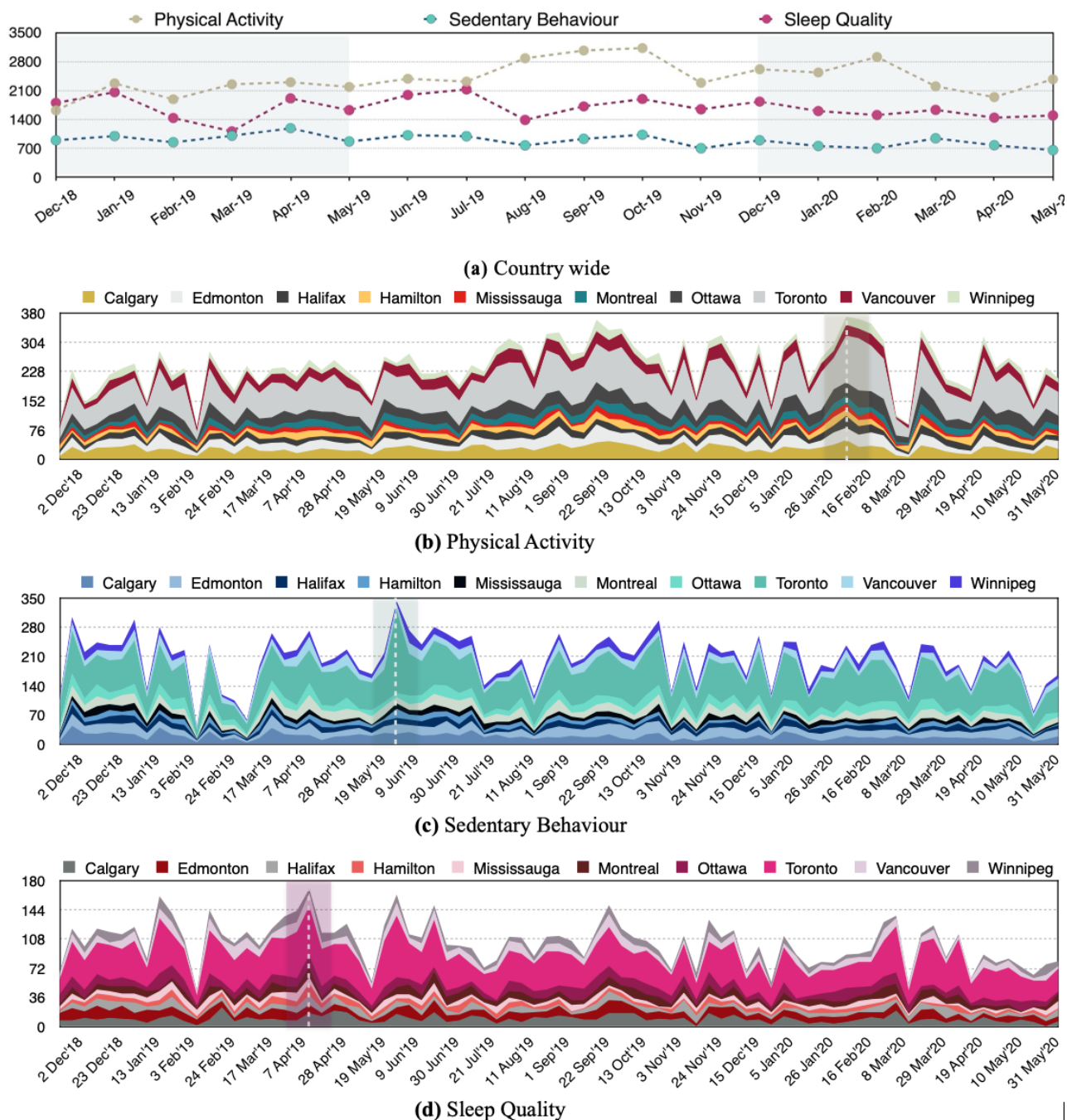
We implemented a pipeline to create the crowdsourcing tasks, referred to as HITs by AMT, post them on AMT, collect the labels through a quality-check process, approve or reject the HITs, and store the results. To minimize noisy and low-quality data, we added a qualification requirement to our tasks and granted the labeling access to workers who had demonstrated a high degree of success in performing a wide range of HITs across MTurk (ie, master qualification). In addition, we added a simple qualification question to each HIT to detect spammers or irresponsible workers. Each HIT contained 4 questions, including the qualification question, and was assigned to 3 workers. Through different iterations of data labeling, workers were paid from US \$0.03 to US \$0.05 after completing each HIT. We collected the labels for the 122,135 tweets used in this study through different iterations, from April 2019 to November 2020. We regularly checked the quality of the submitted tasks to detect low-quality workers during each iteration and revoke their access to our tasks. Before the formal initiation of the process, we pilot-tested the design, response time, and complexity of the HITs in 2 iterations and revised the workflow accordingly. To label the data set provided in this paper, we used AMT as a crowdsourcing service and did not collect any personally identifiable information from the workers (participants) during the data labeling task. The experiments were carried out in accordance with relevant guidelines and the University of Calgary Conjoint Faculties Research Ethics Board's regulations. We implemented the entire workflow in Python (Python Software Foundation) and used Boto3 Python Software Development Kit to connect to and work with AMT.

### Time Adjustment

The Twitter API returns the date and time that a tweet is published in the Universal Time Coordinated. To adjust this time zone based on each tweet's location, we used the bounding box of coordinates, which enabled spatial mapping to tweets' respective city locations, and used a time zone finder in Python. Given that daytime, month, and weekday can be influential factors in twitting about each of the PASS categories, and to better use the date-time data (%a %b %d %H:%M:%S %Y), we extracted weekday (a), month (b), and hour (H) fields and stored them as separate features. [Figure 1A](#) shows the temporal distribution of tweets for each of the PASS categories in the Canadian data set. Moreover, the stacked area charts presented

in Figures 1B-1D detail the frequency of tweets for each of PASS categories for the top 10 Canadian cities.

**Figure 1.** The temporal distribution of tweets for the Canadian data set. To make fair comparisons, we used data from December 1, 2018, to May 31, 2020, for these visualizations and removed data collected during the last 2 days of November 2018 and the first 2 weeks of June 2020.



### Location Inference

The geospatial metadata provided by the Twitter API are derived from three main sources: (1) geotagged location, (2) profile location, and (3) mentioned location in the tweet text. Geotagged location can be defined by exact location (ie, device location) at the time of tweeting, by the assigned Twitter place (ie, at the neighborhood level), or both. Although the exact location field provides the highest level of precision, it is not a default setting, and only a small portion of users share their exact latitude and longitude (eg, only 1%-2% of tweets are geotagged [25]). Thus, to infer the location of each tweet in LPHEADADA, we proposed

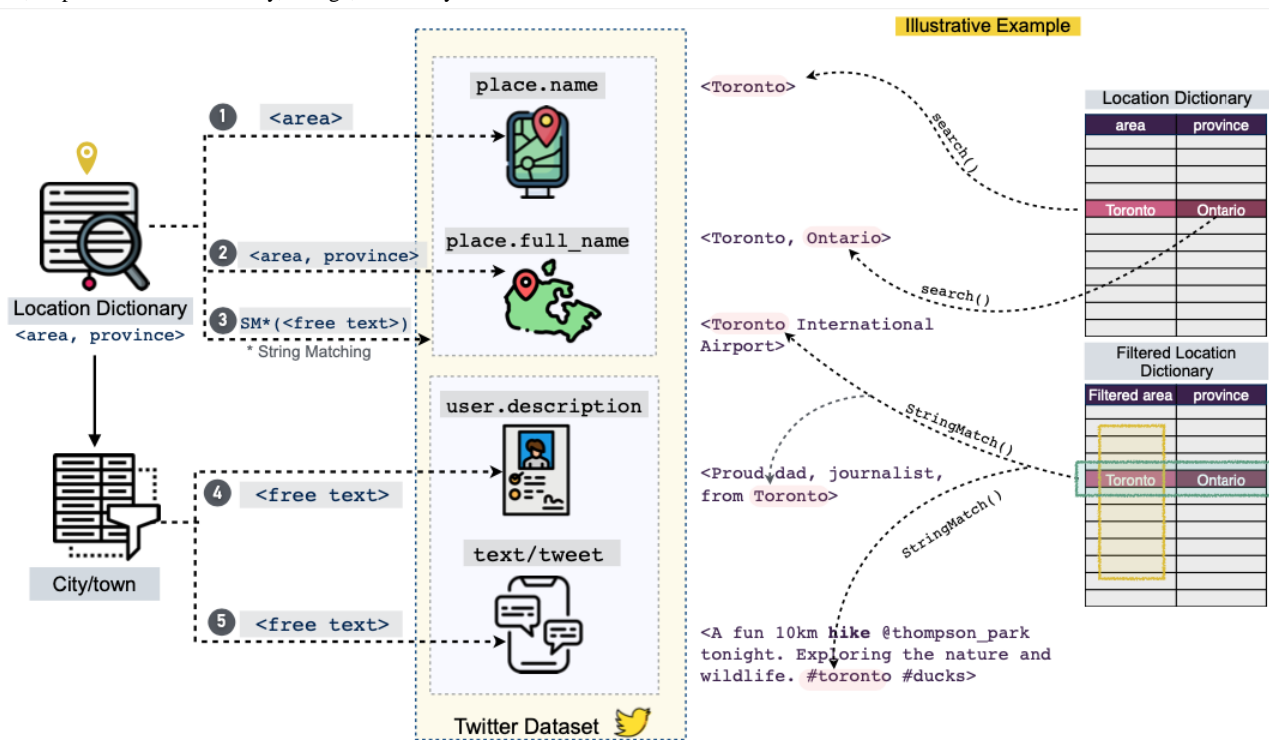
and developed a 5-step process that uses tweet-neighborhood location (ie, *place.name* and *place.full\_name*), profile information (ie, profile description and location), and tweet text to map Twitter’s geospatial metadata for each tweet to physical locations in the form of  $\langle c_i, p_j | s_i \rangle$ , where *c* denotes the city of a tweet and *p*/*s* represent its corresponding province or state, respectively (Figure 2). To demonstrate the proposed process, we used the Canadian geographical names data set (ie, location dictionary) provided by the Geographical Names Board of Canada. Each geographical name provided by this data set is mapped to a province and is classified to a geographic area such as city, town, village, lake, administrative sector, or recreational

center. As illustrated in Figure 2, for each  $t_i$ , we first used a simple search function to map the place name field associated with each tweet to its corresponding  $c_i$  in the location dictionary. If found, the corresponding province field  $p_i$  was defined using equation 1:

$$p_i = \text{second\_component}(c_i)$$

where  $\text{second\_component}(c_i)$  denotes the second component of the field when the first component is  $c_i$  (eg, Ontario in the illustrative example of Figure 2). This will detect geographical areas with the same names but in different provinces (eg, Leduc is a town or city in both Alberta and Quebec).

**Figure 2.** Five-step location inference process. The location dictionary is an external regional geographical metadata used to extract the exact locations of tweets. The area field in this process refers to regions at different scales such as city, region, municipality, town, township municipality, municipal district, dispersed rural community, village, or country.



### Demographic Attribute Inference

The demographic variables of age and gender and the information about the source of each tweet (eg, organizations vs real users) were not available within the data set collected from Twitter. We estimated these variables for each tweet using the M3inference package in Python [26], which uses a multimodal deep neural architecture for joint classification of age (binned into four groups: 18, 19-29, 30-39, and 40 years), gender, and information source of social media data. This approach uses 4 sources of information, namely, username, screen name, biography, and profile image of public profiles, to develop 2 separate pipelines for processing a profile image and each of the 3 text sources of information. The models provided in this package are trained on 14.53M, 2.61M, and 23.86M profiles for each of the gender, age, and organization categories, respectively.

Steps 3-5 of the process deal with unstructured text objects that can come from all 3 sources of geospatial information. To extract the location information from these fields, we developed a string-matching function to detect the longest common substring between the unstructured text of the data set and the area field of the location dictionary (eg, first time boating in Lake Louise #AB is mapped to <Lake Louise, Alberta> instead of to <Louise, Quebec>). To manage the complexity of information extraction from unstructured text, we only used a subset of areas listed in the location dictionary with high population density (eg, city, municipality, town, village, and country). Thus, we excluded areas classified as lake, mountain, river, bridge, or park.

## Results

### Data Records

#### Overview

LPHEADATA is released in accordance with Twitter’s terms and conditions, and the developer’s agreement and policies [27], which prohibits the verbatim release of the collected tweets. However, releasing the tweet IDs is allowed. Data access requires a data use agreement between the data user and Twitter to govern the access and use of the licensed material returned by the Twitter API. Once approved, using the Tweet ID field, Twitter metadata can be rehydrated and downloaded as a JSON (JavaScript Object Notation) file to be mapped to other subsets of data provided in this study (eg, labels, location, time, and demographics). A detailed and demonstrative tutorial on rehydration of the data set using tweet IDs is described on the GitHub page of the data set [28].

LPHEADADA comprises 366,405 labels for 122,135 unique tweets generated by 72,749 unique users between November 28, 2018, and June 19, 2020. This data set is organized into 12 subsets (3 PASS categories for each of the 4 countries). [Table 1](#) provides the demographics of the data set, including the number of tweets per PASS category for each country, labels' characteristics, and demographics characteristics of the users. Each unique tweet is

assigned a unique integer, known as TweetID. Each ID is mapped to the core Twitter metadata and to 3 crowd-generated labels for binary and multi-class classification tasks. [Figure 3](#) visualizes this hierarchy. As illustrated in this figure, for each labeled tweet, LPHEADADA provides the following data categories.



**Table 1.** Characteristics of the data set.<sup>a</sup>

Variable	Canada (N=98,722)			United States (N=10,193)			United Kingdom (N=9154)			Australia (N=4067)		
	PA <sup>b</sup>	SQ <sup>c</sup>	SB <sup>d</sup>	PA	SQ	PA	PA	SQ	SB	PA	SQ	SB
Tweets, n (%)	48,576 (49.2)	32,779 (33.2)	17,367 (17.59)	5053 (49.57)	3065 (30.07)	2074 (20.35)	4076 (44.53)	3001 (32.78)	2077 (22.69)	2216 (54.49)	1312 (32.26)	539 (13.25)
<b>Labels</b>												
<b>Binary, n (%)</b>												
Yes	15,337 (31.57)	11,814 (36.04)	6514 (37.51)	1196 (23.67)	1032 (33.67)	731 (35.25)	1092 (26.79)	766 (25.52)	356 (17.14)	729 (32.9)	499 (38.03)	75 (13.91)
No	33,239 (68.43)	20,965 (63.96)	10,853 (62.49)	3857 (76.33)	2033 (66.33)	1343 (64.75)	2984 (73.21)	2235 (74.48)	1721 (82.86)	1487 (67.1)	813 (61.97)	469 (87.01)
<b>Multi-class, n (%)</b>												
YY	17,298 (35.61)	12,818 (39.1)	7174 (41.31)	1431 (28.32)	1150 (37.52)	831 (40.07)	1227 (30.1)	934 (31.12)	468 (22.53)	846 (38.18)	566 (43.14)	97 (18)
YN	6583 (13.55)	7720 (23.55)	1895 (10.91)	1445 (28.6)	1104 (36.02)	623 (30.04)	1037 (25.44)	1017 (33.89)	916 (44.1)	905 (40.84)	404 (30.79)	147 (27.27)
NY	4407 (9.07)	2242 (6.84)	1471 (8.47)	634 (12.55)	215 (7.01)	147 (7.09)	226 (5.54)	332 (11.06)	118 (5.68)	208 (9.39)	72 (5.49)	33 (6.12)
NN	19,622 (40.39)	9339 (28.49)	6502 (37.44)	1512 (29.92)	564 (18.4)	469 (22.61)	1559 (38.25)	694 (23.13)	572 (27.54)	233 (10.51)	258 (19.66)	258 (47.87)
NC	666 (1.37)	660 (2.01)	325 (1.87)	31 (0.61)	32 (1.04)	10 (0.48)	27 (0.66)	24 (0.8)	3 (0.14)	24 (1.08)	12 (0.91)	4 (0.74)
<b>Users, n</b>												
Unique	22,601	16,984	11,490	4911	2994	2048	3810	2653	2002	1735	1004	517
Valid	21,772	14,919	10,912	3660	2614	1759	2614	2157	1840	1531	858	517
<b>Gender, n (%)</b>												
Female	8471 (38.91)	7270 (48.73)	4486 (41.11)	1448 (39.56)	1537 (58.8)	860 (48.89)	911 (34.85)	1068 (49.51)	625 (33.97)	520 (33.96)	401 (46.74)	205 (39.65)
Male	13,301 (61.09)	7649 (51.27)	6426 (58.89)	2212 (60.44)	1077 (41.2)	899 (51.11)	1703 (65.15)	1089 (50.49)	1215 (66.03)	1011 (66.04)	457 (53.26)	312 (60.35)
<b>Age range (years), n (%)</b>												
≤18	1772 (8.14)	2372 (15.9)	1361 (12.47)	490 (13.39)	518 (19.82)	318 (18.08)	136 (5.2)	248 (11.5)	136 (7.39)	114 (7.45)	132 (15.38)	68 (13.15)
19-29	5804 (26.66)	5575 (37.37)	3358 (30.77)	1469 (40.14)	1450 (55.47)	841 (47.81)	639 (24.45)	784 (36.35)	539 (29.29)	421 (27.5)	301 (35.08)	152 (29.4)
30-39	5609 (25.76)	3265 (21.88)	2605 (23.87)	761 (20.79)	363 (13.89)	302 (17.17)	705 (26.97)	540 (25.03)	486 (26.41)	421 (27.5)	209 (24.36)	139 (26.88)
≥40	8527 (39.16)	3707 (24.85)	3588 (32.88)	940 (25.68)	283 (10.83)	298 (16.94)	1137 (43.5)	585 (27.12)	659 (35.82)	575 (37.56)	216 (25.17)	158 (30.56)
<b>Top 5 cities</b>												
	Toronto	Toronto	Toronto	Los Angeles	Houston	Los Angeles	Cardiff	Leeds	Glasgow	Sydney	Sydney	Sydney
	Ottawa	Calgary	Calgary	Houston	Los Angeles	Houston	East Midlands	Sheffield	Manchester	Melbourne	Melbourne	Melbourne
	Calgary	Edmonton	Ottawa	Manhattan	San Antonio	Brooklyn	Bristol	London	Sheffield	Brisbane	Brisbane	Brisbane
	Vancouver	Ottawa	Edmonton	Chicago	Chicago	Chicago	Lambeth	Liverpool	Leeds	Perth	Adelaide	Perth

Variable	Canada (N=98,722)			United States (N=10,193)			United Kingdom (N=9154)			Australia (N=4067)		
	PA <sup>b</sup>	SQ <sup>c</sup>	SB <sup>d</sup>	PA	SQ	PA	PA	SQ	SB	PA	SQ	SB
	Edmon- ton	Montréal	Vancou- ver	Florida	Brook- lyn	Flori- da	Liver- pool	Scot- land	Edinburgh	Adelaide	Perth	Adelaide
<b>Source, n (%)</b>												
Organiza- tion	3109 (14.28)	727 (4.87)	1105 (10.13)	63 (1.72)	97 (3.71)	76 (4.32)	368 (14.08)	94 (4.36)	143 (7.77)	135 (8.82)	22 (2.56)	26 (5.0)
Individual	18,663 (85.72)	14,192 (95.13)	9807 (89.87)	3391 (92.65)	2551 (97.59)	1638 (93.12)	2246 (85.92)	2063 (95.64)	1697 (92.23)	1396 (91.18)	836 (97.44)	491 (95.0)

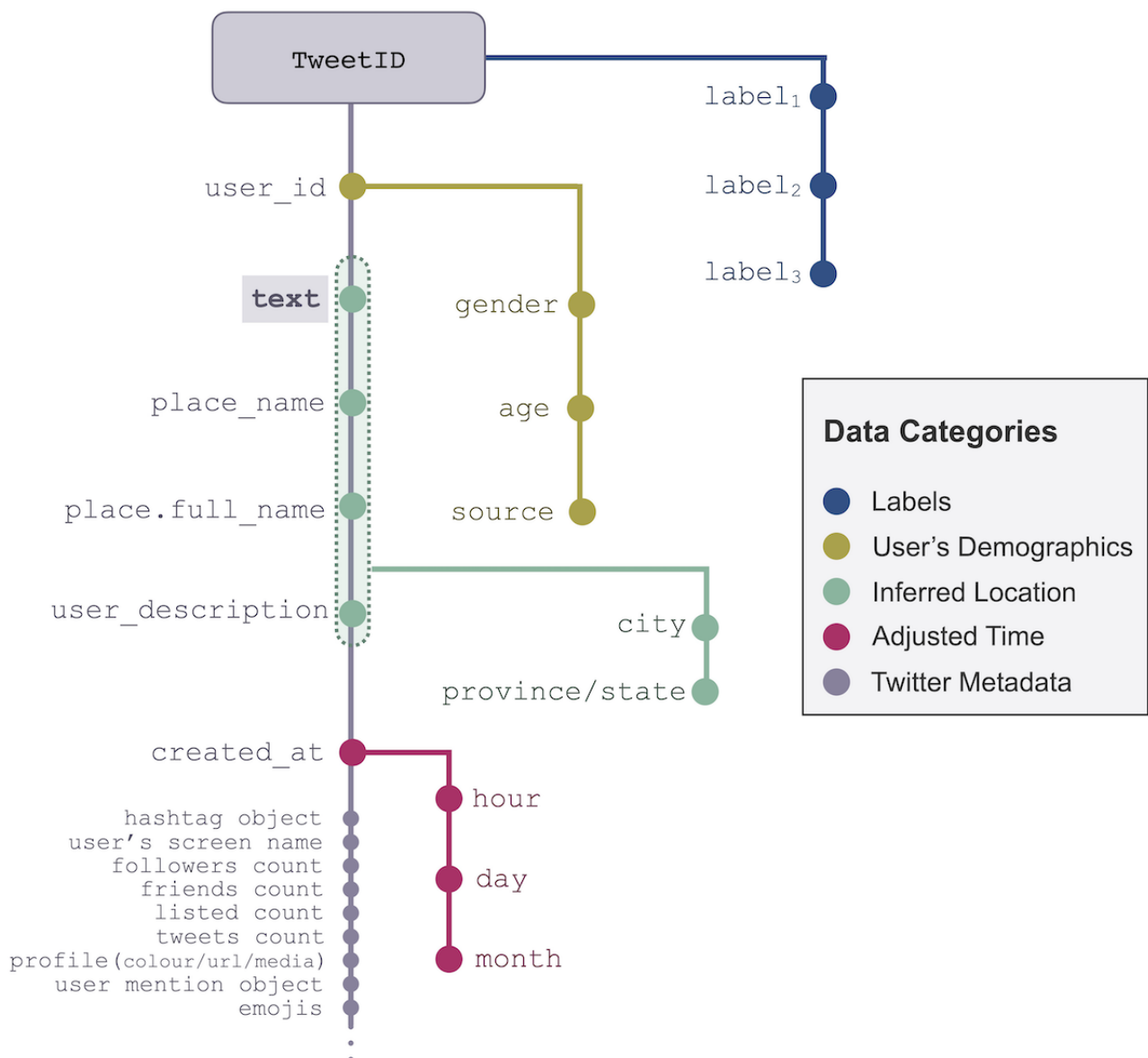
<sup>a</sup>The 3 labels collected for each tweet were consolidated into a single label using majority voting. The discrepancy between the numbers of binary and multi-class labels is because of how majority voting calculates the truth label for each of these categories.

<sup>b</sup>PA: physical activity.

<sup>c</sup>SQ: sleep quality.

<sup>d</sup>SB: sedentary behavior.

**Figure 3.** Overview of information tracking using TweetID. Each tweet or text is identified by a unique TweetID (provided by LPHEADADA). This ID is mapped to metadata that includes user\_id, place\_name, place\_full\_name, user\_description, and created\_at. A total of 3 labels are provided for each TweetID that can be used for developing machine learning models. user\_id was used to infer the demographics of each tweet, including gender, age range, and source. Adjusted time (month/day/hour) was extracted using created\_at, and text, place\_name, place\_full\_name, and user\_description were used to identify the city and state or province mapped to each TweetID. LPHEADADA: Labelled Digital Public Health Data Set.



**Labels**

Let  $L$  denote the set of  $j$  unique class labels,  $t$  represent the tweet text, and  $w_k$  represent the  $k^{th}$  worker who labeled the tweet, where  $k \in \{1,2,3\}$ . Each  $l_j \in L$  is defined based on two conditions: whether the tweet is self-reported ( $c_1 \in \{0,1\}$ ) and whether the tweet reports a recent PASS experience ( $c_2 \in \{0,1\}$ ). For each PASS category across the 4 countries, the data set contains the following 2 subsets of labels for each tweet.

**Multi-class Labels**

In this subset, each tweet  $t$  is mapped to quadruple  $L = \langle tweetID, (w_1, l_{j1}), (w_2, l_{j2}), (w_3, l_{j3}) \rangle$ , where  $j=5$  and each  $l_j$  is defined based on the values of both  $c_1$  and  $c_2$  conditions and can be formulated as  $\{11,10,01,00\}$ . We also let workers choose a fifth option, called *unclear*, to ensure they do not give random labels to tasks that they are not confident of performing successfully. The 11, 10, 01, 00, and *unclear* labels correspond to the YY, YN, NY, NN, and NC labels, respectively, presented in Table 1.

**Binary Labels**

Each label  $l_j$  in this subset is defined based on logical AND relationship between conditions  $c_1$  and  $c_2$ . Thus,  $l_1=1$  if the tweet

presents self-reported PASS surveillance and  $l_1=0$  otherwise. Like the multi-class category, each tweet is mapped to a quadruple, and there is a class called *unclear* ( $l_3$ ) with  $j=3$ . The binary labels did not directly come from the AMT workers and were generated by dichotomizing the collected labels.

**User’s Demographic Data**

In the demographic data set, each tweet  $t$  is mapped to quadruple  $D = \langle tweetID, a, g, o \rangle$ , where  $g \in \{male, female\}$  represents the gender of the user who posted the tweet,  $a \in \{\leq 18, 19-29, 30-39, 40\leq\}$  represents their age range, and  $o \in \{0,1\}$  shows the source of the tweet (ie,  $o=1$  if the tweet was posted by an organization and  $o=0$  otherwise). Table 2 shows the demographic distribution of the Canadian data set based on gender and age range of the unique users associated with tweet IDs. We can see that 79.24% (16,027/20,227) of female users in this data set are inferred to be aged <40 years, whereas this number for the male users is 57.55% (15,754/27,376). In addition, the most populated age category for female users across all PASS categories is 19-29 years, whereas this range for male users is  $\geq 40$  years. Excluding the (female, sleep quality) category, the age range  $\leq 18$  years, regardless of the user sex, is the least populated category across all PASS categories.

**Table 2.** Demographic information of users associated with the tweets originated in Canada (N=42,603).

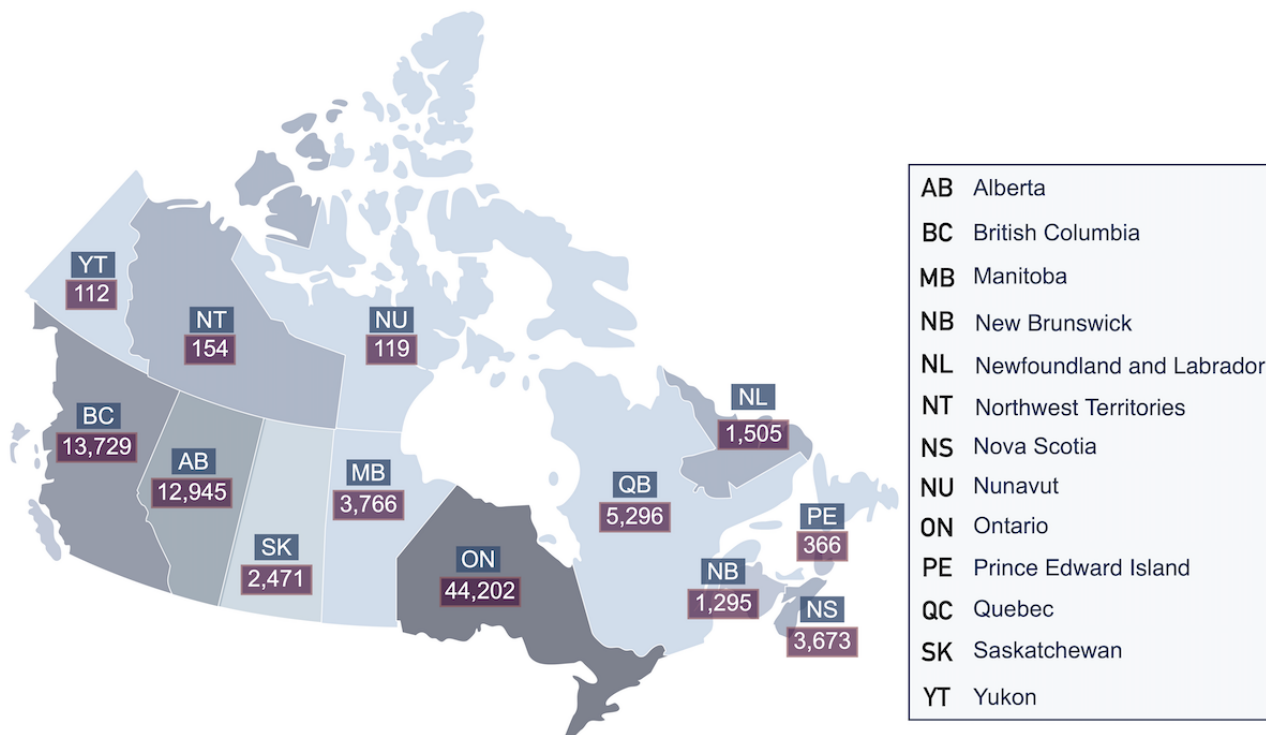
	Age group, n (%)				Total
	Age≤18 years	Age 19-29 years	Age 30-39 years	Age≥40 years	
<b>Physical activity (n=21,772)</b>					
Female	779 (3.58)	3160 (14.51)	2408 (11.06)	2124 (9.76)	8471 (38.91)
Male	993 (4.56)	2704 (12.42)	3201 (14.7)	6403 (29.41)	13,301 (61.09)
<b>Sedentary behavior (n=10,912)</b>					
Female	607 (5.56)	1822 (16.7)	1108 (10.15)	949 (8.7)	4486 (41.11)
Male	754 (6.91)	1536 (14.08)	1497 (13.72)	2639 (24.18)	6426 (58.89)
<b>Sleep quality (n=14,919)</b>					
Female	1267 (8.49)	3307 (22.17)	1569 (10.52)	1127 (7.55)	7270 (48.73)
Male	1105 (7.41)	2268 (15.2)	1696 (11.37)	2580 (17.29)	7649 (51.27)


**Inferred Location Data**

Each row of the location data set is presented in the form of  $A = \langle tweetID, c, p \rangle$ , where  $c$  and  $p$  denote the city and province or state associated with each tweet, respectively. Using the TweetID parameter, the location data can be mapped to other data sets, including labels, user demographics, time, and Twitter metadata. Each of the  $c$  and  $p$  variables is inferred based on the raw variables in Twitter metadata, including text, place objects, and user’s profile description (Figures 2 and 3). For example,

Figure 4 shows the distribution of labeled tweets for each PASS category across Canadian provinces (ie,  $p$ ). For the top 5 provinces, the overall size of the data set is directly proportional to the population size of each province. However, as only English tweets from Twitter users are included in the data set, LPHEADADA represents only English-speaking Quebecor’s and Francophone Quebecor’s tweets in English, placing the province in fourth place. Moreover, with a lower population than British Columbia, Alberta had more sedentary behavior and sleep quality tweets and places in second place (Figure 4).

**Figure 4.** Geospatial details of the Canadian data set.



	AB	BC	MB	NB	NL	NT	NS	NU	ON	PE	QC	SK	YT
Physical Activity 	6,692	7,512	1,726	598	797	80	1,780	56	20,092	192	2,177	1,398	67
Sleep 	4,048	4,032	1,321	439	430	56	1,274	42	15,788	100	2,226	661	25
Sedentary Behaviour 	2,205	2,185	719	258	278	18	619	21	8,322	74	893	412	20

**Temporal Data**

The temporal data set inferred from the *created\_at* field of the Twitter metadata presents the adjusted time of each tweet based on the tweet’s location. Each row of this data set is presented in the form of  $T = \langle tweetID, h, d, m \rangle$ , where *h*, *d*, and *m* represent the hour, weekday, and month associated with each tweet, respectively. The *year* value does not need any adjustment and can be extracted directly from the original *created\_at* field. For example, Figure 1 represents the frequency of tweets in each PASS category across Canada at the national (Figure 1A) and city levels (Figures 1B-1D). The highlighted area in Figure 1A demonstrates the data set’s temporal windows that can be used to compare different aspects of PASS surveillance between 2019 and 2020.

**Twitter Metadata**

In addition to the inferred data records mentioned above, TweetIDs presented in LPHEADATA can be used to retrieve Twitter metadata. This metadata, in addition to the tweet text, place object, time of the tweet, and user IDs, provides more details on the tweet and user objects, including the following.

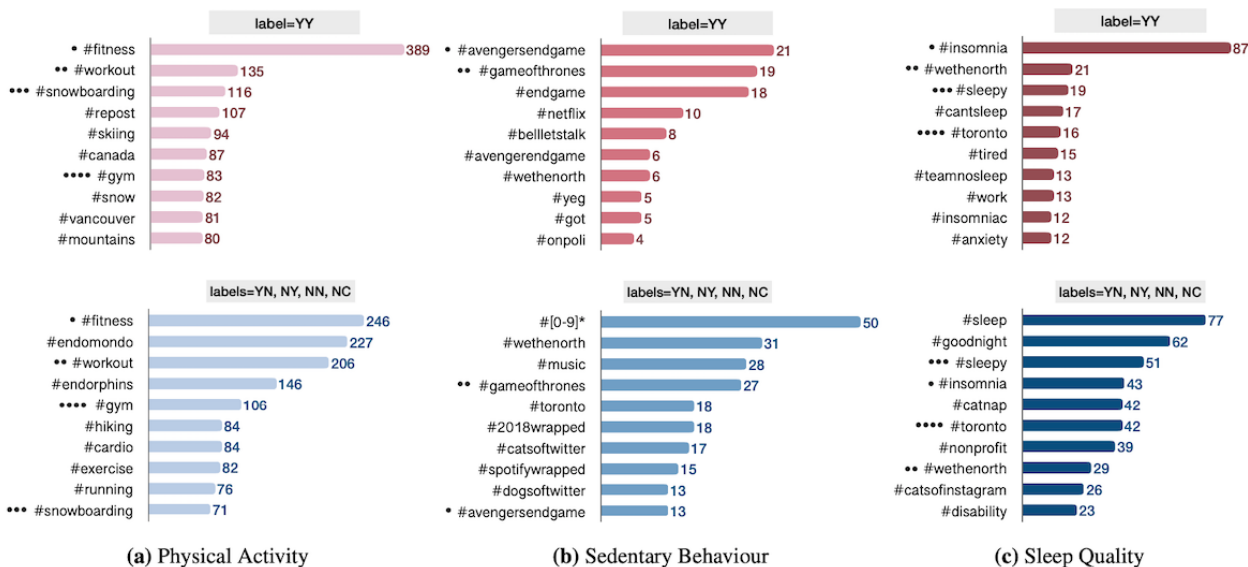
**User Object**

This object comprises user’s screen name, description, follower count, friend count, listed count (ie, the number of public lists that the user is a member of), tweet count (ie, the number of tweets issued by the user), and profile characteristics (eg, image, color, and URLs).

**Tweet Object**

This object comprises hashtags mentioned in each tweet, emojis, user mentions, URLs, and media (eg, images and videos). For example, Figure 5 illustrates the distribution of the top 10 hashtags per label for each of the PASS categories in the entire data set. Hashtags are basically keywords or word strings prefixed with the symbol # that are used for categorizing and communicating tweets pertaining to the same topics. Although the high level of intersection between the hashtags of positive and negative classes in our data set makes this feature a less discriminating feature for the development of ML models (eg, annotated hashtags in Figure 5), this field can still be used by PASS-related advocacy campaigns on Twitter to brand their movement and open up their campaigns to users who need more information about the context [29]. As tagged tweets are easily archived and accessible, the hashtag field can be effectively leveraged to improve the public’s engagement in digital PHS discussions.

**Figure 5.** The distribution of the top 10 hashtags per label for each of the physical activity, sedentary behavior, and sleep quality categories. The label on top of each bar graph shows the class of tweets presented in the graph. For example, the YY category presents all tweets that are self-reported and describe a recent PASS experience. Similarly, YN presents all self-reported tweets but does not present a recent PASS experience. NC presents tweets with an unclear label. The number at the end of each bar presents the frequency of its corresponding hashtag. The intersections between 2 classes of labels for each PASS category are annotated using filled circles (•). Same hashtags are tagged with the same number of circles. This figure is based on all data collected from Canada, the United States, the United Kingdom, and Australia. PASS: physical activity, sedentary behavior, and sleep.



### General Release Notes

We have made our data set publicly available, along with instructions and Jupyter Notebooks [30,31] to illustrate the application of the data. All data and code (written in Python 3) used in this study are available through our GitHub repository [25]. We provide all necessary instructions, required libraries, and sample Jupyter Notebooks, allowing replicating our experiments and using the data set.

## Discussion

### Technical Validation

To verify the quality of crowd-generated labels and set a baseline for the data set, we conducted 4 studies. First, we used a series of statistical inference models to verify the quality of the labels provided in this data set. Second, we evaluated the semantic consistency between the data sets collected from the countries included in our repository. Third, we trained and tested 12 binary classifiers using the labels provided in the data set. Finally, to investigate the structural differences between all subsets of LPHEADADA, we conducted linguistic and lexical analysis and visualized the results for further comparisons. Moreover, to address unseen technical issues of the data set, we provide a public issue tracker for handling bug reports, describing solutions to technical issues, data updates, and other issues and contributions.

### Methods of Label Agreement

To measure the consistency of labels generated by AMT workers, we calculated label consistency (*LC*) as the average entropy of the collected labels for each PASS category [32]. For each tweet  $t_i \in T_s$ , where  $T_s$  denotes the set of all tweets related to surveillance category  $s \in \{\text{physical activity, sleep}$

quality, sedentary behavior} and  $s \in \{\text{physical activity, sleep quality, sedentary behaviour}\}$ ,  $n_j$  defines the number of answers given to the  $j^{th}$  choice ( $j \in \{1,2,3,4,5\}$ ), as we have 5 choices for each tweet), we calculate *LC* as follows:

$$LC = \frac{1}{|s|} \sum_{j=1}^5 n_j \log_2 \frac{|s|}{n_j}$$

$|s|$  denotes the size of the surveillance category  $s$ , and as we collect 3 labels for each tweet, the denominators in the entropy formula receive the constant value of 3. *LC* ranges from 0 to 1, and the values close to 0 show less consistency between workers' input. After calculating *LC* for each PASS category, we had  $LC > 0.52$  for the multi-class labeling and  $LC > 0.73$  for the binary labeling task.

To consolidate the collected labels for each tweet, we used the majority voting (MV) technique (Table 1). Defining the estimated label as  $\hat{l}_w$ , and the submitted label by worker  $w$  as  $l_w$ , the MV approach for a binary labeling task assigns 1 to  $\hat{l}_w$  if  $l_w = 1$  and assigns 0 if otherwise. The discrepancy between the number of binary and multi-class labels presented in Table 1 is caused by the way that MV approach calculates the truth label for each of these categories. In addition to MV, there are models, such as those by David and Skene [33] and Raykar et al [34] and the generative model of labels, abilities, and difficulties [35], that incorporate the error rate of annotators (workers), task complexity, and context-sensitive features into the inference process and can be used to predict truth labels from crowd-labeled data.

### Semantic Consistency

To validate the semantic consistency of the data sets collected from different countries, we transformed the data set of each

PASS category into a semantic space of low dimensionality using latent semantic analysis. For the vector presentation of each data set, to capture high-level semantics of the text, we ran the PASS category of each country through a pretrained word2vec embedding model. This model contains 300-dimensional vectors of 3 million words and phrases trained on 100 billion words from a Google News data set. The resulting 300-dimensional vectors were then averaged for each tweet.

For each *tweet*  $T$  composed of words  $\langle w_1, w_2, \dots, w_n \rangle$ , with  $\boxed{x}$  defining the embedding of  $w_i$ , the embedding of *tweet*  $T$  can be calculated as:

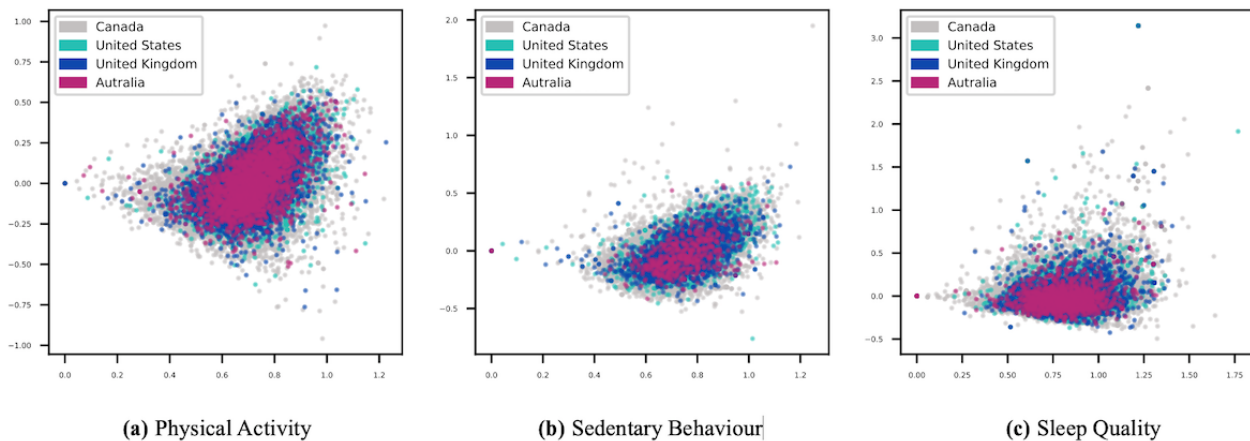
$$\boxed{x}$$

We then applied truncated singular value decomposition on the new vectorized data set and kept the top 2 dimensions of the data set containing the most variance (eg, those directions in vector space of the data set that contain more information). The

scatterplots presented in Figure 6 illustrate our data sets in a 2-dimensional latent semantic space. The high level of overlap between the data sets of each PASS category indicates that the data from different countries cover similar semantic space, but the space is scaled differently based on the size of the data sets.

Moreover, to further investigate the internal consistency of the data sets presented in this paper, we trained 3 convolutional neural network multi-class classifiers (ie, 1 for each PASS category) to classify tweets into Canada, US, Australia, or UK classes. Given the highly imbalanced distribution of the classes in our data set due to the highly unequal number of samples from each country, we used the average precision (AP) metric to measure the discrimination ability of our predictive models. The poor performance of these classifiers ( $AP_{PA}$ , 37%;  $AP_{SB}$ , 32%;  $AP_{SQ}$ , 31%) in detecting each tweet's country implies a high level of semantic and syntactic cohesion among the 4 countries in our data set.

**Figure 6.** Scatter plots of the first 2 dimensions of latent semantic analysis performed on physical activity, sedentary behavior, and sleep categories, and classified based on the geographic source of the data.



### Classification of PASS Categories

For the PASS classification experiment, we used a standard convolutional neural network classifier with 1 layer of convolution with global max-pooling on top of a word2vec embedding trained on 100 billion words of Google News. The vectors have a dimensionality of 300 and were trained using the continuous bag-of-words architecture [36]. We used the binary labels of the data set to train and evaluate the model on each of 12 data sets provided in LPHEADADA. Owing to the imbalanced distribution of binary labels across these data sets (Table 1), in addition to precision, recall, F1, and area under the curve (AUC) scores, we used AP to measure the weighted mean of precision at different thresholds to make the score

robust to heterogeneous and imbalanced class distributions. Like AUC score, AP is a model-wide and threshold-free evaluation metric. However, for imbalance class distributions with the negatives outnumbering the positives, AP is more informative than AUC, as it mainly evaluates the fraction of true positive samples among positive predictions and is more robust to the relationship between false-positive and false-negative rates [37]. As shown in Table 3, for each of the Canada, US, and UK data sets, we find a steady increase in the overall performance of the classifier as the size of the data set increase (ie,  $|PA| > |SQ| > |SB|$ ). Interestingly, the UK data set achieves the highest performance for the PA category among all the countries.

**Table 3.** Binary classification of tweets using bidirectional long short-term memory.<sup>a</sup>

Metrics	Country (%)			
	Canada	United States	United Kingdom	Australia
<b>Physical activity</b>				
Precision	78	75	80	65
Recall	78	76	81	68
F1	78	76	80	64
AUC <sub>ROC</sub> <sup>b</sup>	83	76	84	64
AP <sup>c</sup>	81	75	81	66
<b>Sedentary behavior</b>				
Precision	73	62	79	74
Recall	73	64	83	86
F1	73	62	79	80
AUC <sub>ROC</sub>	76	62	63	57
AP	78	65	66	66
<b>Sleep quality</b>				
Precision	76	70	73	61
Recall	76	70	76	60
F1	75	70	73	60
AUC <sub>ROC</sub>	81	70	66	63
AP	83	72	71	64

<sup>a</sup>The same classifier is used to classify the data from different countries.

<sup>b</sup>AUC<sub>ROC</sub>: area under the receiver operating characteristic curve.

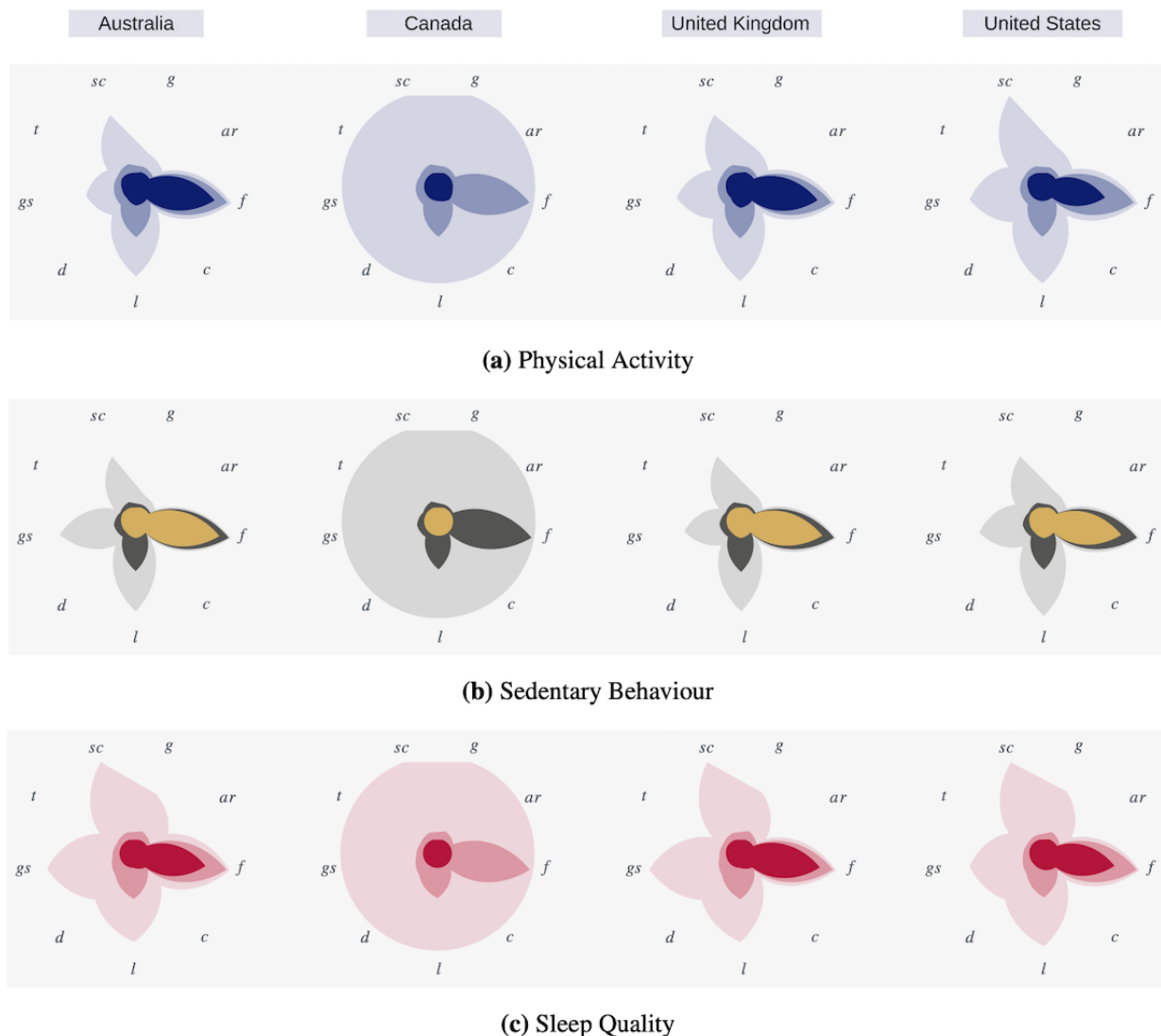
<sup>c</sup>AP: average precision.

## Linguistic Properties

To understand and validate the linguistic properties of each data set, we measured and visually compared the following metrics for each PASS category grouped by country: (1) sentence count, (2) grammar score (ie, number of grammar errors), (3) the average number of syllables per word and the average sentence length (ie, Flesch-Kincaid Grade Level index [38]), (4) the average number of words per sentence and the percentage of words with 3 or more syllables (ie, Gunning Fog index [39]), (5) a combination of average sentence length and percentage of difficult words (ie, Dale-Chall readability [40]), (6) sentence length and number of polysyllables (ie, Linsear Write readability [40]), (7) number of characters (ie, Coleman-Liau Index [40,41]), (8) the average number of characters per word and number of words per sentence (ie, automated readability index [38]), and (9) the text standard score based on number of

sentences, words, syllables, and characters in each tweet (ie, text readability consensus). Figure 7 illustrates these comparisons based on the minimum (red), average (pink), and maximum (light pink) values of each feature. Although all data sets have similar behavior in terms of each feature's minimum value, the Canadian data set has a lower score for the average number of syllables per word and the average sentence length for all PASS categories. Interestingly, the sleep quality data set, compared with other PASS categories, has a higher value for the maximum number of *grammar errors* and *sentence count* metrics, whereas all data sets show the same behavior for the minimum and the average values of these metrics. These location-specific linguistic characteristics should be considered when using these data sets to train and evaluate PASS surveillance ML models. For example, a model trained on the Canadian data set may not present some linguistic features of a data set that originated in Australia and vice versa.

**Figure 7.** Linguistic comparisons across different countries. *ar*: automated readability index; *c*: Coleman-Liau Index; *d*: Dale-Chall readability; *f*: Flesch-Kincaid grade level; *g*: Gunning Fog index; *gs*: grammar score; *l*: Linsear Write readability; *sc*: sentence count; *t*: text readability consensus.



## Limitations

Several limitations should be noted. First, we collected our data set using Twitter's free streaming API, which returns a random sample of about only 1% of global public tweets produced at a given time. However, our data set spans 19 months of tweets posted by users from 4 English-speaking countries, which provides enough diversity and coverage for conducting retrospective and comparative digital public PASS studies. Moreover, the search terms used to filter the data set could have impacted the topics included in our data set, which may influence the generalizability of the results derived from this data set. To address this and to ensure the lists of

context-sensitive terms for filtering all the PASS categories are comprehensive enough, we used domain-specific ontologies, WordNet [42], and NLP techniques (eg, topic modeling, language modeling, and lexical analysis) to detect latent word patterns to identify PASS-related contexts in unstructured text.

Despite these limitations, the curated, validated [43], and labeled data set provided in this paper will allow researchers and practitioners to delve into different aspects of digital PASS surveillance by developing ML, NLP, and exploratory models. We believe that the novelty and comprehensiveness of this data set will help the development, evaluation, and deployment of digital PASS surveillance systems, and it will be an invaluable resource for both public health researchers and practitioners.

## Acknowledgments

This work was supported by a postdoctoral scholarship from the Libin Cardiovascular Institute and the Cumming School of Medicine, University of Calgary. Additionally, this work was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2014-04743). The Public Health Agency of Canada funded the Amazon Mechanical Turk costs.



## Authors' Contributions

ZSHA was responsible for data collection and curation, model development, documentation and maintenance of the data set, data analysis and visualization and wrote the paper. GB and WT reviewed the paper and provided comments. JL contributed to the conception and design of the study and revised the manuscript.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

A complete list of the regular expressions and filtering words for each PASS domain.

[[PDF File \(Adobe PDF File\), 63 KB - publichealth\\_v8i2e32355\\_app1.pdf](#)]

## References

1. Kemp S. Digital 2020: July global statshot. Datareportal. 2020. URL: <https://datareportal.com/reports/digital-2020-july-global-statshot> [accessed 2021-01-08]
2. Fung IC, Tse ZT, Fu KW. The use of social media in public health surveillance. *Western Pac Surveill Response J* 2015;6(2):3-6. [doi: [10.5365/wpsar.2015.6.1.019](https://doi.org/10.5365/wpsar.2015.6.1.019)] [Medline: [26306208](https://pubmed.ncbi.nlm.nih.gov/26306208/)]
3. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection--harnessing the Web for public health surveillance. *N Engl J Med* 2009;360(21):2153-2157. [doi: [10.1056/nejmp0900702](https://doi.org/10.1056/nejmp0900702)] [Medline: [19423867](https://pubmed.ncbi.nlm.nih.gov/19423867/)]
4. Sinnenberg L, Buttenheim AM, Padrez K, Mancheno C, Ungar L, Merchant RM. Twitter as a tool for health research: a systematic review. *Am J Public Health* 2017;107(1):e1-e8. [doi: [10.2105/ajph.2016.303512](https://doi.org/10.2105/ajph.2016.303512)] [Medline: [27854532](https://pubmed.ncbi.nlm.nih.gov/27854532/)]
5. Pershad Y, Hange P, Albadawi H, Oklu R. Social medicine: Twitter in healthcare. *J Clin Med* 2018;7(6):121. [doi: [10.3390/jcm7060121](https://doi.org/10.3390/jcm7060121)] [Medline: [29843360](https://pubmed.ncbi.nlm.nih.gov/29843360/)]
6. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Health Inf Sci Syst* 2014;2:3. [doi: [10.1186/2047-2501-2-3](https://doi.org/10.1186/2047-2501-2-3)] [Medline: [25825667](https://pubmed.ncbi.nlm.nih.gov/25825667/)]
7. Shakeri Hossein Abad Z, Kline A, Sultana M, Noaen M, Nurmambetova E, Lucini F, et al. Digital public health surveillance: a systematic scoping review. *NPJ Digit Med* 2021;4(1):41. [doi: [10.1038/s41746-021-00407-6](https://doi.org/10.1038/s41746-021-00407-6)] [Medline: [33658681](https://pubmed.ncbi.nlm.nih.gov/33658681/)]
8. Teutsch SM. Considerations in planning a surveillance system. In: Lee LM, Teutsch SM, Thacker SB, St. Louis ME, editors. *Principles and practice of public health surveillance*. Oxford: Oxford University Press; 2010:18-28.
9. Declich S, Carter AO. Public health surveillance: historical origins, methods and evaluation. *Bull World Health Organ* 1994;72(2):285-304 [FREE Full text] [Medline: [8205649](https://pubmed.ncbi.nlm.nih.gov/8205649/)]
10. Croner CM. Public health, GIS, and the internet. *Annu Rev Public Health* 2003;24:57-82. [doi: [10.1146/annurev.publhealth.24.012902.140835](https://doi.org/10.1146/annurev.publhealth.24.012902.140835)] [Medline: [12543872](https://pubmed.ncbi.nlm.nih.gov/12543872/)]
11. Graham M, Hale SA, Gaffney D. Where in the world are you? Geolocation and language identification in Twitter. *Prof Geogr* 2014;66(4):568-578. [doi: [10.1080/00330124.2014.907699](https://doi.org/10.1080/00330124.2014.907699)]
12. Doan S, Ritchart A, Perry N, Chaparro JD, Conway M. How do you #relax when you're #stressed? A content analysis and infodemiology study of stress-related tweets. *JMIR Public Health Surveill* 2017;3(2):e35. [doi: [10.2196/publichealth.5939](https://doi.org/10.2196/publichealth.5939)] [Medline: [28611016](https://pubmed.ncbi.nlm.nih.gov/28611016/)]
13. Woo H, Sung Cho H, Shim E, Lee JK, Lee K, Song G, et al. Identification of keywords from Twitter and Web blog posts to detect influenza epidemics in Korea. *Disaster Med Public Health Prep* 2018;12(3):352-359. [doi: [10.1017/dmp.2017.84](https://doi.org/10.1017/dmp.2017.84)] [Medline: [28756796](https://pubmed.ncbi.nlm.nih.gov/28756796/)]
14. Chorionopoulos K, Talvis K. Flutrack.org: open-source and linked data for epidemiology. *Health Informatics J* 2016;22(4):962-974. [doi: [10.1177/1460458215599822](https://doi.org/10.1177/1460458215599822)] [Medline: [26351261](https://pubmed.ncbi.nlm.nih.gov/26351261/)]
15. Jimeno-Yepes A, MacKinlay A, Han B, Chen Q. Identifying diseases, drugs, and symptoms in Twitter. *Stud Health Technol Inform* 2015;216:643-647. [Medline: [26262130](https://pubmed.ncbi.nlm.nih.gov/26262130/)]
16. Aphinyanaphongs Y, Lulejian A, Brown DP, Bonneau R, Krebs P. Text classification for automatic detection of e-cigarette use and use for smoking cessation from Twitter: a feasibility pilot. *Pac Symp Biocomput* 2016;21:480-491 [FREE Full text] [Medline: [26776211](https://pubmed.ncbi.nlm.nih.gov/26776211/)]
17. Mowery D, Smith H, Cheney T, Stoddard G, Coppersmith G, Bryan C, et al. Understanding depressive symptoms and psychosocial stressors on Twitter: a corpus-based study. *J Med Internet Res* 2017;19(2):e48. [doi: [10.2196/jmir.6895](https://doi.org/10.2196/jmir.6895)] [Medline: [28246066](https://pubmed.ncbi.nlm.nih.gov/28246066/)]
18. Müller MM, Salathé M. Crowdbreaks: tracking health trends using public social media data and crowdsourcing. *Front Public Health* 2019;7:81. [doi: [10.3389/fpubh.2019.00081](https://doi.org/10.3389/fpubh.2019.00081)] [Medline: [31037238](https://pubmed.ncbi.nlm.nih.gov/31037238/)]
19. O'Connor K, Pimpalkhute P, Nikfarjam A, Ginn R, Smith KL, Gonzalez G. Pharmacovigilance on Twitter? Mining tweets for adverse drug reactions. *AMIA Annu Symp Proc* 2014;2014:924-933 [FREE Full text] [Medline: [25954400](https://pubmed.ncbi.nlm.nih.gov/25954400/)]
20. Watkins NG, Lovell NH, Larsen ME. SMCT - an innovative tool for mental health analysis of Twitter data. *Annu Int Conf IEEE Eng Med Biol Soc* 2018;2018:4114-4117. [doi: [10.1109/EMBC.2018.8513257](https://doi.org/10.1109/EMBC.2018.8513257)] [Medline: [30441260](https://pubmed.ncbi.nlm.nih.gov/30441260/)]

21. Zhao L, Chen J, Chen F, Wang W, Lu CT, Ramakrishnan N. Simnest: social media nested epidemic simulation via online semi-supervised deep learning. In: 2015 IEEE International Conference on Data Mining. 2015 Presented at: ICDM'15; November 14-17, 2015; Atlantic City p. 639-648. [doi: [10.1109/icdm.2015.39](https://doi.org/10.1109/icdm.2015.39)]
22. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, Lancet Physical Activity Series Working Group. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet* 2012;380(9838):219-229. [doi: [10.1016/s0140-6736\(12\)61031-9](https://doi.org/10.1016/s0140-6736(12)61031-9)] [Medline: [22818936](https://pubmed.ncbi.nlm.nih.gov/22818936/)]
23. Roberts KC, Butler G, Branchard B, Rao DP, Otterman V, Thompson W, et al. The physical activity, sedentary behaviour and sleep (PASS) indicator framework. *Health Promot Chronic Dis Prev Can* 2017;37(8):252-256. [doi: [10.24095/hpcdp.37.8.04](https://doi.org/10.24095/hpcdp.37.8.04)] [Medline: [28800295](https://pubmed.ncbi.nlm.nih.gov/28800295/)]
24. Paolacci G, Chandler J, Ipeirotis PG. Running experiments on Amazon mechanical turk. *Judgm Decis Mak* 2010;5(5):411-419.
25. Filtering tweets by location. Twitter developer platform. URL: <https://developer.twitter.com/en/docs/tutorials/filtering-tweets-by-location> [accessed 2021-01-12]
26. Wang Z, Hale S, Adelani DI, Grabowicz P, Hartman T, Flöck F, et al. Demographic inference and representative population estimates from multilingual social media data. In: Proceedings of the 2019 World Wide Web Conference. New York: Association for Computing Machinery; 2019 Presented at: WWW'19; May 13-17, 2019; San Francisco p. 2056-2067. [doi: [10.1145/3308558.3313684](https://doi.org/10.1145/3308558.3313684)]
27. Developer agreement and policy. Twitter developer platform. 2020. URL: <https://developer.twitter.com/en/developer-terms/agreement-and-policy> [accessed 2020-08-01]
28. Data intelligence for health lab: LPHEADA-labelled-public-HEalth-DAtaset. Github. URL: <https://github.com/data-intelligence-for-health-lab/Lpheada-Labelled-Public-HEalth-DAtaset> [accessed 2022-02-03]
29. Ojo A, Guntuku SC, Zheng M, Beidas RS, Ranney ML. How health care workers wield influence through Twitter hashtags: retrospective cross-sectional study of the gun violence and COVID-19 public health crises. *JMIR Public Health Surveill* 2021;7(1):e24562. [doi: [10.2196/24562](https://doi.org/10.2196/24562)] [Medline: [33315578](https://pubmed.ncbi.nlm.nih.gov/33315578/)]
30. Perez F, Granger BE. IPython: a system for interactive scientific computing. *Comput Sci Eng* 2007;9(3):21-29. [doi: [10.1109/mcse.2007.53](https://doi.org/10.1109/mcse.2007.53)]
31. Kluyver T, Ragan-Kelley B, Pérez F, Granger B, Bussonnier M, Frederic J, Jupyter Development Team. Jupyter notebooks - a publishing format for reproducible computational workflows. In: Proceedings of the 20th International Conference on Electronic Publishing. Amsterdam: IOS Press; 2016 Presented at: ELPUB '16; June 7-9, 2016; Göttingen p. 87-90. [doi: [10.3233/978-1-61499-649-1-87](https://doi.org/10.3233/978-1-61499-649-1-87)]
32. Zheng Y, Li G, Li Y, Shan C, Cheng R. Truth inference in crowdsourcing: is the problem solved? *Proc VLDB Endow* 2017;10(5):541-552. [doi: [10.14778/3055540.3055547](https://doi.org/10.14778/3055540.3055547)]
33. Dawid AP, Skene AM. Maximum likelihood estimation of observer error-rates using the EM algorithm. *J R Stat Soc Ser C Appl Stat* 1979;28(1):20-28. [doi: [10.2307/2346806](https://doi.org/10.2307/2346806)]
34. Raykar VC, Yu S, Zhao LH, Valadez GH, Florin C, Bogoni L, et al. Learning from crowds. *J Mach Learn Res* 2010;11(43):1297-1322.
35. Whitehill J, Wu TF, Bergsma J, Movellan J, Ruvolo P. Whose vote should count more: optimal integration of labels from labelers of unknown expertise. In: Proceedings of the 22nd International Conference on Neural Information Processing Systems. Red Hook: Curran Associates; 2009 Presented at: NIPS '09; December 7-10, 2009; Vancouver p. 2035-2043.
36. Mikolov T, Sutskever I, Chen K, Corrado G, Dean J. Distributed representations of words and phrases and their compositionality. In: Proceedings of the 26th International Conference on Neural Information Processing Systems - volume 2. Red Hook: Curran Associates; 2013 Presented at: NIPS '13; December 5-10, 2013; Lake Tahoe p. 3111-3119.
37. Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One* 2015;10(3):e0118432 [FREE Full text] [doi: [10.1371/journal.pone.0118432](https://doi.org/10.1371/journal.pone.0118432)] [Medline: [25738806](https://pubmed.ncbi.nlm.nih.gov/25738806/)]
38. Kincaid JP, Fishburne Jr RP, Rogers RL, Chissom BS. Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. Naval Education and Training Command. 1975. URL: <https://www.readcube.com/articles/10.21236%2Fada006655> [accessed 2021-07-09]
39. Gunning R. Technique of clear writing. Toronto: McGraw-Hill; 1952.
40. Mohammadhassanzadeh H, Sketris I, Traynor R, Alexander S, Winquist B, Stewart SA. Using natural language processing to examine the uptake, content, and readability of media coverage of a pan-Canadian drug safety research project: cross-sectional observational study. *JMIR Form Res* 2020;4(1):e13296. [doi: [10.2196/13296](https://doi.org/10.2196/13296)] [Medline: [31934872](https://pubmed.ncbi.nlm.nih.gov/31934872/)]
41. Coleman M, Liau TL. A computer readability formula designed for machine scoring. *J Appl Psychol* 1975;60(2):283-284. [doi: [10.1037/h0076540](https://doi.org/10.1037/h0076540)]
42. Miller GA. WordNet: a lexical database for English. *Commun ACM* 1995;38(11):39-41. [doi: [10.1145/219717.219748](https://doi.org/10.1145/219717.219748)]
43. Shakeri Hossein Abad Z, Butler GP, Thompson W, Lee J. Crowdsourcing for machine learning in public health surveillance: lessons learned from Amazon mechanical turk. *J Med Internet Res* 2022;24(1):e28749 [FREE Full text] [doi: [10.2196/28749](https://doi.org/10.2196/28749)] [Medline: [35040794](https://pubmed.ncbi.nlm.nih.gov/35040794/)]

## Abbreviations

**AMT:** Amazon Mechanical Turk  
**AP:** average precision  
**API:** application programming interface  
**AUC:** area under the curve  
**DPHS:** digital public health surveillance  
**HIT:** human intelligence task  
**LC:** label consistency  
**LPHEADA:** Labelled Digital Public Health Dataset  
**ML:** machine learning  
**MV:** majority voting  
**NLP:** natural language processing  
**PASS:** physical activity, sedentary behavior, and sleep  
**PHS:** public health surveillance

*Edited by Y Khader; submitted 23.07.21; peer-reviewed by N Zhang, M Lotto, K Kishore; comments to author 16.08.21; revised version received 02.10.21; accepted 13.10.21; published 14.02.22.*

*Please cite as:*

*Shakeri Hossein Abad Z, Butler GP, Thompson W, Lee J*

*Physical Activity, Sedentary Behavior, and Sleep on Twitter: Multicountry and Fully Labeled Public Data Set for Digital Public Health Surveillance Research*

*JMIR Public Health Surveill 2022;8(2):e32355*

*URL: <https://publichealth.jmir.org/2022/2/e32355>*

*doi: [10.2196/32355](https://doi.org/10.2196/32355)*

*PMID: [35156938](https://pubmed.ncbi.nlm.nih.gov/35156938/)*

©Zahra Shakeri Hossein Abad, Gregory P Butler, Wendy Thompson, Joon Lee. Originally published in JMIR Public Health and Surveillance (<https://publichealth.jmir.org>), 14.02.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

# Diagnostic Accuracy of an At-Home, Rapid Self-test for Influenza: Prospective Comparative Accuracy Study

Rachel E Geyer<sup>1\*</sup>, MPH; Jack Henry Kotnik<sup>1,2\*</sup>, BA; Victoria Lyon<sup>1</sup>, MPH; Elisabeth Brandstetter<sup>3,4</sup>, MPH; Monica Zigman Suchsland<sup>1</sup>, MPH; Peter D Han<sup>4,5</sup>, MS; Chelsey Graham<sup>6</sup>, MEng; Misja Ilcisin<sup>7</sup>, BS; Ashley E Kim<sup>3</sup>, BS; Helen Y Chu<sup>3,4</sup>, MD; Deborah A Nickerson<sup>4,5</sup>, PhD; Lea M Starita<sup>4,5</sup>, PhD; Trevor Bedford<sup>4,5,7</sup>, PhD; Barry Lutz<sup>2,4</sup>, PhD; Matthew J Thompson<sup>1</sup>, MPH, MD, DPhil

<sup>1</sup>Department of Family Medicine, University of Washington, Seattle, WA, United States

<sup>2</sup>Department of Bioengineering, University of Washington, Seattle, WA, United States

<sup>3</sup>Department of Medicine, University of Washington, Seattle, WA, United States

<sup>4</sup>Brotman Baty Institute, University of Washington, Seattle, WA, United States

<sup>5</sup>Department of Genome Sciences, University of Washington, Seattle, WA, United States

<sup>6</sup>Massachusetts Institute of Technology, Cambridge, MA, United States

<sup>7</sup>Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA, United States

\*these authors contributed equally

**Corresponding Author:**

Rachel E Geyer, MPH

Department of Family Medicine

University of Washington

4225 Roosevelt Way NE Suite 308

Seattle, WA, 98195

United States

Phone: 1 2066163961

Email: [geyerr@uw.edu](mailto:geyerr@uw.edu)

## Abstract

**Background:** Rapid diagnostic tests (RDTs) for influenza used by individuals at home could potentially expand access to testing and reduce the impact of influenza on health systems. Improving access to testing could lead to earlier diagnosis following symptom onset, allowing more rapid interventions for those who test positive, including behavioral changes to minimize spread. However, the accuracy of RDTs for influenza has not been determined in self-testing populations.

**Objective:** This study aims to assess the accuracy of an influenza RDT conducted at home by lay users with acute respiratory illness compared with that of a self-collected sample by the same individual mailed to a laboratory for reference testing.

**Methods:** We conducted a comparative accuracy study of an at-home influenza RDT (Ellume) in a convenience sample of individuals experiencing acute respiratory illness symptoms. Participants were enrolled in February and March 2020 from the Greater Seattle region in Washington, United States. Participants were mailed the influenza RDT and reference sample collection materials, which they completed and returned for quantitative reverse-transcription polymerase chain reaction influenza testing in a central laboratory. We explored the impact of age, influenza type, duration, and severity of symptoms on RDT accuracy and on cycle threshold for influenza virus and ribonuclease P, a marker of human DNA.

**Results:** A total of 605 participants completed all study steps and were included in our analysis, of whom 87 (14.4%) tested positive for influenza by quantitative reverse-transcription polymerase chain reaction (70/87, 80% for influenza A and 17/87, 20% for influenza B). The overall sensitivity and specificity of the RDT compared with the reference test were 61% (95% CI 50%-71%) and 95% (95% CI 93%-97%), respectively. Among individuals with symptom onset  $\leq 72$  hours, sensitivity was 63% (95% CI 48%-76%) and specificity was 94% (95% CI 91%-97%), whereas, for those with duration  $>72$  hours, sensitivity and specificity were 58% (95% CI 41%-74%) and 96% (95% CI 93%-98%), respectively. Viral load on reference swabs was negatively correlated with symptom onset, and quantities of the endogenous marker gene ribonuclease P did not differ among reference standard positive and negative groups, age groups, or influenza subtypes. The RDT did not have higher sensitivity or specificity among those who reported more severe illnesses.

**Conclusions:** The sensitivity and specificity of the self-test were comparable with those of influenza RDTs used in clinical settings. False-negative self-test results were more common when the test was used after 72 hours of symptom onset but were not related to inadequate swab collection or severity of illness. Therefore, the deployment of home tests may provide a valuable tool to support the management of influenza and other respiratory infections.

(*JMIR Public Health Surveill* 2022;8(2):e28268) doi:[10.2196/28268](https://doi.org/10.2196/28268)

## KEYWORDS

influenza; influenza; rapid testing; acute respiratory illness; self-collection; self-testing; mHealth; mobile health; home collection; home testing; mobile phone

## Introduction

### Background

In the most recent influenza season in the United States (October 2019 to April 2020), an estimated 39 to 62 million people were infected, resulting in 18 to 26 million health care visits and 24,000 to 62,000 deaths [1]. The economic impacts are proportional—the 2018 seasonal influenza cost the United States an estimated US \$11.2 billion, including US \$3.2 billion in direct medical costs and an estimated 20.1 million productive hours lost [2]. Negative impacts on health and the economy may be improved by early interventions to diagnose those with influenza and intervene with antiviral treatment or behavioral changes to reduce transmission.

Diagnosis of influenza based on clinical features alone is inaccurate; therefore, several clinical guidelines support laboratory testing of respiratory tract specimens (usually nasal or nasopharyngeal) to detect the influenza virus. Increasingly, laboratory testing for influenza has shifted to in-clinic testing using point-of-care (POC) devices [3]. Rapid diagnostic tests (RDTs) are a class of POC tests that can be performed with a few simple steps and typically do not require instrumentation or special supplies, raising the possibility for untrained individuals to use these tests outside of clinical settings [4]. Influenza RDTs for home use could potentially expand access to testing and lower costs, thus facilitating earlier diagnosis and reducing the time from symptom onset to appropriate care, such as receiving antiviral treatment or making behavioral changes to minimize spread [5]. The advantages of home testing for influenza and other respiratory viruses could be even more critical in pandemic situations, where isolating cases and limiting contact with potential cases are essential components of containing outbreaks [6,7].

Several studies have already investigated the accuracy of self-swabbing and self-testing for influenza. A recent systematic review of 13 studies found that influenza was detected by self-collected nasal or midturbinate samples, with similar accuracy to samples collected by health care professionals [8]. RDTs tested in routine health care settings have shown sensitivities and specificities of 60% to 70% and 90% to 100%, respectively [9,10]; however, owing to the novelty of home testing, few RDTs have been studied in the home environment. There are currently no Food and Drug Administration tests approved for the detection of influenza at home.

A primary hurdle to at-home testing for influenza or other respiratory viruses is that RDTs are typically less accurate than

laboratory-based assays, even when used by health care workers [11,12]. Numerous variables affect the performance of the test, including quality of the sample, infection prevalence, timely testing after illness onset, and lower viral load in less severe cases. These variables have not been well-studied in POC settings [9,13-15] or in at-home populations; to our knowledge, only 1 publicly available study has attempted to assess the accuracy and feasibility of performing an entire self-test at home using an RDT [16].

### Objective

In this study, we assess the accuracy of an influenza RDT conducted at home by lay users with influenza-like-illness compared with that of a self-collected sample by the same individual mailed to a laboratory for reference testing.

## Methods

### Study Design

We conducted a prospective, comparative accuracy study of an at-home influenza RDT in a convenience sample of individuals experiencing acute respiratory illness (ARI). The study was conducted as a substudy within the Seattle Flu Study (SFS), which has conducted city-wide community surveillance for influenza and other respiratory viral infections. The SFS involved same-day self-swab samples [17]; participants who qualified and enrolled in the self-test substudy reported here received an additional at-home influenza RDT. The RDT results were compared with the results of a self-collected midturbinate nasal swab sample returned by mail and tested by a laboratory quantitative reverse-transcription polymerase chain reaction (qRT-PCR) assay as described below [17]. Participants also answered a questionnaire that included information about their symptoms, risk factors, and demographics.

### Ethical Approval

The study was approved by the University of Washington Human Subjects Division (STUDY00006181) and informed consent was obtained prior to study enrollment. Reporting of this study adheres to STARD (Standards for Reporting of Diagnostic Accuracy Studies) guidance [18].

### Recruitment

Participants were recruited using targeted web-based advertisements on websites, including Facebook, Instagram, Twitter, and Google. Additional recruitment occurred through in-person referrals from community kiosks set up by the SFS, health care providers, travel clinics, immigrant and refugee

health screenings, local schools, and workplaces. Potential participants were directed to a study website, where they were screened for eligibility based on age, zip code, symptoms, time from symptom onset, and their smartphone operating system. Participants were compensated with US \$20 in the form of an electronic gift card for completing the study procedures; if they returned the reference swab but did not complete the surveys, they were still compensated with US \$15.

### Sample Size

The recruitment goal was to enroll 3000 participants into the study, assuming 2400 (80%) participants would complete all steps and an influenza prevalence rate of 12.5%, which would have provided us with 300 influenza-positive samples. Following the study's completion, electronic gift cards of up to US \$20 were sent to all participants.

### Participants

Participants were enrolled from February 19 to March 9, 2020, from the Greater Seattle area of Washington, United States, which has a population of 744,000. Eligibility criteria included those who self-identified as having a cough or at least two new or worsening ARI symptoms (ie, feeling feverish, headache, chills or shivering, sore throat, nausea or vomiting, runny or stuffy nose, malaise, muscle or body aches, trouble breathing, diarrhea, rash, ear pain, or discharge) in the previous 72 hours [19,20]. In addition, participants were required to be aged  $\geq 5$  years, residing or working within a list of eligible zip codes, able to understand the study instructions in English, and able to use the study app on a Bluetooth-enabled device to conduct study procedures and the Ellume Home Flu Test (EHFT).

### Pretest Data Collection

The participants consented electronically; a parent or legal guardian consented for participants aged <18 years, and assent forms were provided for participants aged 13 to 18 years using REDCap (Research Electronic Data Capture; Vanderbilt University) [21], hosted at the University of Washington Institute of Translational Health Sciences. After consenting, the REDCap instrument obtained the participant's home address and contact information to allow the delivery of an influenza kit through courier.

### Influenza Kit Components and Delivery

Influenza kits were fabricated by the study team to comply with US regulations for shipping biological substances (Category B) [22]. The influenza kit contained 1 instructional quick start guide; an influenza RDT labeled as a research device (EHFT), which included a midturbinate swab, buffer fluid, dropper, and Bluetooth-enabled sample analyzer; and a reference sample kit containing 1 midturbinate swab (Copan, FLOQSwabs 56380CS01), 1 tube with 3 mL of the viral transport medium (VTM; catalog #220220; Becton, Dickinson and Company Ltd), 1 specimen transport bag with absorbent sleeve (cat. #11215-684; VWR International LLC), 1 return box (S-16524; ULIN), and 1 return mailer (S-3355; ULIN) overpack. If participants reported errors with the EHFT, a second kit was sent out as soon as possible during staffing hours (within 12 hours).

Influenza kits were sent to the participants' homes within 24 hours of enrollment, with most sent within 2 hours. Enrollments received after business hours were processed and mailed the following morning. Each kit included a unique barcode number (located at the following three places in each kit: on the 3 mL tube, return mailer, and quick start guide) to link surveys, EHFT results, and reference test results for analysis. Participants were asked to confirm that the barcode on their kit's tube matched that on the quick start guide and enter this barcode in the REDCap survey. In cases where multiple kits were sent to the same household, barcodes entered in REDCap allowed reference samples received in the laboratory to be linked to the correct participants, even when kits were switched between participants. Once participants completed their at-home study procedures, they were instructed to mail their reference sample, using the materials provided, to the University of Washington research laboratory via the US Postal Service within 24 hours of completing their test.

### At-Home Test and Data Collection

Upon receiving their influenza kit at home, participants were instructed to complete a questionnaire on REDCap (sent via email). This included questions about their symptoms and exposure risks, including housing, health conditions, recent travel, and demographics (Multimedia Appendix 1).

Participants were instructed to download the EHFT app onto their Bluetooth-enabled device. The app provided an instructional video, followed by step-by-step on-screen instructions for sample self-collection using the included custom midturbinate swab. Participants were instructed to insert and rotate the swab 3 times around their nasal cavities on both nostrils. They then placed the swab sample in the buffer, added this buffer fluid to the analyzer, and waited for 12 minutes to process the sample. The analyzer then sent test results directly to the EHFT app on the user's device via Bluetooth and a secure research database. As participants were using an experimental research device, they were blinded to the EHFT test results. Participants received a *thank you* screen in the EHFT app once their sample was processed, which instructed them to refer to the study instructions for completing their reference sample and contact their health care provider if they were concerned about their symptoms.

Participants were also asked to obtain a second midturbinate swab using the swab included in their influenza kit, following written and photographic instructions on both the quick start guide and the REDCap survey (Multimedia Appendix 1). They were instructed to insert the swab halfway (approximately 1 inch) into either nostril, press against the side, and rotate 5 times. They were then instructed to place the swab into the collection tube, repackage all components to meet US regulations for shipping biological substances (UN3373 Category B) [23], and return via the US Postal Service to the University of Washington research laboratory.

Participants received a follow-up survey 7 days after enrollment, which included questions about their illness duration and severity, recent travel, and feedback for the research team (Multimedia Appendix 1).

## Reference Testing

Returned kits received in the laboratory were examined, and any evidence of damage to the sample or packaging was documented. Samples were split into 2 aliquots of 1 mL. One aliquot was frozen at  $-80^{\circ}\text{C}$ , and the other was stored at  $4^{\circ}\text{C}$  until extraction. All samples were run in duplicate. Approximately 200  $\mu\text{L}$  of VTM were extracted using Magna Pure 96 small-volume total nucleic acids extraction kit (product #06543588001; Roche). Purified total nucleic acids were tested against a panel of respiratory pathogens using the TaqMan OpenArray platform (ThermoFisher) for qRT-PCR. The OpenArray panel included probe sequences for influenza A H3N2 and influenza A H1N1 and pan influenza A; influenza B; influenza C; respiratory syncytial virus (RSV) A and B; human coronavirus 229E, NL63, OC43, and HKU1; adenovirus; human rhinovirus; human metapneumovirus; human parechovirus; enterovirus A, B, C, D, D68, and G; human bocavirus; and *Streptococcus pneumoniae*, *Mycoplasma pneumoniae*, and *Chlamydia pneumoniae*. The OpenArray panel also included probes for the human gene ribonuclease P (RNase P) as an indicator of sample quality. All quantitative data were captured as relative cycle threshold (Crt) values, which is approximately 10 cycles less than the equivalent quantitative polymerase chain reaction cycle threshold.

Laboratory personnel did not have access to EHFT results or clinical information when interpreting the reference assay. A reference test was considered positive for a pathogen if qRT-PCR generated a fluorescent signal for the channel-specific pathogen within 40 polymerase chain reaction (PCR) cycles. The EHFT or laboratory results were not visible to the participants.

In May 2020, following the completion of data collection, participants were asked if they wanted to opt in to receive the results of their reference swab. If participants opted for results, the second aliquot of their sample was thawed and tested using the Clinical Laboratory Improvement Amendments-waived GeneXpert Xpress (Cepheid) with *Xpert Xpress Flu/RSV* cartridges, following the manufacturer's instructions. Cepheid results distinguished influenza A, influenza B, and RSV. Participants were notified if influenza or RSV were detected. In addition, participants who consented to the study between March 4 and March 9 were notified in real time if SARS-CoV-2 was detected in their reference sample.

## Data Analysis

A participant flow diagram was created demonstrating each major step in the study and participant dropout. Summary statistics were calculated for participant demographics, risk factors, ARI symptoms, symptom severity, and other detected pathogens. Pearson chi-square test with Yate continuity correction was calculated for risk factors, symptom presence, and symptom severity between participants who were reference test positive (PCR positive) and negative (PCR negative). *P* values  $<.05$  were considered statistically significant. We calculated symptom onset as the difference between the self-reported symptom onset date and the exact time at which the EHFT was completed. Participants were instructed to collect the reference swabs immediately after taking the EHFT.

We calculated the sensitivity, specificity, and positive and negative likelihood ratios (with 95% CI) for the overall performance of the index test (EHFT) compared with the reference test (OpenArray qRT-PCR) and independently for influenza A and B. In addition, we analyzed data by subgroups that had previously been shown to affect viral load [24-26], namely symptom onset before testing and illness severity measured as the total number of symptoms (1-9 symptoms) and disruption of daily life caused by their illness (1-5 scale, 1 being not at all and 5 being very much). For each subgroup, we calculated the sensitivity, specificity, and positive and negative likelihood ratios with a 95% CI. We performed pairwise comparisons of the mean level of impact on activities between subgroups using 1-way analysis of variance and Tukey honestly significant difference test. Where appropriate, Pearson correlations were calculated.

The average influenza Crt value was used as a proxy for relative viral load; lower Crt values correspond to higher viral loads and, thus, fewer cycles to generate a sufficient OpenArray signal [27]. Each additional Crt cycle is equivalent to a roughly 2-fold reduction in the genomic copies of viral RNA. Means and SDs were calculated for the following subgroups: symptom onset, influenza subtype, child and adult (5-17 years vs  $\geq 18$  years), and true positive (TP) versus false negative (FN) subgroups. Pairwise comparisons of mean influenza Crt for the subgroups were performed using the Student 2-tailed *t* test. Multiple linear regression models were fitted separately for average Crt values as a function of symptom onset, adjusted for age and number of symptoms and their level of impact on daily activities, both adjusted for age and symptom onset.

The RNase P Crt value was used as an indicator of reference sample quality; a lower Crt corresponds to more endogenous human DNA in the sample, indicating a greater likelihood of sufficient material collected on the swab [25,28,29]. Median RNase P Crt values were compared between age groups and between TP, false positive (FP), FN, and true negative test result subgroups using a Kruskal-Wallis test on ranks and Dunn multiple comparisons post hoc test with a Holm-Bonferroni correction. Median RNase P Crt values were compared between PCR-positive and PCR-negative groups, influenza A-positive and influenza B-positive groups, and between child (aged  $\leq 18$  years) and adult (aged  $>18$  years) groups using the Mann-Whitney *U* test.

Participants with missing or indeterminate EHFT or reference samples were removed from the analysis. The analysis was conducted using R (version 1.3.1056; R Foundation for Statistical Computing) [30].

## Results

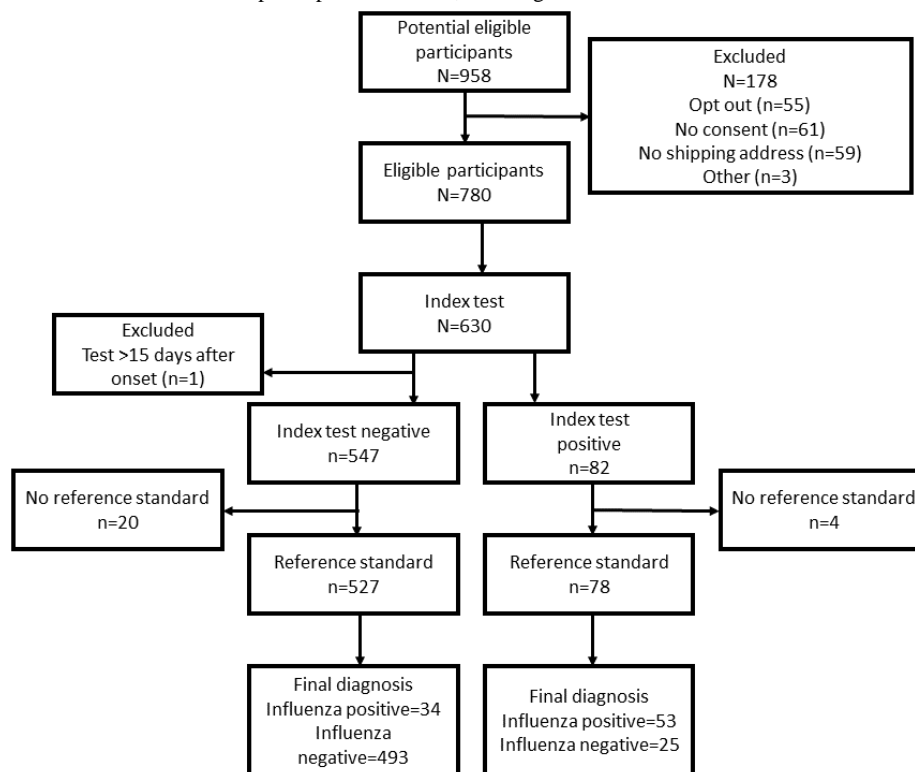
### Participant Recruitment and Retention

A total of 958 participants met the inclusion criteria (Figure 1), of whom 780 (81.4%) completed the consent form, provided a viable shipping address, and were sent an influenza kit. Of these 780 participants who received their kit, 630 (80.8%) completed the index test. One of the individuals who completed the index test 34 days after symptom onset was excluded. Of those 630

participants who completed the index test, 605 (96%) returned their reference samples to the laboratory and were included in our analysis. The final study sample included in this analysis was the 605 participants who completed both the index and reference tests.

Most participants were recruited through advertisements on Facebook, Instagram, or Twitter (184/605, 30.4%); friend or family referrals (181/605, 29.9%); and other web-based media (138/605, 22.8%). Google advertisements and health care provider referrals recruited more influenza-positive participants than other recruitment methods.

**Figure 1.** Recruitment flow for the Home Flu Test participants in Seattle, Washington.



## Participant Demographics

Most participants were aged 25-44 years, female, White, and had completed a bachelor's degree or more (Table 1). Most had private health insurance, had received the 2019 influenza vaccination, and were nonsmokers. Age ( $P=.02$ ) and education ( $P=.02$ ) were significantly associated with a positive influenza PCR test; however, receipt of the 2019-2020 influenza vaccine was not associated with a positive influenza PCR test. The most frequent symptoms reported were fatigue (438/605, 72.4%), sore throat (414/605, 68.4%), cough (355/605, 58.7%), headache (342/605, 56.5%), and fever (292/605, 48.3%; Table 2). Fever, cough, chills or shivering, sweats, nausea or vomiting, myalgia (all  $P$  values  $<.001$ ), fatigue ( $P=.007$ ), runny nose ( $P=.04$ ), and difficulty breathing ( $P=.05$ ) were all significantly associated with a positive PCR test. Participants with positive influenza PCR were also more likely to report moderate or severe levels of fever ( $P<.001$ ), cough ( $P<.001$ ), fatigue ( $P<.001$ ), myalgia ( $P<.001$ ), and sore throat ( $P=.04$ ) compared with those with a

negative PCR test. The TaqMan OpenArray identified 30.7% (186/605) of participants who tested positive for  $\geq 1$  respiratory pathogen other than influenza, some of which were coinfections with influenza A or B (Multimedia Appendix 2). Other than influenza, the most common respiratory pathogens identified were rhinovirus (77/186, 41.2%) and seasonal human coronavirus (50/186, 26.9%).

Almost all (604/605, 99.8%) participants completed the EHFT within 15 days of symptoms onset (range 0.6-14.4 days; Multimedia Appendix 2), with an average time from symptom onset to EHFT testing of 2.9 (SD 1.5) days. Of the total 605 participants, 344 (56.9%) took the EHFT within 72 hours (3 days) after symptom onset, 249 (41.2%) between 4 and 7 days, and 12 (1.9%) between 8 and 15 days after symptom onset. The longest median time interval segment between symptom onset and EHFT testing was between symptom onset and study enrollment (median 48 hours) compared with time from enrollment to kit shipping (median 1.25 hours) and from kit shipping to testing (median 5.33 hours).



**Table 1.** Characteristics of participants with and without influenza detected on reference test (N=605).

Demographics	Overall, n (%)	Influenza PCR <sup>a</sup> positive (n=87), n (%)	Influenza PCR negative (n=518), n (%)	P value
<b>Age (years)</b>				
5-12	29 (4.8)	10 (13.8)	19 (3.7)	.02
13-17	4 (0.7)	1 (1.1)	3 (0.6)	.02
18-24	54 (8.9)	6 (6.9)	48 (9.3)	.02
25-34	222 (36.8)	23 (26.4)	199 (38.4)	.02
35-44	159 (26.2)	26 (29.9)	133 (25.7)	.02
45-64	112 (18.5)	18 (20.7)	94 (18.1)	.02
≥65	12 (2)	1 (1.1)	11 (2.1)	.02
<b>Gender</b>				
Male	217 (36)	35 (40.2)	182 (35.1)	.07
Female	360 (59.4)	47 (54)	313 (60.4)	.07
Other	4 (0.7)	2 (2.3)	2 (0.4)	.07
<b>Education</b>				
Less than high school	3 (0.5)	2 (2.3)	1 (0.2)	.02
Graduated high school or GED <sup>b</sup>	15 (2.5)	0 (0)	15 (2.9)	.02
Some college	55 (9.2)	5 (5.7)	50 (9.7)	.02
Bachelor's degree	234 (38.6)	37 (42.5)	197 (38)	.02
Advanced degree	236 (38.9)	28 (32.2)	208 (40.2)	.02
<b>Race<sup>c</sup></b>				
White	430 (71)	69 (79.3)	361 (69.7)	.11
Black	9 (1.5)	3 (3.4)	6 (1.2)	.25
Asian	137 (22.8)	10 (11.5)	127 (24.5)	.01
AI <sup>d</sup> , AN <sup>e</sup> , NH <sup>f</sup> , or PI <sup>g</sup>	10 (1.7)	0 (0)	10 (1.9)	N/A <sup>h</sup>
Other	21 (3.5)	3 (3.4)	18 (3.5)	.99
Hispanic	31 (5.1)	2 (2.3)	29 (5.6)	.31
<b>Health insurance<sup>c</sup></b>				
Public	42 (6.9)	4 (4.6)	38 (7.3)	.48
Private	519 (85.8)	76 (87.4)	443 (85.5)	.91
Other	7 (1.2)	1 (1.1)	6 (1.2)	.99
No insurance	11 (1.8)	3 (3.4)	8 (1.5)	.43
Taken 2019-2020 influenza vaccination	424 (70)	62 (71.3)	362 (69.9)	.78
Not a current smoker	516 (85.3)	71 (81.6)	445 (85.9)	.44

<sup>a</sup>PCR: polymerase chain reaction.

<sup>b</sup>GED: General Education Development.

<sup>c</sup>Total values may exceed 100% as participants could select multiple responses.

<sup>d</sup>AI: American Indian.

<sup>e</sup>AN: Alaskan Native.

<sup>f</sup>NH: Native Hawaiian.

<sup>g</sup>PI: Pacific Islander.

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

**Table 2.** Presence and severity of symptoms reported by participants with and without influenza detected on reference test (N=605).

Symptoms	Overall, n (%)	Influenza PCR <sup>a</sup> positive (n=87), n (%)	Influenza PCR negative (n=518), n (%)	P value
<b>Feeling feverish<sup>b</sup></b>	292 (48.3)	76 (87.4)	216 (41.7)	<.001
Mild	119 (19.7)	14 (16.1)	105 (20.3)	<.001
Moderate	122 (20.2)	34 (39.1)	88 (17)	<.001
Severe	51 (8.4)	28 (32.2)	23 (4.4)	<.001
Headache	342 (56.5)	63 (72.4)	279 (53.9)	.002
<b>Cough<sup>b</sup></b>	355 (58.7)	72 (82.8)	283 (54.6)	<.001
Mild	168 (27.8)	17 (19.5)	151 (29.2)	<.001
Moderate	150 (24.8)	37 (42.5)	113 (21.8)	<.001
Severe	36 (6)	18 (20.7)	18 (3.5)	<.001
Chills or shivering	227 (37.5)	68 (78.2)	159 (30.7)	<.001
Sweats	142 (23.5)	43 (49.4)	99 (19.1)	<.001
<b>Sore, itchy, or scratchy throat<sup>b</sup></b>	414 (68.4)	57 (65.5)	357 (68.9)	.53
Mild	174 (28.8)	27 (31)	147 (28.4)	.04
Moderate	189 (31.2)	18 (20.7)	171 (33)	.04
Severe	49 (8.1)	11 (12.6)	38 (7.3)	.04
Nausea or vomiting	72 (11.9)	21 (24.1)	51 (9.8)	<.001
Running or stuffy nose	358 (59.2)	61 (70.1)	297 (57.3)	.04
<b>Feeling more tired than usual<sup>b</sup></b>	438 (72.4)	74 (85.1)	364 (70.3)	.007
Mild	105 (17.4)	11 (12.6)	94 (18.1)	<.001
Moderate	217 (35.9)	27 (31)	190 (36.7)	<.001
Severe	116 (19.2)	36 (41.4)	80 (15.4)	<.001
<b>Muscle or body aches<sup>b</sup></b>	300 (49.6)	65 (74.7)	235 (45.4)	<.001
Mild	99 (16.4)	4 (4.6)	95 (18.3)	<.001
Moderate	146 (24.1)	37 (42.5)	109 (21)	<.001
Severe	55 (9.1)	24 (27.6)	31 (6)	<.001
Increased trouble with breathing	134 (22.1)	27 (31)	107 (20.7)	.047
Diarrhea <sup>c</sup>	69 (11.4)	6 (6.9)	63 (12.2)	.19
Rash <sup>c</sup>	9 (1.5)	0 (0)	9 (1.7)	.44
Ear pain or discharge <sup>c</sup>	54 (8.9)	11 (12.6)	43 (8.3)	.31

<sup>a</sup>PCR: polymerase chain reaction.

<sup>b</sup>Included survey questions on severity for a subset of symptoms presented in this table.

<sup>c</sup>Symptoms reported for children (aged <18 years) only.

## Accuracy of RDT

Of the total 605 participants, 87 (14.4%) participants tested positive for influenza using the reference test (70/87, 80% for influenza A and 17/87, 20% for influenza B). The overall sensitivity and specificity of the EHFT compared with the reference test were 61% (95% CI 50-71) and 95% (95% CI 93-97), respectively (see Table 3 for accuracy values). The sensitivity and specificity of the EHFT for influenza A were 60% (95% CI 48-72) and 99% (95% CI 98-100), respectively, and those for influenza B were 65% (95% CI 38-86) and 96%

(95% CI 94-98), respectively. Influenza B yielded a higher rate of FP tests, which resulted in a positive predictive value (PPV) of 33% (95% CI 18-52).

Subgroup analysis of EHFT accuracy based on the time from symptom onset to conducting the EHFT (Table 3) showed that the sensitivity and specificity of influenza A detection at ≤72 hours of symptom onset were 62.5% (95% CI 48-77) and 100% (95% CI 98-100), respectively, whereas at >72 hours they were 57% (95% CI 37-75) and 99% (95% CI 97-100), respectively. The sensitivity and specificity of influenza B at ≤72 hours were 64% (95% CI 31-89) and 95% (95% CI 92-97), respectively,

whereas, at >72 hours, they were 67% (95% CI 22-96) and 98% (95% CI 95-99), respectively.

There were no associations between illness severity and EHFT sensitivity (Multimedia Appendix 3). Neither the number of symptoms nor disruption to daily activities were significantly associated with EHFT sensitivity, nor was there a meaningful association between either of these measures and mean influenza Crt. However, measures of illness severity were correlated with each other; individuals who reported more disruption to daily activities also reported a greater number of symptoms ( $r=0.54$ ;

$P<.001$ ), and individuals who were PCR positive reported significantly higher scores for both these measures compared with individuals who were PCR negative ( $P<.001$ ; Multimedia Appendix 3).

Of the 25 FP results, 22 (88%) occurred when the EHFT indicated the presence of influenza B, and all occurred within 96 hours of symptom onset (Multimedia Appendix 3). Other respiratory pathogens, namely RSV, human metapneumovirus, human coronavirus, and *Streptococcus pneumoniae*, were detected in 32% (7/22) of the influenza B FP samples.

**Table 3.** Accuracy of the Ellume home influenza test compared with laboratory reference polymerase chain reaction values (N=605).

Accuracy	True positive, n (%)	False positive, n (%)	False negative, n (%)	True negative, n (%)	Sensitivity (95% CI)	Specificity (95% CI)	PPV <sup>a</sup> (95% CI)	NPV <sup>b</sup> (95% CI)
<b>Overall</b>								
All	53 (8.8)	25 (4.1)	34 (5.6)	493 (81.5)	61 (50-71)	95 (93-97)	68 (56-78)	94 (91-95)
≤72 hours (n=344)	32 (9.3)	17 (4.9)	19 (5.5)	276 (80.2)	63 (48-76)	94 (91-97)	65 (50-78)	94 (90-96)
>72 hours (n=261)	21 (8)	8 (3.1)	15 (5.7)	217 (83.1)	58 (41-74)	96 (93-98)	72 (53-87)	94 (90-96)
<b>Influenza A</b>								
All	42 (6.9)	3 (0.5)	28 (4.6)	532 (87.9)	60 (48-72)	99 (98-100)	93 (82-99)	95 (93-97)
≤72 hours (n=344)	25 (7.3)	1 (0.3)	15 (4.4)	303 (88.1)	63 (46-77)	100 (98-100)	96 (80-100)	95 (92-97)
>72 hours (n=261)	17 (6.5)	2 (0.8)	13 (5)	229 (87.7)	57 (37-75)	99 (97-100)	89 (67-99)	95 (91-97)
<b>Influenza B</b>								
All	11 (1.8)	22 (3.6)	6 (1)	566 (93.6)	65 (38-86)	96 (94-98)	33 (18-52)	99 (98-100)
≤72 hours (n=344)	7 (2)	16 (4.7)	4 (1.2)	317 (92.2)	64 (31-89)	95 (92-97)	30 (13-53)	99 (97-100)
>72 hours (n=261)	4 (1.5)	6 (2.3)	2 (0.8)	249 (95.4)	67 (22-96)	98 (95-99)	40 (12-74)	99 (97-100)

<sup>a</sup>PPV: positive predictive value.

<sup>b</sup>NPV: negative predictive value.

### Influenza Crt

The average influenza Crt value for all influenza PCR-positive samples was 18.82 and was significantly lower for samples collected ≤72 hours after symptom onset (17.9, SD 4.4), than those collected at >72 hours (20.14, SD 4.2;  $P=.02$ ; Table 4). The mean influenza Crt (16.8, SD 4.0) for individuals who were TP was significantly lower than the mean Crt value (22.0, SD

3.0) for individuals who were FN ( $P<.001$ ). The mean influenza Crt values were also significantly lower for influenza B (16.7, SD 4.4) than for influenza A (19.3, SD 4.3;  $P=.04$ ). Multiple regression estimated that the average Crt increased by 1.34 (95% CI 39-2.29) Crt for each additional day after symptom onset ( $P=.006$ ), and Crt values decreased with a greater number of symptoms ( $P=.03$ ; Multimedia Appendix 4).

**Table 4.** Influenza relative cycle threshold (Crt) of various subgroups (N=87).

Group	Influenza-positive participants, n (%)	Influenza Crt raw, mean (SD)	P value
<b>Influenza subtype</b>			
Influenza A	70 (80)	19.34 (4.3)	.04
Influenza B	17 (20)	16.71 (4.4)	.04
<b>72-hour testing cutoff</b>			
≤72 hours	N/A <sup>a</sup>	17.9 (4.4)	.02
>72 hours	N/A	20.14 (4.2)	.02
<b>Age<sup>b</sup> (years)</b>			
5-17	11 (13)	16.6 (4.5)	.10
≥18	73 (84)	19.1 (4.4)	.10
<b>True positives and false negatives</b>			
True positives	53 (61)	16.8 (4.0)	<.001
False negatives	34 (39)	22.0 (3.0)	<.001

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

<sup>b</sup>A total of 3 participants who were influenza positive were missing age data.

## User Experiences With Study Procedures and Specimen Quality

Overall, participants stated they were *somewhat confident* (216/560, 38.5%) or *very confident* (337/560, 60.1%) that they completed the reference swab correctly and experienced only mild discomfort (431/560, 76.8%; [Multimedia Appendix 5](#)). This was similar for the EHFT, for which participants stated they were *somewhat confident* (188/567, 33.5%) or *very confident* (371/567, 66.1%) and experienced only mild discomfort (427/567, 76.1%). Only 0.3% (2/567) of participants reported errors with the Bluetooth component of the EHFT device and were sent new devices. No other issues with the device were reported to the study team.

Reference sample RNase P Crt values ranged from 10.89 cycles to 33.3 cycles (median 21.1, SD 4.3, IQR 16.7-24.1). RNase P Crt did not vary significantly between influenza-positive samples (median 22.7, SD 4.32) and influenza-negative samples (median 20.7, SD 4.32;  $P=.05$ ) or based on age or influenza subtype ([Multimedia Appendix 5](#)). All samples had Rnase P Crt values well below the 40-cycle cutoff value, and only 4 were >30 (31.5, 32.4, 32.5, and 33.3). Median values for all subgroups assessed were well within the recommended quality range of 28 Crt [31]. No samples were excluded on the basis of their RNase P Crt values.

In 7.8% (46/587) of the returned kits, ≥1 error was noted, indicating that these participants did not correctly follow ≥1 provided instruction. Of these 46 kits, 23 (50%) were returned with a packaging error (either missing the outer box or specimen bag sealed incorrectly), and 27 (59%) were returned with incorrect labeling on the VTM tube ([Multimedia Appendix 5](#)).

## Discussion

### Principal Findings

This study demonstrated the feasibility of implementing an unsupervised at-home diagnostic test. The vast majority of participants were able to complete the multiple procedures required to evaluate a home influenza test without direct supervision, including surveys, 2 midturbinate swabs, app-guided directions to complete an influenza RDT, and returning reference samples by mail to a central laboratory. The influenza positivity within our study sample was 14.4% (87/605), with 80% (70/87) of influenza A cases, which is consistent with both the prevalence and relative proportion of influenza strains reported in the local area during the study period [32]. The EHFT had moderate sensitivity (61%) and high specificity (95%) compared with laboratory PCR on self-collected swabs. Specificity was slightly higher (99%) for influenza A than for influenza B (96%), whereas sensitivity was slightly higher (65%) for influenza B than for influenza A (60%), although the CIs were wide and overlapping. The small proportion of participants who were PCR positive for influenza B and the high rate of influenza B FPs resulted in a much lower PPV for influenza B (33%) when analyzed independently from influenza A (68%).

TP EHFT results had significantly lower influenza Crt values (corresponding to higher viral load) than FN EHFT results, suggesting that lower viral load may have affected the sensitivity of EHFT. To further investigate this relationship, we assessed EHFT accuracy across 2 additional variables known to affect viral load, namely, symptom onset and illness severity [24-26,33,34]. EHFT sensitivity was related to symptom onset, with a moderate improvement in sensitivity (6%) when the test was conducted within 72 hours of symptom onset. Furthermore, we noted a linear relationship between symptom onset and Crt value, where each additional day between symptom onset and testing corresponded to an average of 1.3 additional cycles (ie,

more than a 2-fold decrease) in the estimated quantity of the virus. In contrast, neither did illness severity appear to influence EHFT sensitivity, with no relationship observed between test accuracy and number of symptoms, nor did it have a greater impact on daily activities. Nevertheless, these 2 measures of illness severity were correlated with each other, which is consistent with the expectation that individuals who report more symptoms face greater disruption to their daily activities. Notably, both measures of influenza severity were significantly higher in individuals who were PCR positive than in individuals who were PCR negative, despite not predicting viral load in this study.

We did not find any evidence that the quality of self-sampling affected EHFT accuracy or that particular demographic groups were more capable of collecting a self-swab than others. In contrast, we found that negative reference swabs had a lower RNase P Crt value than the positive samples. If sample quality affected the reference swab, we would expect the opposite—negative reference swabs to be of poorer quality and thus higher RNase P Crt values. The median RNase P values for the test accuracy subgroups were well within the acceptable range.

### Comparison With Prior Studies

The accuracy of the EHFT we reported was comparable with that of other influenza RDTs. A 2017 meta-analysis of 134 studies of influenza RDTs showed pooled estimates of sensitivity (61%, 95% CI 53.3-68.3) and specificity (98.9%, 95% CI 98.4-99.3) [14] that are comparable with those reported in this study. These results are consistent with those of 2 other meta-analyses [10,14]. The similarity in test accuracy is even more notable considering that published studies on influenza RDT accuracy were conducted in health care settings, with sampling and RDTs performed by health care workers and researchers rather than patients themselves.

The time from symptom onset to testing, or symptom onset, has a critical impact on viral load and influenza RDT accuracy [13,35]. One of the disadvantages of mailed testing kits is the delay in testing following symptom onset because of the time needed to distribute swabbing materials [16,36]. Elliot et al [36] reported an average of 4 days between symptom onset and self-swabbing compared with 2 days for clinician-collected samples. Similarly, we found that influenza Crt decreased with a longer symptom onset [36]. Studies that have lower mean times from onset to testing tend to report higher sensitivity values [13]. In contrast, 1 study found that testing too early can lead to increased FNs, primarily if the RDT is used within 12 hours of symptom onset [15], suggesting that there might be a *sweet spot* for RDT testing for influenza that must be balanced with other factors that affect test sensitivity. In addition, we did not find a relationship between viral shedding and illness severity [24-26,33,34]. This may be as the measures of illness severity we used lacked validity in our setting or population, and the range of illness severities in our population was too narrow. A more robust understanding of viral load dynamics, especially in less severely ill populations, will help delineate the conditions under which an at-home RDT for influenza is most appropriate.

The quality of self-collected swabs did not appear to affect the EHFT accuracy. Participants reported high confidence in completing both the EHFT swab and reference test swab. Moreover, reference swabs had RNase P Crt values (when corrected to an equivalent cycle threshold value) that tended to be high but within the range of those reported in other studies of both clinician-collected and self-collected midturbinate swabs [28,37]. The EHFT swab instructions asked the participants to swab both nostrils. A possible explanation for the low RNase P Crt values on the reference swab, which was completed after the EHFT swab, is that there was less human cellular debris available for collection. This may also have affected the influenza Crt values. Although there was variability in RNase P Crt between individuals, variability was not observed between TP, FN, FP, and true negative groups, suggesting that sample collection did not affect EHFT accuracy.

Our findings of inferior accuracy of the EHFT for influenza B (including FP results) are consistent with other literature [38-40] and may have been because of several factors. The prevalence of influenza B in the study catchment area was low (3%-4%) during the study period [32]; low disease prevalence is known to affect predictive values [41,42]. Other studies of influenza RDTs have also noted higher rates of FPs for influenza B than influenza A [38-40], suggesting nonspecific reactivity with antibodies used for influenza B detection.

### Strengths and Limitations

This is one of the first studies to report the accuracy of an influenza RDT used by unsupervised participants and the first to do so for an RDT designed specifically for home use. Our study design included remote web-based recruitment, shipping of influenza kits complete with RDT and reference sample collection materials, and completion of all stages of the study by the user without direct supervision from the study staff. The high response and completion rate of study procedures (605/780, 77.6%) were matched by high self-reported confidence for both the EHFT and reference sample procedures. There was a 22.4% (175/780) dropout of participants who were sent a kit but did not participate; it is unclear whether this introduced additional bias but is consistent with other mail-based testing studies [43]. Although occasional errors in the required shipping procedures occurred, the vast majority of samples were returned to the research laboratory in the appropriate condition. Participants likely had milder symptoms than those attending clinical settings, and although this may have affected RDT sensitivity with lower viral shedding, they represented the population in which this RDT would be used. Our findings support this type of study design for the assessment of self-tests for influenza and other respiratory viruses such as SARS-CoV-2.

This study had several limitations. First, participants were more highly educated, were English speakers with private insurance and access, and had the ability to use a Bluetooth mobile app device. Future studies may address this limitation through varied recruitment approaches outside of the social media advertisements used in this study; however, this type of study design depends on internet- and app-based data collection, which inherently limits the study population. Second, many individuals (261/605, 43.1%) did not conduct the EHFT within 72 hours

of illness onset, likely because of the time elapsed from recruitment, shipment of kits, and participants' availability to complete the EHFT on receipt. Future studies of at-home tests should explore solutions to identify symptomatic individuals earlier in their illness and expediently provide tests. For example, cohort studies have used regular, self-reported symptom surveys via SMS text messaging or email to identify influenza early and prompt testing [44,45] or through prepositioning of influenza kits or other at-home testing devices. Third, although participants reported confidence in self-collection of swabs, confirmed by markers of human DNA in these samples, there remains some uncertainty regarding the validity of this type of reference sample. Finally, we recruited less than one-third of the desired sample size (and only a small number of influenza B infections) because of delays in study initiation and premature closure as a result of the COVID-19 pandemic in the local area. A larger sample size would have provided tighter CIs around the estimates of test accuracy. In addition, it is possible that the circulation of SARS-CoV-2 in the community from which we recruited affected the accuracy of the EHFT. The enrollment criteria were intended to be specific to ARI and thus may have incidentally recruited individuals with COVID-19 who presented with symptoms similar to influenza. Unfortunately, because of privacy protocols implemented by our laboratory in conjunction with the county public health office, we were not able to analyze the reference samples for SARS-CoV-2 and were unable to determine the impact it may have had on the EHFT accuracy. On the basis of the seroprevalence of SARS-CoV-2 around this time [46], it seems unlikely that there were enough COVID-19 cases circulating in the Seattle area to cause a major reduction in test performance, although it may have had a small impact on PPV for ARI.

### Implications for Clinicians, Researchers, and Policy Makers

RDTs designed for home use have the potential to be purchased over the counter or prescribed by health care providers and coupled, if necessary, with in-person or telemedicine consultations to guide care [47]. The accuracy of the influenza EHFT reported here is similar to that of many RDTs used in clinical settings, which supports its use in similar populations,

provided suitable precautions are in place, particularly to mitigate the risk of FN results. These could include using clinical prediction rules to assist potential RDT users in quantifying their pretest probability of influenza. Our findings support the use of the EHFT among individuals within 72 hours of symptom onset and suggest the need for further research to understand other indicators of viral load that could be used to select individuals for whom this type of RDT should or should not be recommended.

Our study design provides a model for comparative accuracy studies of RDTs for influenza and other respiratory pathogens, including SARS-CoV-2, in home settings. We recommend that future study designs prioritize minimizing the time from symptom onset through study enrollment to conduct the index test, particularly for infections such as influenza, where viral shedding declines rapidly after symptom onset. Strategies could include prepositioning test kits and encouraging early completion of RDT with the onset of illness. Given the potentially important relationship between influenza severity and viral load (and hence self-test sensitivity), we also encourage the use of more accurate ways of measuring illness severity from self-reported surveys. Finally, rather than blinding participants to RDT results, revealing self-test results would facilitate recruitment and allow exploration of the impacts of positive and negative self-test results on participants' health-seeking and other behaviors.

### Conclusions

Using an entirely community-based, remote recruitment study design, our findings showed that the EHFT had comparable accuracy to many influenza RDTs used in clinical settings. However, the sensitivity of the EHFT was only moderate and was higher when the test was used within 72 hours of symptom onset when virus shedding was likely the highest. Our findings support a new form of trial design, in which recruitment and self-sampling for reference testing can be performed successfully by lay users in the communities and populations in which these tests will be implemented. Such study designs could be used to assess the accuracy of tests for other viral respiratory tract pathogens, such as SARS-CoV-2 and RSV. Home tests have the potential to expand access to testing for infectious diseases, with potential benefits for individuals and the health care system.

### Acknowledgments

The authors would like to acknowledge Rosemichelle Marzan, Grace Kim, Ashley Song, Kendall Escene, and Angel Wong for their critical help in constructing the study kits.

The authors would also like to acknowledge Lincoln C Pothan and her logistics team for orchestrating the extremely expedient shipping and distribution of test kits to participants. Finally, the authors would like to acknowledge Jaclyn Ruckle, Jen Mooney, and the formative team for designing persuasive advertisements to assist with study recruitment. The Seattle Flu Study was funded by Gates Ventures. Neither was the funder involved in the design of the study, nor did they have any ownership over the management and conduct of the study, the data, or the rights to publish.

### Conflicts of Interest

HYC has received research support from GlaxoSmithKline, Novavax, and Sanofi Pasteur. BL is a cofounder of and holds equity in a company (Anavasi Diagnostics, Inc) that is developing a point-of-care respiratory disease test. MJT has received reimbursement as a medical advisor for Click Diagnostics (now Visby Medical), Inflammatrix, and Roche Molecular Diagnostics, which are all

companies that are involved in point-of-care tests for infectious disease pathogens. MJT is employed part time as Senior Director of Clinical Science at Inflammatrix Inc. JHK also holds equity in Anavasi Diagnostics; the study in this manuscript was completed before this company was founded.

---

#### Multimedia Appendix 1

Additional details on participant data collection.

[DOCX File , 343 KB - [publichealth\\_v8i2e28268\\_app1.docx](#) ]

---

#### Multimedia Appendix 2

Additional participant demographics.

[DOCX File , 18 KB - [publichealth\\_v8i2e28268\\_app2.docx](#) ]

---

#### Multimedia Appendix 3

Additional details on rapid diagnostic test accuracy.

[DOCX File , 27 KB - [publichealth\\_v8i2e28268\\_app3.docx](#) ]

---

#### Multimedia Appendix 4

Regression adjusted mean influenza relative cycle threshold.

[DOCX File , 19 KB - [publichealth\\_v8i2e28268\\_app4.docx](#) ]

---

#### Multimedia Appendix 5

User experience with study procedures and specimen quality.

[DOCX File , 31 KB - [publichealth\\_v8i2e28268\\_app5.docx](#) ]

---

## References

1. 2021-2022 U.S. flu season: preliminary in-season burden estimates. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/flu/about/burden/preliminary-in-season-estimates.htm> [accessed 2022-01-26]
2. Putri WC, Muscatello DJ, Stockwell MS, Newall AT. Economic burden of seasonal influenza in the United States. *Vaccine* 2018 Jun 22;36(27):3960-3966. [doi: [10.1016/j.vaccine.2018.05.057](#)] [Medline: [29801998](#)]
3. Clinical Laboratory Improvement Amendments (CLIA). Centers for Medicare & Medicaid Services. URL: <https://www.cms.gov/Regulations-and-Guidance/Legislation/CLIA> [accessed 2022-01-26]
4. Basile K, Kok J, Dwyer DE. Point-of-care diagnostics for respiratory viral infections. *Expert Rev Mol Diagn* 2018 Jan;18(1):75-83 [FREE Full text] [doi: [10.1080/14737159.2018.1419065](#)] [Medline: [29251007](#)]
5. Egilmezer E, Walker GJ, Bakthavathsalam P, Peterson JR, Gooding JJ, Rawlinson W, et al. Systematic review of the impact of point-of-care testing for influenza on the outcomes of patients with acute respiratory tract infection. *Rev Med Virol* 2018 Sep;28(5):e1995. [doi: [10.1002/rmv.1995](#)] [Medline: [30101552](#)]
6. Qualls N, Levitt A, Kanade N, Wright-Jegede N, Dopson S, Biggerstaff M, CDC Community Mitigation Guidelines Work Group. Community mitigation guidelines to prevent pandemic influenza - United States, 2017. *MMWR Recomm Rep* 2017 Apr 21;66(1):1-34. [doi: [10.15585/mmwr.rr6601a1](#)] [Medline: [28426646](#)]
7. Fong MW, Gao H, Wong JY, Xiao J, Shiu EY, Ryu S, et al. Nonpharmaceutical measures for pandemic influenza in nonhealthcare settings-social distancing measures. *Emerg Infect Dis* 2020 May;26(5):976-984 [FREE Full text] [doi: [10.3201/eid2605.190995](#)] [Medline: [32027585](#)]
8. Seaman CP, Tran LT, Cowling BJ, Sullivan SG. Self-collected compared with professional-collected swabbing in the diagnosis of influenza in symptomatic individuals: a meta-analysis and assessment of validity. *J Clin Virol* 2019 Sep;118:28-35. [doi: [10.1016/j.jcv.2019.07.010](#)] [Medline: [31400670](#)]
9. Chartrand C, Leeftang MM, Minion J, Brewer T, Pai M. Accuracy of rapid influenza diagnostic tests: a meta-analysis. *Ann Intern Med* 2012 Apr 03;156(7):500-511 [FREE Full text] [doi: [10.7326/0003-4819-156-7-201204030-00403](#)] [Medline: [22371850](#)]
10. Merckx J, Wali R, Schiller I, Caya C, Gore GC, Chartrand C, et al. Diagnostic accuracy of novel and traditional rapid tests for influenza infection compared with reverse transcriptase polymerase chain reaction: a systematic review and meta-analysis. *Ann Intern Med* 2017 Sep 19;167(6):394-409. [doi: [10.7326/M17-0848](#)] [Medline: [28869986](#)]
11. Rumbold B, Wenham C, Wilson J. Self-tests for influenza: an empirical ethics investigation. *BMC Med Ethics* 2017 May 05;18(1):33 [FREE Full text] [doi: [10.1186/s12910-017-0192-y](#)] [Medline: [28476115](#)]
12. Green DA, StGeorge K. Rapid antigen tests for influenza: rationale and significance of the FDA reclassification. *J Clin Microbiol* 2018 Oct;56(10):e00711-e00718 [FREE Full text] [doi: [10.1128/JCM.00711-18](#)] [Medline: [29899007](#)]

13. Chan MC, Lee N, Ngai KL, Leung TF, Chan PK. Clinical and virologic factors associated with reduced sensitivity of rapid influenza diagnostic tests in hospitalized elderly patients and young children. *J Clin Microbiol* 2014 Feb;52(2):497-501 [FREE Full text] [doi: [10.1128/JCM.02316-13](https://doi.org/10.1128/JCM.02316-13)] [Medline: [24478479](https://pubmed.ncbi.nlm.nih.gov/24478479/)]
14. Bruning AH, Leeftang MM, Vos JM, Spijker R, de Jong MD, Wolthers KC, et al. Rapid tests for influenza, respiratory syncytial virus, and other respiratory viruses: a systematic review and meta-analysis. *Clin Infect Dis* 2017 Sep 15;65(6):1026-1032 [FREE Full text] [doi: [10.1093/cid/cix461](https://doi.org/10.1093/cid/cix461)] [Medline: [28520858](https://pubmed.ncbi.nlm.nih.gov/28520858/)]
15. Tanei M, Yokokawa H, Murai K, Sakamoto R, Amari Y, Boku S, et al. Factors influencing the diagnostic accuracy of the rapid influenza antigen detection test (RIADT): a cross-sectional study. *BMJ Open* 2014 Jan 02;4(1):e003885 [FREE Full text] [doi: [10.1136/bmjopen-2013-003885](https://doi.org/10.1136/bmjopen-2013-003885)] [Medline: [24384898](https://pubmed.ncbi.nlm.nih.gov/24384898/)]
16. Zigman Suchsland ML, Rahmatullah I, Lutz B, Lyon V, Huang S, Kline E, Seattle Flu Study investigators. Evaluating an app-guided self-test for influenza: lessons learned for improving the feasibility of study designs to evaluate self-tests for respiratory viruses. *BMC Infect Dis* 2021 Jun 29;21(1):617 [FREE Full text] [doi: [10.1186/s12879-021-06314-1](https://doi.org/10.1186/s12879-021-06314-1)] [Medline: [34187397](https://pubmed.ncbi.nlm.nih.gov/34187397/)]
17. Chu H, Boeckh M, Englund J, Famulare M, Lutz B, Nickerson D, et al. The Seattle Flu Study: a multiarm community-based prospective study protocol for assessing influenza prevalence, transmission and genomic epidemiology. *BMJ Open* 2020 Oct 07;10(10):e037295 [FREE Full text] [doi: [10.1136/bmjopen-2020-037295](https://doi.org/10.1136/bmjopen-2020-037295)] [Medline: [33033018](https://pubmed.ncbi.nlm.nih.gov/33033018/)]
18. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* 2016 Nov 14;6(11):e012799 [FREE Full text] [doi: [10.1136/bmjopen-2016-012799](https://doi.org/10.1136/bmjopen-2016-012799)] [Medline: [28137831](https://pubmed.ncbi.nlm.nih.gov/28137831/)]
19. Powers JH, Guerrero ML, Leidy NK, Fairchok MP, Rosenberg A, Hernández A, et al. Development of the Flu-PRO: a patient-reported outcome (PRO) instrument to evaluate symptoms of influenza. *BMC Infect Dis* 2016 Jan 05;16:1 [FREE Full text] [doi: [10.1186/s12879-015-1330-0](https://doi.org/10.1186/s12879-015-1330-0)] [Medline: [26729246](https://pubmed.ncbi.nlm.nih.gov/26729246/)]
20. Ebell MH, Afonso AM, Gonzales R, Stein J, Genton B, Senn N. Development and validation of a clinical decision rule for the diagnosis of influenza. *J Am Board Fam Med* 2012;25(1):55-62 [FREE Full text] [doi: [10.3122/jabfm.2012.01.110161](https://doi.org/10.3122/jabfm.2012.01.110161)] [Medline: [22218625](https://pubmed.ncbi.nlm.nih.gov/22218625/)]
21. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009 Apr;42(2):377-381 [FREE Full text] [doi: [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]
22. Guidance on regulations for the transport of infectious substances 2019--2020. World Health Organization. URL: <https://apps.who.int/iris/bitstream/handle/10665/325884/WHO-WHE-CPI-2019.20-eng.pdf> [accessed 2022-01-26]
23. Guidance on regulations for the transport of infection substances. World Health Organization. URL: [https://www.who.int/csr/resources/publications/biosafety/WHO\\_CDS\\_EPR\\_2007\\_2.pdf](https://www.who.int/csr/resources/publications/biosafety/WHO_CDS_EPR_2007_2.pdf) [accessed 2022-01-26]
24. Lee N, Chan PK, Hui DS, Rainer TH, Wong E, Choi K, et al. Viral loads and duration of viral shedding in adult patients hospitalized with influenza. *J Infect Dis* 2009 Aug 15;200(4):492-500 [FREE Full text] [doi: [10.1086/600383](https://doi.org/10.1086/600383)] [Medline: [19591575](https://pubmed.ncbi.nlm.nih.gov/19591575/)]
25. Rodrigues Guimarães Alves V, Perosa AH, de Souza Luna LK, Cruz JS, Conte DD, Bellei N. Influenza A(H1N1)pdm09 infection and viral load analysis in patients with different clinical presentations. *Mem Inst Oswaldo Cruz* 2020;115:e200009 [FREE Full text] [doi: [10.1590/0074-02760200009](https://doi.org/10.1590/0074-02760200009)] [Medline: [32428082](https://pubmed.ncbi.nlm.nih.gov/32428082/)]
26. Lau LL, Cowling BJ, Fang VJ, Chan K, Lau EH, Lipsitch M, et al. Viral shedding and clinical illness in naturally acquired influenza virus infections. *J Infect Dis* 2010 May 15;201(10):1509-1516 [FREE Full text] [doi: [10.1086/652241](https://doi.org/10.1086/652241)] [Medline: [20377412](https://pubmed.ncbi.nlm.nih.gov/20377412/)]
27. Crt, a relative threshold method for qPCR data analysis on the QuantStudio 12K Flex system with OpenArray technology. Thermo Fisher Scientific. URL: [https://assets.thermofisher.com/TFS-Assets/LSG/brochures/CO28730-Crt-Tech-note\\_FLR.pdf](https://assets.thermofisher.com/TFS-Assets/LSG/brochures/CO28730-Crt-Tech-note_FLR.pdf) [accessed 2022-01-26]
28. Van Wesenbeeck L, Meeuws H, D'Haese D, Ispas G, Houspie L, Van Ranst M, et al. Sampling variability between two mid-turbinate swabs of the same patient has implications for influenza viral load monitoring. *Virology* 2014 Dec 24;11:233 [FREE Full text] [doi: [10.1186/s12985-014-0233-9](https://doi.org/10.1186/s12985-014-0233-9)] [Medline: [25539740](https://pubmed.ncbi.nlm.nih.gov/25539740/)]
29. CDC protocol of realtime RTPCR for swine influenza A(H1N1). World Health Organization. URL: <https://pesquisa.bvsalud.org/portal/resource/pt/des-17766> [accessed 2022-01-26]
30. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2013.
31. Respiratory tract microbiota profiling experiments. Thermo Fisher Scientific. URL: [https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0017952\\_RespiratoryTractMicrobiotaProfiling\\_OA\\_AG.pdf](https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0017952_RespiratoryTractMicrobiotaProfiling_OA_AG.pdf) [accessed 2022-01-26]
32. Influenza. King County. URL: <https://www.kingcounty.gov/depts/health/communicable-diseases/disease-control/influenza.aspx> [accessed 2022-01-26]
33. de Jong MD, Simmons CP, Thanh TT, Hien VM, Smith GJ, Chau TN, et al. Fatal outcome of human influenza A (H5N1) is associated with high viral load and hypercytokinemia. *Nat Med* 2006 Oct;12(10):1203-1207 [FREE Full text] [doi: [10.1038/nm1477](https://doi.org/10.1038/nm1477)] [Medline: [16964257](https://pubmed.ncbi.nlm.nih.gov/16964257/)]



34. Bagga B, Woods CW, Veldman TH, Gilbert A, Mann A, Balaratnam G, et al. Comparing influenza and RSV viral and disease dynamics in experimentally infected adults predicts clinical effectiveness of RSV antivirals. *Antivir Ther* 2013;18(6):785-791. [doi: [10.3851/IMP2629](https://doi.org/10.3851/IMP2629)] [Medline: [23714753](https://pubmed.ncbi.nlm.nih.gov/23714753/)]
35. Cheng CK, Cowling BJ, Chan KH, Fang VJ, Seto WH, Yung R, et al. Factors affecting QuickVue Influenza A + B rapid test performance in the community setting. *Diagn Microbiol Infect Dis* 2009 Sep;65(1):35-41. [doi: [10.1016/j.diagmicrobio.2009.05.003](https://doi.org/10.1016/j.diagmicrobio.2009.05.003)] [Medline: [19679233](https://pubmed.ncbi.nlm.nih.gov/19679233/)]
36. Elliot AJ, Bermingham A, Charlett A, Lackenby A, Ellis J, Sadler C, et al. Self-sampling for community respiratory illness: a new tool for national virological surveillance. *Euro Surveill* 2015 Mar 12;20(10):21058. [doi: [10.2807/1560-7917.es2015.20.10.21058](https://doi.org/10.2807/1560-7917.es2015.20.10.21058)] [Medline: [25788252](https://pubmed.ncbi.nlm.nih.gov/25788252/)]
37. Thompson MG, Ferber JR, Odouli R, David D, Shifflett P, Meece JK, et al. Results of a pilot study using self-collected mid-turbinate nasal swabs for detection of influenza virus infection among pregnant women. *Influenza Other Respir Viruses* 2015 May;9(3):155-160 [FREE Full text] [doi: [10.1111/irv.12309](https://doi.org/10.1111/irv.12309)] [Medline: [25711978](https://pubmed.ncbi.nlm.nih.gov/25711978/)]
38. Iovine NM, Ison T, Payton T, Morris JG, Rand K. High percentage of false-positive results for influenza B obtained with a rapid influenza point-of-care test. *Clin Infect Dis* 2014 Aug 15;59(4):604-605. [doi: [10.1093/cid/ciu315](https://doi.org/10.1093/cid/ciu315)] [Medline: [24795330](https://pubmed.ncbi.nlm.nih.gov/24795330/)]
39. Nivin B, Zancocchio T. Cluster of false-positive influenza B virus rapid antigen test results in a New York City hospital. *J Clin Microbiol* 2012 Sep;50(9):3141 [FREE Full text] [doi: [10.1128/JCM.01452-12](https://doi.org/10.1128/JCM.01452-12)] [Medline: [22718940](https://pubmed.ncbi.nlm.nih.gov/22718940/)]
40. Otto MP, Toyer A, Poggi C, Janvier F. *Diagn Microbiol Infect Dis* 2019 Aug;94(4):342-343. [doi: [10.1016/j.diagmicrobio.2019.03.002](https://doi.org/10.1016/j.diagmicrobio.2019.03.002)] [Medline: [30971360](https://pubmed.ncbi.nlm.nih.gov/30971360/)]
41. Leeftang MM, Rutjes AW, Reitsma JB, Hooft L, Bossuyt PM. Variation of a test's sensitivity and specificity with disease prevalence. *CMAJ* 2013 Aug 06;185(11):E537-E544 [FREE Full text] [doi: [10.1503/cmaj.121286](https://doi.org/10.1503/cmaj.121286)] [Medline: [23798453](https://pubmed.ncbi.nlm.nih.gov/23798453/)]
42. Grijalva CG, Poehling KA, Edwards KM, Weinberg GA, Staat MA, Iwane MK, et al. Accuracy and interpretation of rapid influenza tests in children. *Pediatrics* 2007 Jan;119(1):e6-11. [doi: [10.1542/peds.2006-1694](https://doi.org/10.1542/peds.2006-1694)] [Medline: [17200259](https://pubmed.ncbi.nlm.nih.gov/17200259/)]
43. Sanner K, Wikström I, Strand A, Lindell M, Wilander E. Self-sampling of the vaginal fluid at home combined with high-risk HPV testing. *Br J Cancer* 2009 Sep 01;101(5):871-874 [FREE Full text] [doi: [10.1038/sj.bjc.6605194](https://doi.org/10.1038/sj.bjc.6605194)] [Medline: [19654577](https://pubmed.ncbi.nlm.nih.gov/19654577/)]
44. Marinsek N, Shapiro A, Clay I, Bradshaw B, Ramirez E, Min J. Measuring COVID-19 and influenza in the real world via person-generated health data. medRxiv. URL: <https://www.medrxiv.org/content/10.1101/2020.05.28.20115964v1> [accessed 2022-01-26]
45. Smolinski MS, Crawley AW, Baltrusaitis K, Chunara R, Olsen JM, Wójcik O, et al. Flu near you: crowdsourced symptom reporting spanning 2 influenza seasons. *Am J Public Health* 2015 Oct;105(10):2124-2130. [doi: [10.2105/AJPH.2015.302696](https://doi.org/10.2105/AJPH.2015.302696)] [Medline: [26270299](https://pubmed.ncbi.nlm.nih.gov/26270299/)]
46. McCulloch DJ, Jackson ML, Hughes JP, Lester S, Mills L, Freeman B, et al. Seroprevalence of SARS-CoV-2 antibodies in Seattle, Washington: October 2019-April 2020. *PLoS One* 2021 May 27;16(5):e0252235 [FREE Full text] [doi: [10.1371/journal.pone.0252235](https://doi.org/10.1371/journal.pone.0252235)] [Medline: [34043706](https://pubmed.ncbi.nlm.nih.gov/34043706/)]
47. Lyon V. Feasibility and acceptance of mobile point-of-care self testing: a case study on the Flu@home mobile app. In: *Proceedings of the AMIA Clinical Informatics Conference. 2020 Presented at: AMIA Clinical Informatics Conference; May 19-21, 2020; Seattle, Washington* URL: <https://d3w3ioujxcalzn.cloudfront.net>

## Abbreviations

- ARI:** acute respiratory illness
- Crt:** relative cycle threshold
- EHFT:** Ellume Home Flu Test
- FN:** false negative
- FP:** false positive
- PCR:** polymerase chain reaction
- POC:** point-of-care
- PPV:** positive predictive value
- RDT:** rapid diagnostic test
- REDCap:** Research Electronic Data Capture
- RNase P:** ribonuclease P
- qRT-PCR:** quantitative reverse-transcription polymerase chain reaction
- RSV:** respiratory syncytial virus
- SFS:** Seattle Flu Study
- STARD:** Standards for Reporting of Diagnostic Accuracy Studies
- TP:** true positive
- VTM:** viral transport medium

*Edited by T Sanchez; submitted 05.03.21; peer-reviewed by Y Man, H Mehdizadeh, J Seitz; comments to author 19.10.21; revised version received 02.11.21; accepted 30.11.21; published 22.02.22.*

*Please cite as:*

*Geyer RE, Kotnik JH, Lyon V, Brandstetter E, Zigman Suchsland M, Han PD, Graham C, Ilcisin M, Kim AE, Chu HY, Nickerson DA, Starita LM, Bedford T, Lutz B, Thompson MJ*

*Diagnostic Accuracy of an At-Home, Rapid Self-test for Influenza: Prospective Comparative Accuracy Study*

*JMIR Public Health Surveill* 2022;8(2):e28268

URL: <https://publichealth.jmir.org/2022/2/e28268>

doi: [10.2196/28268](https://doi.org/10.2196/28268)

PMID: [35191852](https://pubmed.ncbi.nlm.nih.gov/35191852/)

©Rachel E Geyer, Jack Henry Kotnik, Victoria Lyon, Elisabeth Brandstetter, Monica Zigman Suchsland, Peter D Han, Chelsey Graham, Misja Ilcisin, Ashley E Kim, Helen Y Chu, Deborah A Nickerson, Lea M Starita, Trevor Bedford, Barry Lutz, Matthew J Thompson. Originally published in JMIR Public Health and Surveillance (<https://publichealth.jmir.org>), 22.02.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

# An Evaluation of the Text Illness Monitoring (TIM) Platform for COVID-19: Cross-sectional Online Survey of Public Health Users

Heather A Joseph<sup>1</sup>, MPH; Susan Z Ingber<sup>2</sup>, MSPP; Chelsea Austin<sup>3</sup>, MPH; Caroline Westnedge<sup>4</sup>, MPH; F V Strona<sup>5</sup>, MPH; Leslie Lee<sup>6</sup>, MPH; Ami B Shah<sup>4</sup>, MPH; Lauren Roper<sup>6</sup>, MPH; Anita Patel<sup>6</sup>, PharmD, MS

<sup>1</sup>Division of Environmental Health Science and Practice, National Center for Environmental Health, Centers for Disease Control and Prevention, Chamblee, GA, United States

<sup>2</sup>Division of Viral Hepatitis, National Center for HIV, Viral Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention, Atlanta, GA, United States

<sup>3</sup>Department of Health Policy and Behavioral Sciences, School of Public Health, Georgia State University, Atlanta, GA, United States

<sup>4</sup>Federal Civilian Division, General Dynamics Information Technology, Atlanta, GA, United States

<sup>5</sup>Division of STD Prevention, National Center for HIV, Viral Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention, Atlanta, GA, United States

<sup>6</sup>Office of the Director, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, GA, United States

**Corresponding Author:**

Heather A Joseph, MPH

Division of Environmental Health Science and Practice

National Center for Environmental Health

Centers for Disease Control and Prevention

Mailstop 106

4770 Buford Hwy

Chamblee, GA, 30341

United States

Phone: 1 4046392636

Email: [hbj7@cdc.gov](mailto:hbj7@cdc.gov)

## Abstract

**Background:** The US public health response to the COVID-19 pandemic has required contact tracing and symptom monitoring at an unprecedented scale. The US Centers for Disease Control and Prevention and several partners created the Text Illness Monitoring (TIM) platform in 2015 to assist US public health jurisdictions with symptom monitoring for potential novel influenza virus outbreaks. Since May 2020, 142 federal, state, and local public health agencies have deployed TIM for COVID-19 symptom monitoring.

**Objective:** The aim of this study was to evaluate the utility, benefits, and challenges of TIM to help guide decision-making for improvements and expansion to support future public health emergency response efforts.

**Methods:** We conducted a brief online survey of previous and current TIM administrative users (admin users) from November 28 through December 21, 2020. Closed- and open-ended questions inquired about the onboarding process, decision to use TIM, groups monitored with TIM, comparison of TIM to other symptom monitoring systems, technical challenges and satisfaction with TIM, and user support. A total of 1479 admin users were invited to participate.

**Results:** A total of 97 admin users from 43 agencies responded to the survey. Most admin users represented the Indian Health Service (35/97, 36%), state health departments (26/97, 27%), and local or county health departments (18/97, 19%), and almost all were current users of TIM (85/94, 90%). Among the 43 agencies represented, 11 (26%) used TIM for monitoring staff exclusively, 13 (30%) monitored community members exclusively, and 19 (44%) monitored both staff and community members. Agencies most frequently used TIM to monitor symptom development in contacts of cases among community members (28/43, 65%), followed by symptom development among staff (27/43, 63%) and among staff contacts of cases (24/43, 56%). Agencies also reported using TIM to monitor patients with COVID-19 for the worsening of symptoms among staff (21/43, 49%) and community members (18/43, 42%). When asked to compare TIM to previous monitoring systems, 78% (40/51) of respondents rated TIM more favorably than their previous monitoring system, 20% (10/51) said there was no difference, and 2% (1/51) rated the previous monitoring system more favorably than TIM. Most respondents found TIM favorable in terms of time burden, staff

burden, timeliness of the data, and the ability to monitor large population sizes. TIM compared negatively to other systems in terms of effort to enroll participants (ie, persons TIM monitors) and accuracy of the data. Most respondents (76/85, 89%) reported that they would highly or somewhat recommend TIM to others for symptom monitoring.

**Conclusions:** This evaluation of TIM showed that agencies used TIM for a variety of purposes and rated TIM favorably compared to previously used monitoring systems. We also identified opportunities to improve TIM; for example, enhancing the flexibility of alert deliveries would better meet admin users' varying needs. We also suggest continuous program evaluation practices to assess and respond to implementation gaps.

(*JMIR Public Health Surveill* 2022;8(2):e32680) doi:[10.2196/32680](https://doi.org/10.2196/32680)

## KEYWORDS

COVID-19; contact tracing; SMS text system; symptom monitoring

## Introduction

Monitoring exposed individuals during a public health crisis, such as the COVID-19 pandemic, is critical for implementation of an effective public health response. Ongoing symptom monitoring conducted by clinical providers and public health officials has traditionally been done via telephone calls or in-person screening, both time-consuming processes requiring extensive health department resources. Text-based communication is increasingly used for public health interventions [1-3]. Two-way SMS text messaging can scale up public health agencies' ability to monitor on a predetermined schedule (ie, daily or weekly) or conduct a one-time follow-up with individuals.

The US Centers for Disease Control and Prevention (CDC) has recommended the use of technology to enhance partner services for sexually transmitted infections since 2008 [4]. For example, CDC's Toolkit for Technology-Based Partnership Services provides guidance on using text messaging and mobile apps to help providers initiate appointment setup for partners of infected patients, as well as patient check-in and monitoring medication compliance [5]. Text platforms have also been used for monitoring postpartum depression and mental health [6,7], symptom monitoring for infectious disease contact tracing during Ebola [8], and symptom monitoring and prophylaxis medication adherence for avian influenza and influenza-like illness [9,10]. In sum, the use of text-based monitoring systems continues to grow, and preliminary reports are encouraging regarding their usability, acceptability, and effectiveness in an acute infectious disease outbreak. However, because the aims, populations monitored, and platforms vary greatly, it is critical to conduct evaluation to improve delivery and build the evidence base for emergent needs and systems.

In 2020, CDC recommended that those potentially exposed to SARS-CoV-2 be monitored for 14 days [11]. This posed a burden to public health systems, which could have led to substantial under monitoring. In the context of the COVID-19 pandemic, the World Health Organization has encouraged the use of electronic data capture tools to support efficient contact tracing and active surveillance of close contacts on a large scale [12]. Though evidence for the effectiveness of text-based active surveillance or monitoring systems among community-based contacts of cases of COVID-19 has been limited to date, several reports indicate some promise [13-16].

The Text Illness Monitoring (TIM) platform was developed in 2015, through a collaboration between CDC, the National Association of County and City Health Officials, and Compliant Campaign (a third-party contractor) to assist US jurisdictions with monitoring individuals at potential risk for novel influenza virus infection. In 2016, the Michigan Department of Health and Human Services asked CDC to pilot-test TIM during a swine flu outbreak at nine county fairs [17]. The pilot evaluated TIM's ability to enhance detections of H3N2v virus infections among household members of symptomatic fair attendees and its feasibility and acceptability for use in future outbreak investigations of novel influenza viruses or similar threats. Among an estimated 500 households with a member who exhibited symptoms, representatives of 87 (17.4%) households were enrolled in TIM. Ultimately, the system detected two H3N2v virus infections among the enrolled household members, and 80% of survey respondents indicated they would participate in another TIM-based monitoring event.

Early in the COVID-19 pandemic, there was no existing system that was free, rapidly available, and easy to scale up for symptom monitoring of large, diverse populations. At the time, public health entities, like state and local health departments, were overwhelmed with trying to identify broader control actions, conduct surveillance, organize laboratory activities, and prepare the nation for the response to the pandemic. As a result, CDC reconfigured TIM to meet these pressing challenges of symptom monitoring for large-scale contact tracing and employee monitoring, while promoting and providing technical support for TIM to domestic public health agencies at no cost. CDC began to use TIM to monitor deployed staff in February 2020, as a pilot. Shortly thereafter, multiple federal agencies began using TIM to monitor symptoms among employees. State, local, and tribal public health authorities quickly followed by using TIM to monitor staff and community members for development and worsening of symptoms. Use of TIM was voluntary for federal agencies and public health authorities. Over the spring and summer of 2020, CDC continued to expand monitoring of field-deployed and remotely deployed responders; CDC staff administering TIM provided regular and detailed feedback to the development team to support enhancements.

In response to heightened interest in TIM, ad hoc user feedback, and suggestions for improving the system, CDC initiated an evaluation of the TIM system in accordance with CDC's Framework for Program Evaluation to facilitate selection and prioritization of system enhancements [18]. The evaluation team

convened internal stakeholders who established the following evaluation questions: Who are the administrative users (admin users)? What are the most important reasons for adopting TIM? How is TIM being used by public health agencies? How does TIM compare to other monitoring systems? To what extent has TIM contributed to earlier COVID-19 identification? What are the challenges of using TIM and how can these be addressed? How satisfied are admin users with TIM and CDC support?

For the purpose of better understanding the needs of admin users, we also wanted to discern if there were differences in these outcomes by the type of agency or the populations being monitored.

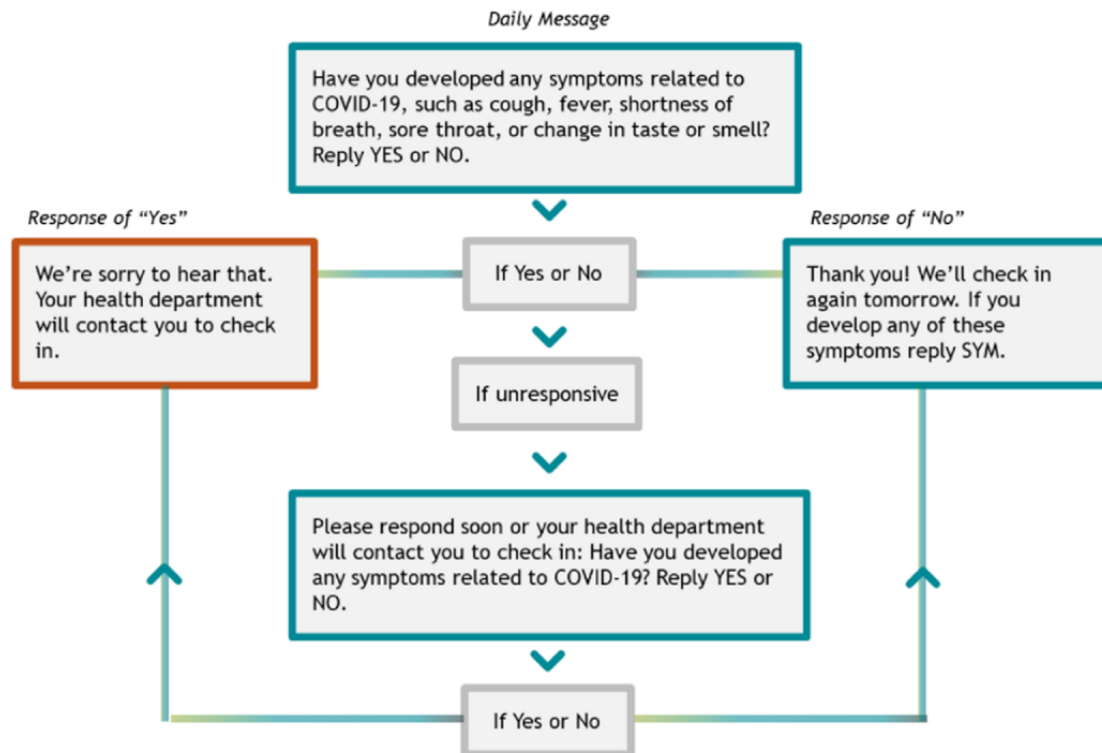
## Methods

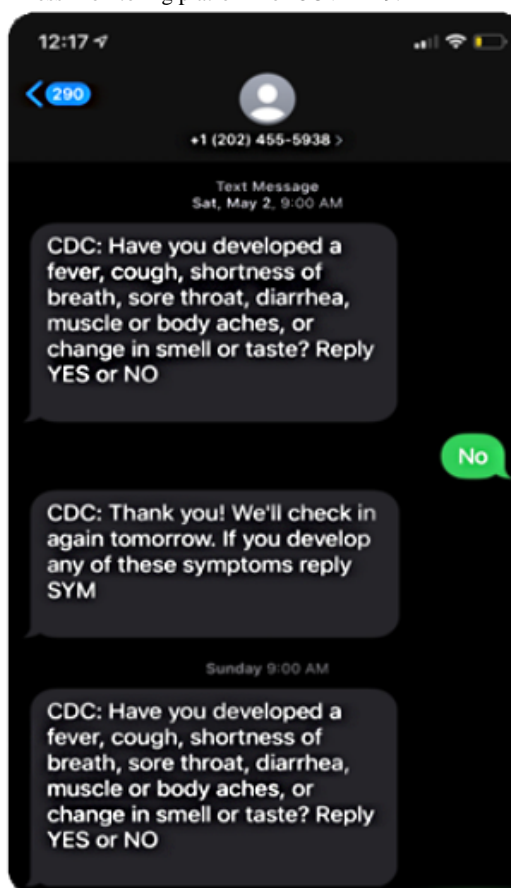
### Text Illness Monitoring Platform

TIM was designed to be a simple, low-resource tool to implement. In the context of the current use for the COVID-19 response in the United States, the CDC TIM team enrolls public health officials (ie, “admin users”), who can then create “campaigns”—text messaging workflows in which participants are enrolled for monitoring—for their jurisdictions. A typical campaign workflow for the daily message is shown in Figure 1.

Persons whom TIM monitors for symptoms (ie, “participants”) receive two or more text messages each day for the monitoring period designated by the admin user. As shown in Figure 2, texts to participants include questions about whether they have symptoms consistent with COVID-19 and require a “YES” or “NO” response. TIM generates an alert when a participant responds “YES” or “SYM” (ie, symptom). TIM can read other response variations (ie, “Y,” “Yea,” or “symptom”) that confirm experiencing symptoms and generates a symptom alert. If a participant does not respond to the first message by a preset time, TIM sends a follow-up reminder. If the participant does not respond by a preset time to the reminder text, TIM will create a nonresponse alert. A dashboard feature allows TIM admin users to log in, track, and follow up with participants who report symptoms or fail to respond. Admin users can view alerts outside of the dashboard if they opt to receive email notifications for alerts. Participants can opt out of TIM at any time by texting “QUIT,” “END,” “CANCEL,” “UNSUBSCRIBE,” “REMOVE,” “OPT OUT,” “OPTOUT,” or “STOP.” Admin users can also stop messages at any time or can set a specific number of days for the monitoring period via the dashboard. If a participant sends an unexpected response to TIM, the system sends a standard response to prompt the participant to send an appropriate response (eg, “You are enrolled in the text symptom monitoring program. If you develop symptoms, reply SYM.”).

Figure 1. Daily message from a typical Text Illness Monitoring campaign for COVID-19.



**Figure 2.** Example text exchange on the Text Illness Monitoring platform for COVID-19.

Reflecting the capability for local tailoring, admin users can customize the technical attributes of their TIM campaigns (eg, number of monitoring days, language selection, and time limit between initial and reminder messages). TIM allows for multiple admin users with tiers of access to control interactions for various user levels that were specifically designed to manage state-level versus local-level views for the system.

Between May 1 and December 29, 2020, a total of 682 TIM campaigns were established. This included 1479 staff from 142 public health agencies in 20 states who were designated as TIM admin users, and 97,184 individuals across the United States who were enrolled as TIM participants. The daily average number of persons monitored was 6838. For the week of November 28 to December 4, 2020 (ie, the first week of data collection for this evaluation), the weekly average of participants reporting symptoms was 3.8%, and the weekly average nonresponse to alerts was 17.4%.

### Survey Design

To validate the instrument, we piloted survey questions through seven semistructured, open-ended, 60-minute phone interviews with a sample of TIM admin users. Eligible respondents for the validation phase and the online survey were TIM admin users associated with a current or formerly active campaign established at least 2 weeks prior to the interview or survey completion. We conducted phone interviews with 7 admin users who had a wide range of experiences with TIM, based on administrative data from the TIM Support Team (see [Multimedia Appendix 1](#) for interview instrument). We incorporated insights

from the validation stage into the survey ([Multimedia Appendix 2](#)) that was deployed to admin users via REDCap (Research Electronic Data Capture). REDCap is a secure, Health Insurance Portability and Accountability Act (HIPAA)-compliant, web-based application developed by Vanderbilt University and used by CDC to capture research data and create databases.

### Data Collection

The TIM Support Team emailed a survey invitation to all admin users (n=1479). The anonymous survey took approximately 15 minutes to complete and, as indicated on the introduction page, multiple admin users per agency were eligible to respond. No incentives for survey completion were provided. Respondents indicated their state and agency name. The survey invited respondents to answer questions, all optional, about the following topics: onboarding process, decision to use TIM, use of TIM-generated data and reports, technical challenges of using TIM, utility of TIM, and satisfaction with TIM and user support. The survey also included two questions that prompted respondents to report the following for their agency: populations monitored and uses of TIM. A respondent may have worked for an agency with multiple campaigns; we sought to collect this information about all of the campaigns for a particular agency. Survey question types included “check all that apply,” yes-or-no, multiple-choice, Likert-scale, and open-ended questions to collect both quantitative and qualitative feedback. For example, we asked respondents to compare TIM to the most recent system used prior to TIM through a series of questions about cost, staffing requirements, staff hours, burden to enroll participants, number of participants that could be monitored,

data accuracy, data completeness, and data timeliness. Participants rated these on a 5-point scale with a neutral response option in the middle (ie, “the same”). We categorized “better” and “somewhat better” responses as “favorable,” and we categorized “somewhat worse” and “much worse” responses as “unfavorable.” The data collection period ran from November 28 to December 21, 2020. A copy of the survey questions is presented as [Multimedia Appendix 2](#).

### Analysis

The quantitative data were analyzed in SAS 9.4 (SAS Institute Inc) using standard analysis techniques, including basic descriptive statistics. The primary unit of analysis is the respondent (ie, admin user). For the analysis on the use of TIM, we report the agency as the unit of analysis because the use of TIM is an agency-level decision. We also sought to discern if there were differences by type of agency and populations monitored, as this information could be used to refine the application to better meet admin users’ needs. Due to low sample sizes, crude frequencies and the Fisher exact test were used to determine statistical differences in terms of agency type or population monitored. We report the results of this statistical test when differences were statistically significant ( $P < .05$ ). Since all questions were optional, there were missing data. Consequently, denominators vary by question.

Qualitative data collected via open-ended questions were input into a separate data set for coding. Two members of the research team used an inductive approach to develop an initial set of codes on a subsample of the extracted qualitative data (ie, approximately 10% of all extracted content) [19]. Both coders then independently coded samples of data and met twice weekly to reconcile coded content and update their code list and definitions [20]. This process was repeated until intercoder agreement reached 84% [19,21]. From there, a single coder finished coding the remaining qualitative data (ie, approximately one-third of the total sample). Both coders then performed content analysis to determine those themes that emerged most prominently [22]. To enable a better understanding of theme salience, coders quantified the frequency of theme mentions across the data set, by number of survey respondents and by number of unique agencies. We also assessed whether patterns may vary by agency type.

### Compliance

CDC determined that the data collection was nonresearch; thus, no human subjects review was conducted in accordance with applicable federal law and CDC policy. The Paperwork Reduction Act applied, but a public health emergency waiver was obtained.

Regarding informed consent, the survey home page included introductory information about the reason for the data collection, voluntary participation, nonidentifiable reporting of the findings, length of time for survey completion, and whom to contact about TIM or the survey. There was no documentation of consent.

## Results

### Response Rate

During the time of data collection from November 28 through December 21, 2020, 67 agencies in 18 states were actively using TIM. A total of 180 monitoring campaigns were in progress, and 10,414 participants were being monitored.

Of 1479 admin users contacted to respond to the survey, 100 (6.8%) responded. Out of those 100 respondents, 2 (2.0%) were ineligible because they had not established a campaign. We also found 1 (1.0%) incomplete duplicate of a complete survey. These 3 (3.0%) records were dropped, resulting in an analytic sample of 97 respondents, representing 43 distinct agencies. Out of 97 respondents, 11 (11%) did not specify their agency name. Specifically, 2 respondents from Florida were assigned to a single “Florida\_Unspecified” group and 9 Indian Health Service (IHS) respondents were assigned to a single “IHS\_Unspecified” group.

The response rate for individual admin users was low (100/1479, 6.8%), which could have been the result of many users never accessing or using the system or due to the frequent turnover of public health staff at the local level. We could not retrospectively identify “active admin users” at a given time through the administrative data provided by the platform developer. However, we were able to determine an agency-level response rate, which was higher. Staff from 30.3% (43/142) of agencies that had ever used TIM responded, while staff from 64% (43/67) of agencies that were actively using TIM at the time of the data collection responded to the survey.

### Respondent Characteristics

Respondents represented a diverse sample of public health agencies (Table 1). Almost all respondents were current admin users of TIM, and slightly over one-third reported using the system for 3 months or less. The mean number of months using TIM was 4.5 (SD 2.55), with a range of less than 1 month to 10 months. Though only 39% (37/94) of respondents indicated that they were primary or secondary points of contact for TIM within their agency, most were responsible for managing one or more features of TIM (eg, administration of campaigns, participants, alerts, user features, and data and reporting features).

**Table 1.** Characteristics of respondents of the online survey regarding the Text Illness Monitoring (TIM) platform for COVID-19, November to December 2020.

Characteristic	Respondents (N=97), n (%)
<b>Agency type<sup>a</sup> (N=97)<sup>b</sup></b>	
Indian Health Service	35 (36)
State	26 (27)
Local or county	18 (19)
Federal (other than Indian Health Service)	12 (12)
Tribal nation	6 (6)
Currently using TIM (n=94)	85 (90)
<b>Number of months using TIM (n=93)</b>	
<1-3	36 (39)
4-6	34 (37)
≥7	23 (25)
<b>Role (n=94)</b>	
Primary or secondary point of contact for TIM	37 (39)
Responsible for managing one or more features of TIM	89 (95)

<sup>a</sup>To simplify subsequent analysis for “agency type,” state and local agencies were combined into one category, and tribal nation agencies and the Indian Health Service were combined into another.

<sup>b</sup>Denominators vary because of nonresponse, in the form of missing responses and responses of “I don’t know.”

### Reasons for Adopting TIM

As shown in Table 2, respondents indicated that the most common reasons for adoption were the ability to monitor large populations and that TIM is a better alternative to phone

screening. Compared to respondents from state and local agencies, those from the IHS and tribal nation agencies more frequently indicated that the rationale for adoption was that TIM was a better alternative to in-person screening (20/40, 50%, vs 9/40, 23%;  $P=.02$ ).



**Table 2.** Key survey results regarding evaluation of the Text Illness Monitoring (TIM) platform for COVID-19, November to December 2020.

Survey item and responses	Value, n (%)
<b>Respondent level<sup>a</sup></b>	
<b>Reasons for adopting TIM (n=90)<sup>b</sup></b>	
Could monitor large numbers of people	53 (59)
Better alternative to screening via phone	42 (47)
Better alternative to screening in person	34 (38)
No cost	27 (30)
Created and maintained by US Centers for Disease Control and Prevention	19 (21)
<b>Systems previous to TIM</b>	
Used a system before TIM (n=90)	64 (71)
Used in-person screening (n=63)	31 (49)
Used other contact tracing software or symptom monitoring application (n=63)	11 (17)
<b>Identification of COVID-19 symptoms (n=87) and confirmed cases (n=86)</b>	
TIM identified symptomatic participants (yes)	70 (80)
TIM identified confirmed cases (yes)	44 (51)
<b>Reported that TIM identified symptomatic participants in a timely way (n=69)</b>	
Somewhat	25 (36)
A lot	30 (43)
Very much so	14 (20)
<b>Satisfied or very satisfied with the TIM Technical Support Team (n=26)</b>	
Timeliness	25 (96)
Communication	24 (92)
Extent to which issues were resolved	22 (85)
<b>Would recommend TIM for symptom activity monitoring (n=85)</b>	
Highly or somewhat recommend	76 (89)
Neither recommend nor discourage	6 (7)
Highly or somewhat discourage	3 (4)
<b>Other tools used alongside TIM (n=92)</b>	
Spreadsheets	64 (70)
Agency or personal phones	43 (47)
In-person screening	30 (33)
Pen and paper	29 (32)
Contact tracing software	19 (21)
Data analysis software	5 (5)
None	5 (5)
<b>Agency level<sup>c</sup></b>	
<b>Populations monitored (n=43)</b>	
Monitoring staff only	11 (26)
Monitoring community members only	13 (30)
Monitoring both staff and community members	19 (44)
<b>Uses of TIM (n=43)</b>	
Among staff, monitoring for development of symptoms	27 (63)

Survey item and responses	Value, n (%)
Among staff, monitoring cases for worsening of symptoms	21 (49)
Among staff, monitoring contacts of cases for development of symptoms	24 (56)
Among community members, monitoring cases for development of symptoms	20 (47)
Among community members, monitoring cases for worsening of symptoms	18 (42)
Among community members, monitoring contacts of cases for development of symptoms	28 (65)

<sup>a</sup>Questions were asked in a way that indicated respondents could answer for themselves or their agency.

<sup>b</sup>Denominators vary because of nonresponse, in the form of missing responses and responses of “I don’t know.”

<sup>c</sup>Questions were asked in a way that indicated respondents should answer for their agency.

### Uses of TIM

We report uses of TIM by agency, rather than by individual respondent (Table 2). We found a mix of uses and populations monitored. In terms of populations monitored, agencies reported using TIM to exclusively monitor illness in staff, community members, or both groups.

Agencies most frequently used TIM to monitor for symptom development in contacts of cases among community members, followed by symptom development among staff and among staff contacts of cases. Agencies also reported using TIM to monitor COVID-19 cases for the worsening of symptoms among staff and community members. Another reported use was for monitoring symptom development in contacts of cases among community members.

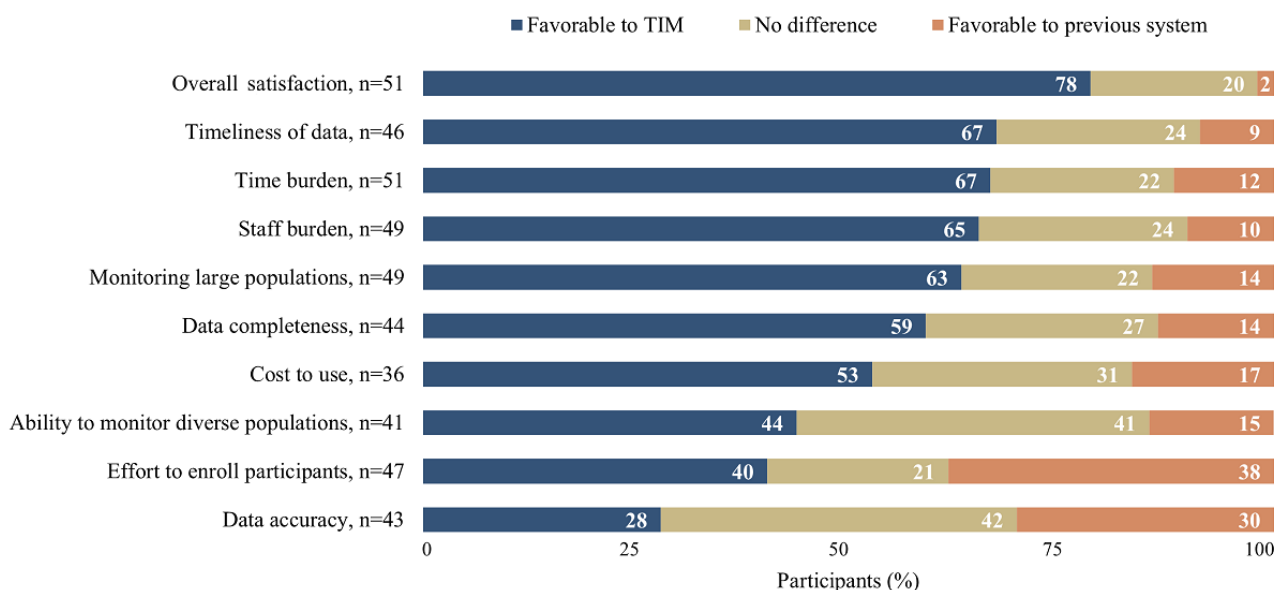
Federal agencies (3/3, 100%) and IHS and tribal nation agencies (12/19, 63%) most frequently used TIM to monitor for development of symptoms among staff, while state and local agencies (18/21, 86%) most frequently used TIM to monitor community contacts of cases for development of symptoms.

In conjunction with using TIM for symptom monitoring, agencies reported using spreadsheets, agency or personal phones, in-person screening, pen and paper, contact tracing software, data analysis software, and no other tools.

### Comparison of TIM to Previously Used Monitoring Systems

Most respondents reported using another monitoring system before TIM. These systems included in-person screening and other contact tracing software or symptom monitoring applications (Table 2). As shown in Figure 3, among those who had used a previous system, most rated TIM favorably overall compared to the previous system they used; 11 respondents indicated “I don’t know” to all questions. TIM compared most favorably to the previous system in terms of the timeliness of the data, time burden, staff burden, and the ability to monitor large population sizes. Compared to TIM, respondents indicated that their previous systems required less effort to enroll participants and yielded more accurate data.

**Figure 3.** Evaluation of the Text Illness Monitoring (TIM) platform for COVID-19, comparison of TIM to previous monitoring system, November to December 2020.



Compared to IHS and tribal nation agencies, state and local agencies were more likely to consider TIM to be favorable for monitoring larger populations (19/23, 83%, vs 10/23, 43%;  $P=.01$ ).

### Identification of COVID-19 Symptoms and Confirmed Cases

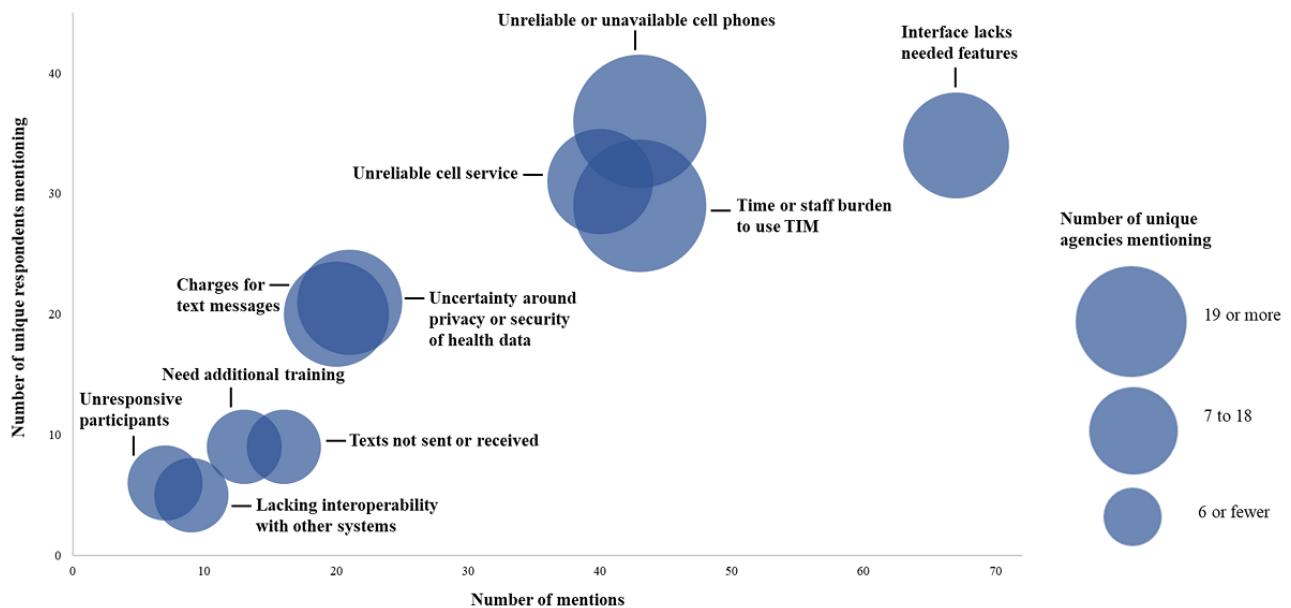
As shown in Table 2, most respondents reported that TIM identified participants who developed symptoms of COVID-19

and participants who were later confirmed to be cases. Respondents indicated that TIM identified symptomatic participants in a timely way either “very much so,” “a lot,” or “somewhat.”

### Challenges of Using TIM

Respondents were asked a series of open-ended questions about technical challenges and concerns regarding the use of TIM.

**Figure 4.** Evaluation of the Text Illness Monitoring (TIM) platform for COVID-19, reported challenges using TIM, November to December 2020. Content analysis themes were mapped according to their number of mentions (x-axis), the number of unique respondents (y-axis), and the number of agencies reporting the challenge (bubble size). Thus, larger bubbles placed further along the x-axis and y-axis indicate more prevalent challenges.



The most requested feature enhancements were the ability to customize campaign time zones, to receive symptom alerts by email or text message, and to delete or move participants from campaigns at any time. Participants reported time or staff burden accruing through efforts to enroll large numbers of participants and difficulty sorting or navigating through many pages of data generated by nonresponse alerts. Of the unique agencies reporting “unreliable cell service” as a concern, a little over half (10/17, 59%) were IHS or tribal nation affiliated. Additionally, of the unique agencies reporting “unreliable or unavailable cell phones,” well over half (12/19, 63%) were IHS or tribal nation affiliated.

Some themes co-occurred across the data set, suggesting underlying relationships among reported challenges. Slightly less than half (19/43, 44%) of the “time or staff burden” mentions co-occurred with “interface lacks needed features” mentions. Additionally, nearly half (7/16, 44%) of the “texts not received or sent” mentions co-occurred with other themes, namely unreliable cell service (4/7, 57%) and unreliable or unavailable cell phones (2/7, 29%).

### Satisfaction With CDC Support and Overall Experience

CDC provided onboarding support that included sending welcome emails to introduce TIM, conducting TIM orientations and demonstrations, and providing the TIM user guide and frequently asked questions (FAQs). CDC also staffed a help

desk to provide daily support for admin users after they established their campaigns. Among those who submitted technical support requests to the TIM Support Team, most respondents indicated that they were satisfied with the timeliness of the TIM Support Team’s responses to their tech support requests (Table 2).

The survey also collected data about whether respondents would recommend TIM to others for symptom monitoring. Most respondents (76/85, 89%) reported that they would highly or somewhat recommend TIM, while a few were neutral (6/85, 7%) or would discourage others from using TIM (3/85, 4%). This did not vary by agency type or population monitored.

The survey also collected data about whether respondents would recommend TIM to others for symptom monitoring. Most respondents (76/85, 89%) reported that they would highly or somewhat recommend TIM, while a few were neutral (6/85, 7%) or would discourage others from using TIM (3/85, 4%). This did not vary by agency type or population monitored.

The survey also collected data about whether respondents would recommend TIM to others for symptom monitoring. Most respondents (76/85, 89%) reported that they would highly or somewhat recommend TIM, while a few were neutral (6/85, 7%) or would discourage others from using TIM (3/85, 4%). This did not vary by agency type or population monitored.

## Discussion

### Principal Findings

Evaluation of emergency public health response activities can provide timely and actionable insights about successes and areas for improvement to responders and decision makers. In the context of the COVID-19 response, CDC has encouraged state and local health departments to implement and use evaluation findings on topics ranging from mask wearing to COVID-19 mitigation strategies in schools. With this same intention, this evaluation of TIM can help support its expanded use for the current COVID-19 pandemic and provide guidance for TIM use during future public health emergencies. The evaluation also provides evidence for prioritizing specific system

enhancements to better support TIM admin users with implementation.

In sum, admin users who responded to the survey represented the IHS, state health departments, and local or county health departments. Agencies represented by survey respondents used TIM to exclusively monitor staff, monitor community members, or both. Admin users' agencies most frequently used TIM to monitor symptom development in contacts of cases among community members, followed by symptom development among staff and among staff contacts of cases. Agencies also used TIM to monitor patients with COVID-19 for the worsening of symptoms among staff and community members. Most respondents rated TIM more favorably compared to their previous monitoring system. Most respondents found TIM favorable in terms of time burden, staff burden, timeliness of the data, and the ability to monitor large population sizes. TIM compared negatively to other systems in terms of effort to enroll participants and accuracy of the data. Most would highly or somewhat recommend TIM to others for symptom monitoring.

TIM is intended to aid in the timely detection, treatment, and prevention of transmission of viruses with pandemic potential, such as COVID-19. Passive surveillance methods can miss infections, while other active surveillance strategies, like in-person or phone screening, can be very time- and resource-intensive. This evaluation supports the results of the initial study of TIM's use during a swine flu outbreak at agricultural fairs [17]. In that use case, TIM successfully identified two cases among the 392 individuals monitored for illness over a 4-week period. Stewart et al [17] reported that two types of text messages were sent: one using formal language and another using informal language. The informal version was associated with more staff follow-up time due to false alerts and unrecognized text responses. The version of TIM deployed for the COVID-19 response provided admin users with the flexibility to customize the level of formality for their TIM campaigns, but also defaulted the system to use more formal and straightforward language to avoid similar monitoring challenges as those experienced during the swine flu outbreak. No survey respondents in this evaluation indicated that participants had difficulty understanding the expected text response.

SMS text-based systems like TIM are important tools for large outbreak investigations that require significant public health resources, as they are not only scalable but also cost-effective. These systems are easier to use than mobile apps, which may require downloading, favor more technologically savvy admin users, and can invoke privacy concerns [23]. Anecdotally, new admin users typically created a campaign and began enrolling participants into TIM within 1 to 2 weeks of gaining access to the platform (personal communication, CW). Survey respondents often reported uses for both community members and staff, reflecting TIM's flexibility; in fact, some admin user agencies implemented multiple concurrent campaigns for different types of participants. Agencies used TIM to monitor participants for the development of symptoms, as well as the worsening of symptoms among confirmed COVID-19 cases. Often used with other support tools and activities, TIM was also

integrated into other program outbreak response operations, such as contact tracing.

According to respondents, TIM not only successfully identified symptomatic participants—some of whom were later identified as confirmed COVID-19 cases—but also did so in a timely manner. The survey results indicated that admin users were generally satisfied with TIM, comparing it favorably to previous systems used, especially in terms of cost, timeliness of data, data completeness, staff burden, time burden, and ability to monitor large populations. This parallels the evaluation findings of an Australian SMS text messaging program administered during a 2013 poultry farm outbreak of avian influenza to monitor for symptom development among exposed individuals. The study found that monitoring via SMS text messages was less time-consuming and more cost-effective than conducting telephone follow-up [9]. TIM was rated less favorably in terms of effort to enroll participants and data accuracy. We did not define the term “accuracy” in the survey instrument. However, the pilot interviews reflected that some admin users were concerned that monitored participants may be less likely to accurately report their symptoms via text compared to speaking to monitoring staff via phone. Given that survey respondents may have interpreted the term “accuracy” differently, we are unable to make a conclusion. This could be explored and verified with current TIM admin users.

TIM is similar to other COVID-19 text-based systems in that resources and staff time are required for monitoring daily reports for follow-up, oversight, and data protection safeguards [14,15]. Our qualitative findings highlighted respondents' particular challenges with TIM, including lack of interface features and cell phone and service access. CDC's implementation of the most feasible requested interface enhancements to TIM may alleviate some of the staff time burden. Since the majority of respondents who mentioned “texts not received or sent” also reported cell phone and service access issues, message interruptions likely came from outside of the TIM platform.

Some respondents indicated needing to set up multiple systems to ensure 100% capture (eg, using phone calls to augment texts). While this may be required for full participation, direct texting may be the most desirable mechanism for most participants. An evaluation of a COVID-19 text message system used in Maine, United States, found that the majority of participants who agreed to be monitored via an automated system preferred direct text (60.2%) versus texted weblink (21.15%), telephone call (7.8%), or emailed weblink (7.3%) [15]. The vast majority (89%) of admin user respondents for this evaluation indicated their overall satisfaction with TIM and reported that they would recommend TIM for managing symptom monitoring activities.

The survey also collected data about admin user experience with technical support issues and requests. The CDC TIM Support Team is composed of one full-time technical lead, one full-time technical coordinator, one 50%-time data analyst, and one 25%-time senior lead, along with the application contractor. Most of the team consisted of Emergency Operation Center responder staff, which has a high turnover requiring regular training. The team provided orientations and onboarding assistance for new, potential admin users; updated and shared

training materials; hosted TIM question-and-answer sessions; and provided daily responses to technical support requests. Respondents were overwhelmingly satisfied with the support provided by the TIM Support Team as well as the documented guidance provided for onboarding. This illustrates the essential nature of support staff when developing SMS text monitoring services meant to be used by diverse audiences.

## Recommendations

Based on these findings, we offer several recommendations for the TIM platform and others considering the use of SMS-based tools for symptom monitoring. First, providing alternative delivery mechanisms to the dashboard, such as text messages for viewing symptom alerts, may improve user experience by eliminating the need for admin users to log in to TIM to view alerts data on the TIM dashboard. Although email alerts are supported by TIM, this option is limited to allowing only one admin user to be able to receive symptom and nonresponse alerts.

Second, the promotion of TIM to other potential admin users could be improved by highlighting the comparative advantages of TIM reported by the survey respondents. Additionally, CDC can clarify the current limitations of TIM's customizability and provide guidance during the admin user onboarding process for assessing the appropriateness of this tool for certain populations. For example, administrators in rural areas with poor cellular coverage, or where populations lack mobile phones, may want to consider deploying an alternative or complementary system that addresses these circumstances. In some public health emergencies, it may be warranted to provide mobile phones to individuals who lack such devices and only require symptom monitoring for a limited period of time [24]. On a larger scale, the value of rural infrastructure that includes expanded mobile service is highlighted by the necessary capacity of public health authorities to conduct routine public health contact tracing and symptom monitoring [25].

Future evaluations of TIM could include a comparison of quantitative metrics reflecting retention, accuracy, and cost between TIM and the alternative and existing systems based on phone screening or in-person visits. We also suggest adopting continuous program evaluation practices to assess and respond to implementation gaps. Future evaluations could also be expanded to include participants' experience with TIM. Additionally, the TIM team can use administrative data to flag issues in real time. For example, multiple nonresponse alerts can indicate that a participant is not receiving texts. Additionally, the TIM team can categorize and tag help-desk tickets to prioritize potential areas for added functionality or training. Along similar lines, this evaluation and others [16] highlight the importance of engaging with primary and secondary users of tools such as TIM to validate assumptions and understand user perspectives. Last, the cohort of TIM admin users who are no longer using the system may also represent an important audience from which to collect feedback on user experience and reasons for discontinuation. While this study did attempt to recruit this population, few previous admin users responded; therefore, targeted recruitment efforts are likely required.

## Limitations

This evaluation has some notable limitations. Though the survey response rate was higher in terms of agency-level representation, the low response rate (6.8%) among all current and previous TIM admin users potentially introduces bias because survey respondents may not share the same opinions as nonrespondents. The low response rate may have been due to the high work demands and high turnover among staff working on the COVID-19 response during the first peak event combined with the holiday season. Additionally, multiple admin users per agency were eligible to respond to the survey, which could introduce sampling bias. However, an agency may have had multiple, ongoing campaigns, each of which may have been managed by a different TIM user at the agency. The data presented in this manuscript are not meant to be generalizable to all past or current TIM admin users. Also, the survey's cross-sectional design and single data collection point precluded analysis across time points in the COVID-19 pandemic.

The survey was shared with all current and previous TIM admin users and, therefore, had to accommodate varying levels of familiarity and expertise for TIM implementation. Consequently, a large proportion of respondents (up to 36%) reported "I don't know" responses to the questions that invited comparisons of TIM to previously used systems and were removed from that analysis. It is unclear why so many participants responded this way to these questions, but some potential reasons include being responsible for limited aspects of TIM administrative activities; not being trained in, or not having experience with, specific features of the system; and lack of knowledge about the previous or alternative systems that were used.

Most respondents were the single survey respondent for an agency (n=27). Most other instances of multiple respondents from an agency were limited to 2, 3, 4, 5, and 6 respondents (n=9, 1, 1, 1, and 1, respectively). However, there were 3 agencies that had high representation: Florida Health Department (n=17), "IHS\_Unspecified" (n=9), and CDC (n=9). Because Florida manages county-level health departments at the state level, we suspect that respondents from the Florida Health Department who reported state affiliations actually work at the county level. Out of 9 respondents representing "IHS\_Unspecified," 8 were from four different regions. Nevertheless, there is potential for respondents from the same agency to be similar and, thus, bias the results. We were unable to conduct intraclass correlation due to the limited sample size.

Additionally, CDC was unable to obtain several important administrative data elements that would have helped tell a more complete story about symptom and nonresponse alert rates, since the TIM admin user agreement precluded CDC access to these data. TIM's allowance for episodic use (ie, allowing admin users to start and stop use at any time) meant that many enrolled admin users had the option to monitor symptoms for periods of time as short as days or weeks, potentially going on to never use the platform again. Consequently, some of the outcomes under study (eg, comparisons to other systems and identification of COVID-19 cases) may have been difficult to realize or observe for some admin users.

## Implications and Conclusions

TIM could continue to play a valuable role in state and local health departments' COVID-19 responses when case numbers decrease and when more intensive contact tracing efforts resume to bring transmission rates to zero. TIM could potentially be used for post-COVID-19 vaccination safety monitoring. It can also be applied to other large-scale infectious disease outbreaks that feature a finite period for symptom monitoring, such as

avian influenza. Other possible use cases for TIM beyond infectious disease outbreak response include monitoring of adverse reactions to vaccination, medication adherence, and health and symptoms of persons with certain chronic diseases or substance use disorders. While the potential uses of tools such as TIM are limitless, the role of evaluation to understand the user experience will remain essential for ensuring their successful implementation.

## Acknowledgments

We would like to thank the following individuals who played a key role in the development and implementation of TIM: Lisa Austin (Vigilant Watch Integration, Inc), Grant Ballew (Compliant Campaign), Milan Chuttani (Oak Ridge Institute for Science and Education), Jason Domsy (MicroAutomation), Seth Edmunds (Centers for Disease Control and Prevention), Rachel Gallegos (Centers for Disease Control and Prevention), Todd Milner (Vigilant Watch Integration, Inc), Ethan Morris (Centers for Disease Control and Prevention), Vinutha Setty (General Dynamics Information Technology), Brittany Sunshine (Centers for Disease Control and Prevention), Abbey Wojno (Centers for Disease Control and Prevention), and Nasha Wonsley (General Dynamics Information Technology). The findings and conclusions in this paper are those of the authors and do not necessarily represent the official position of Centers for Disease Control and Prevention.

## Conflicts of Interest

None declared.

### Multimedia Appendix 1

Interview instrument.

[PDF File (Adobe PDF File), 227 KB - [publichealth\\_v8i2e32680\\_app1.pdf](#)]

### Multimedia Appendix 2

Survey instrument.

[PDF File (Adobe PDF File), 133 KB - [publichealth\\_v8i2e32680\\_app2.pdf](#)]

## References

1. Hightow-Weidman L, Beagle S, Pike E, Kuruc J, Leone P, Mobley V, et al. "No one's at home and they won't pick up the phone": Using the internet and text messaging to enhance partner services in North Carolina. *Sex Transm Dis* 2014 Mar;41(2):143-148. [doi: [10.1097/OLQ.0000000000000087](#)] [Medline: [24413497](#)]
2. Udeagu CN, Bocour A, Shah S, Ramos Y, Gutierrez R, Shepard CW. Bringing HIV partner services into the age of social media and mobile connectivity. *Sex Transm Dis* 2014 Oct;41(10):631-636. [doi: [10.1097/OLQ.0000000000000181](#)] [Medline: [25211262](#)]
3. Hochberg CH, Berringer K, Schneider JA. Next-generation methods for HIV partner services: A systematic review. *Sex Transm Dis* 2015 Sep;42(9):533-539 [FREE Full text] [doi: [10.1097/OLQ.0000000000000335](#)] [Medline: [26267881](#)]
4. Centers for Disease Control and Prevention (CDC). Recommendations for partner services programs for HIV infection, syphilis, gonorrhea, and chlamydial infection. *MMWR Recomm Rep* 2008 Nov 07;57(RR-9):1-83; quiz CE1 [FREE Full text] [Medline: [18987617](#)]
5. Toolkit for Technology-based Partner Services. US Centers for Disease Control and Prevention. 2020. URL: <https://www.cdc.gov/std/program/ips/default.htm> [accessed 2021-05-24]
6. Berrouguet S, Baca-García E, Brandt S, Walter M, Courtet P. Fundamentals for future mobile health (mHealth): A systematic review of mobile phone and web-based text messaging in mental health. *J Med Internet Res* 2016 Jun 10;18(6):e135 [FREE Full text] [doi: [10.2196/jmir.5066](#)] [Medline: [27287668](#)]
7. Lawson A, Dalfen A, Murphy KE, Milligan N, Lancee W. Use of text messaging for postpartum depression screening and information provision. *Psychiatr Serv* 2019 May 01;70(5):389-395. [doi: [10.1176/appi.ps.201800269](#)] [Medline: [30717643](#)]
8. Tracey LE, Regan AK, Armstrong PK, Dowse GK, Effler PV. EbolaTracks: An automated SMS system for monitoring persons potentially exposed to Ebola virus disease. *Euro Surveill* 2015 Jan 08;20(1):20999 [FREE Full text] [doi: [10.2807/1560-7917.es2015.20.1.20999](#)] [Medline: [25613652](#)]
9. Stephenson LM, Biggs JS, Sheppard V, Oakman TL. An evaluation of the use of short message service during an avian influenza outbreak on a poultry farm in Young. *Commun Dis Intell Q Rep* 2016 Jun 30;40(2):E195-E201 [FREE Full text] [Medline: [27522128](#)]

10. Stockwell MS, Reed C, Vargas CY, Camargo S, Garretson AF, Alba LR, et al. MoSAIC: Mobile Surveillance for Acute Respiratory Infections and Influenza-Like Illness in the Community. *Am J Epidemiol* 2014 Dec 15;180(12):1196-1201 [FREE Full text] [doi: [10.1093/aje/kwu303](https://doi.org/10.1093/aje/kwu303)] [Medline: [25416593](https://pubmed.ncbi.nlm.nih.gov/25416593/)]
11. Science Brief: Options to Reduce Quarantine for Contacts of Persons with SARS-CoV-2 Infection Using Symptom Monitoring and Diagnostic Testing. Centers for Disease Control and Prevention. 2020 Dec 02. URL: <https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/scientific-brief-options-to-reduce-quarantine.html> [accessed 2022-02-02]
12. Contact Tracing in the Context of COVID-19: Interim Guidance. Geneva, Switzerland: World Health Organization; 2020 May 10. URL: <https://apps.who.int/iris/handle/10665/332049> [accessed 2021-05-22]
13. Yap KY, Xie Q. Personalizing symptom monitoring and contact tracing efforts through a COVID-19 web app. *Infect Dis Poverty* 2020 Jul 13;9(1):93 [FREE Full text] [doi: [10.1186/s40249-020-00711-5](https://doi.org/10.1186/s40249-020-00711-5)] [Medline: [32660568](https://pubmed.ncbi.nlm.nih.gov/32660568/)]
14. Barrett PM, Bambury N, Kelly L, Condon R, Crompton J, Sheahan A, on behalf of the regional Department of Public Health. Measuring the effectiveness of an automated text messaging active surveillance system for COVID-19 in the south of Ireland, March to April 2020. *Euro Surveill* 2020 Jun;25(23):2000972 [FREE Full text] [doi: [10.2807/1560-7917.ES.2020.25.23.2000972](https://doi.org/10.2807/1560-7917.ES.2020.25.23.2000972)] [Medline: [32553064](https://pubmed.ncbi.nlm.nih.gov/32553064/)]
15. Krueger A, Gunn JKL, Watson J, Smith AE, Lincoln R, Huston SL, et al. Characteristics and outcomes of contacts of COVID-19 patients monitored using an automated symptom monitoring tool- Maine, May-June 2020. *MMWR Morb Mortal Wkly Rep* 2020 Aug 07;69(31):1026-1030 [FREE Full text] [doi: [10.15585/mmwr.mm6931e2](https://doi.org/10.15585/mmwr.mm6931e2)] [Medline: [32759918](https://pubmed.ncbi.nlm.nih.gov/32759918/)]
16. 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 09;22(9):e21956 [FREE Full text] [doi: [10.2196/21956](https://doi.org/10.2196/21956)] [Medline: [32791493](https://pubmed.ncbi.nlm.nih.gov/32791493/)]
17. Stewart RJ, Rossow J, Eckel S, Bidol S, Ballew G, Signs K, et al. Text-based illness monitoring for detection of novel influenza A virus infections during an influenza A (H3N2)v virus outbreak in Michigan, 2016: Surveillance and survey. *JMIR Public Health Surveill* 2019 Apr 26;5(2):e10842 [FREE Full text] [doi: [10.2196/10842](https://doi.org/10.2196/10842)] [Medline: [31025948](https://pubmed.ncbi.nlm.nih.gov/31025948/)]
18. Framework for Program Evaluation in Public Health. *MMWR*; 48 (No. RR-11). Atlanta, GA: US Centers for Disease Control and Prevention; 1999 Sep 17. URL: <https://www.cdc.gov/mmwr/PDF/rr/rr4811.pdf> [accessed 2021-07-23]
19. Saldaña J. *The Coding Manual for Qualitative Researchers*. 3rd edition. Thousand Oaks, CA: SAGE Publications; 2016.
20. MacQueen KM, McLellan E, Kay K, Milstein B. Codebook development for team-based qualitative analysis. *Cult Anthropol Methods* 2016 Sep 14;10(2):31-36. [doi: [10.1177/1525822X980100020301](https://doi.org/10.1177/1525822X980100020301)]
21. Ryan G. Measuring the typicality of text: Using multiple coders for more than just reliability and validity checks. *Hum Organ* 1999 Sep;58(3):313-322. [doi: [10.17730/humo.58.3.g224147522545rln](https://doi.org/10.17730/humo.58.3.g224147522545rln)]
22. Patton MQ. *Qualitative Research & Evaluation Methods*. 4th edition. Thousand Oaks, CA: SAGE Publications; 2014.
23. Megnin-Viggars O, Carter P, Melendez-Torres GJ, Weston D, Rubin GJ. Facilitators and barriers to engagement with contact tracing during infectious disease outbreaks: A rapid review of the evidence. *PLoS One* 2020;15(10):e0241473 [FREE Full text] [doi: [10.1371/journal.pone.0241473](https://doi.org/10.1371/journal.pone.0241473)] [Medline: [33120402](https://pubmed.ncbi.nlm.nih.gov/33120402/)]
24. Joseph HA, Wojno AE, Winter K, Grady-Erickson O, Hawes E, Benenson GA, et al. The Check and Report Ebola (CARE+) program to monitor travelers for Ebola after arrival to the United States, 2014-2016. *Public Health Rep* 2019;134(6):592-598 [FREE Full text] [doi: [10.1177/0033354919878165](https://doi.org/10.1177/0033354919878165)] [Medline: [31600452](https://pubmed.ncbi.nlm.nih.gov/31600452/)]
25. Howard B, Morris T. *Tribal Technology Assessment: The State of Internet Service on Tribal Lands*. Phoenix, AZ: American Indian Policy Institute, Arizona State University; 2019. URL: [https://aipi.asu.edu/sites/default/files/tribal\\_tech\\_assessment\\_compressed.pdf](https://aipi.asu.edu/sites/default/files/tribal_tech_assessment_compressed.pdf) [accessed 2022-01-30]

## Abbreviations

- admin user:** administrative user
- CDC:** Centers for Disease Control and Prevention
- FAQs:** frequently asked questions
- HIPAA:** Health Insurance Portability and Accountability Act
- IHS:** Indian Health Service
- REDCap:** Research Electronic Data Capture
- TIM:** Text Illness Monitoring

*Edited by T Sanchez; submitted 05.08.21; peer-reviewed by A Krueger, K Taira; comments to author 23.09.21; revised version received 28.10.21; accepted 30.11.21; published 07.02.22.*

*Please cite as:*

*Joseph HA, Ingber SZ, Austin C, Westnedge C, Strona FV, Lee L, Shah AB, Roper L, Patel A*

*An Evaluation of the Text Illness Monitoring (TIM) Platform for COVID-19: Cross-sectional Online Survey of Public Health Users*

*JMIR Public Health Surveill 2022;8(2):e32680*

*URL: <https://publichealth.jmir.org/2022/2/e32680>*

*doi: [10.2196/32680](https://doi.org/10.2196/32680)*

*PMID: [34882572](https://pubmed.ncbi.nlm.nih.gov/34882572/)*

©Heather A Joseph, Susan Z Ingber, Chelsea Austin, Caroline Westnedge, F V Strona, Leslie Lee, Ami B Shah, Lauren Roper, Anita Patel. Originally published in JMIR Public Health and Surveillance (<https://publichealth.jmir.org>), 07.02.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.



Original Paper

# COVID-19 Assessment and Testing in Rural Communities During the Pandemic: Cross-sectional Analysis

Jonathan Fitzsimon<sup>1,2</sup>, MBChB, CCFP; Oliver Gervais<sup>3</sup>, MD, CCFP; Chelsea Lanos<sup>4</sup>, MSc

<sup>1</sup>Department of Family Medicine, University of Ottawa, Ottawa, ON, Canada

<sup>2</sup>Arnprior Regional Health, Arnprior, ON, Canada

<sup>3</sup>Department of Family Medicine, Discipline of Emergency Medicine, Memorial University, St. John's, NL, Canada

<sup>4</sup>County of Renfrew Paramedic Service, Renfrew, ON, Canada

**Corresponding Author:**

Jonathan Fitzsimon, MBChB, CCFP

Department of Family Medicine

University of Ottawa

Suite 201

600 Peter Morand Crescent

Ottawa, ON, K1G 5Z3

Canada

Phone: 1 613 878 8225

Fax: 1 343 761 2921

Email: [jfitzsi2@uottawa.ca](mailto:jfitzsi2@uottawa.ca)

## Abstract

**Background:** The COVID-19 pandemic exacerbated the need for urgent improvements in access to health care for rural, remote, and underserved communities. The Renfrew County Virtual Triage and Assessment Centre (VTAC) was designed to provide access to COVID-19 testing and assessment with a family physician. The goal was to protect emergency departments and 911 paramedics while ensuring that nobody was left at home, suffering in silence. Residents were encouraged to call their own family physician for any urgent health needs. If they did not have a family physician or could not access their usual primary care provider, then they could call VTAC. This study reports on the output of a service model offering access to assessment and COVID-19 testing through a blend of virtual and in-person care options by a multidisciplinary team.

**Objective:** The purpose of this study was to assess the ability of VTAC to provide access to COVID-19 assessment and testing across rural, remote, and underserved communities.

**Methods:** We conducted a cross-sectional analysis of the data derived from the cases handled by VTAC between March 27, 2020 (launch day), and September 30, 2020.

**Results:** Residents from all 19 census subdivisions and municipalities of Renfrew County accessed VTAC. A total of 10,086 family physician assessments were completed (average 64 per day). Of these, 8535 (84.6%) assessments were to unique patient users. Thirty physicians provided care. Using digital equipment setup in the patients' home, 31 patients were monitored remotely. VTAC community paramedics completed 14,378 COVID-19 tests and 3875 home visits.

**Conclusions:** Renfrew County's experience suggests that there is tremendous synergy between family physicians and community paramedics in providing access to COVID-19 assessment and COVID-19 testing. The blended model of virtual and in-person care is well suited to provide improved access to other aspects of health care post pandemic, particularly for patients without a family physician.

(*JMIR Public Health Surveill* 2022;8(2):e30063) doi:[10.2196/30063](https://doi.org/10.2196/30063)

**KEYWORDS**

healthcare; virtual care; access; COVID-19; pandemic; assessment; testing; community paramed; digital health; online health; physician; virtual health

## Introduction

On March 11, 2020, the World Health Organization announced that the novel coronavirus outbreak first identified in China could be characterized as a pandemic [1]. Since then, the World Health Organization released a response plan to help countries prevent and delay outbreaks and improve patient care [2]. In response to this, the Ontario provincial government asked local health authorities to establish COVID-19 assessment centers [3].

In many urban settings, large buildings were repurposed, allowing health care workers to assess patients and perform COVID-19 testing. While functional and realistic for larger, densely populated areas, this was impractical for rural and remote areas. Patients suspected of having COVID-19 in these regions often relied on emergency departments (EDs) to provide basic care, threatening to increase risk of community spread, reduce capacity for responding to genuine emergencies (including severe symptoms due to COVID-19 infection), and dissuade patients from seeking treatment for other conditions for fear of infection [4].

We created the Renfrew County Virtual Triage and Assessment Centre (VTAC) to improve access to COVID-19 assessment and testing in rural, remote, and underserved communities. Through a local advertising campaign, using social media, radio interviews, and roadside notice boards, the residents were encouraged to contact their own family physician for urgent health needs. Those without a family physician or unable to reach their usual primary care provider could call VTAC. Patients call a toll-free telephone number and speak to a trained medical receptionist, with experience working in a physician's office, who could provide advice and schedule appointments with a physician when required. Family physicians provide virtual care by telephone and video encounters. Community paramedics offer COVID-19 testing at scheduled drive-through sites, ad hoc pop-up sites, and in-home settings for vulnerable housebound patients. Referral to existing community health care services enables patients to access a wide array of community support without resorting to the ED.

The primary purpose of this paper was to describe and quantify the output of the Renfrew County VTAC in the first 6 months of operation, from March 2020 to September 2020. We anticipate that our findings could help improve VTAC service delivery in Renfrew County and guide the implementation of

similar systems in other communities. Secondly, we gathered data that highlight the impact an ongoing system such as VTAC might have on increasing access to care for rural and underserved populations.

## Methods

### Study Design

We conducted a cross-sectional analysis of the data derived from the cases handled by VTAC between March 27, 2020 (launch day), and September 30, 2020.

### Setting

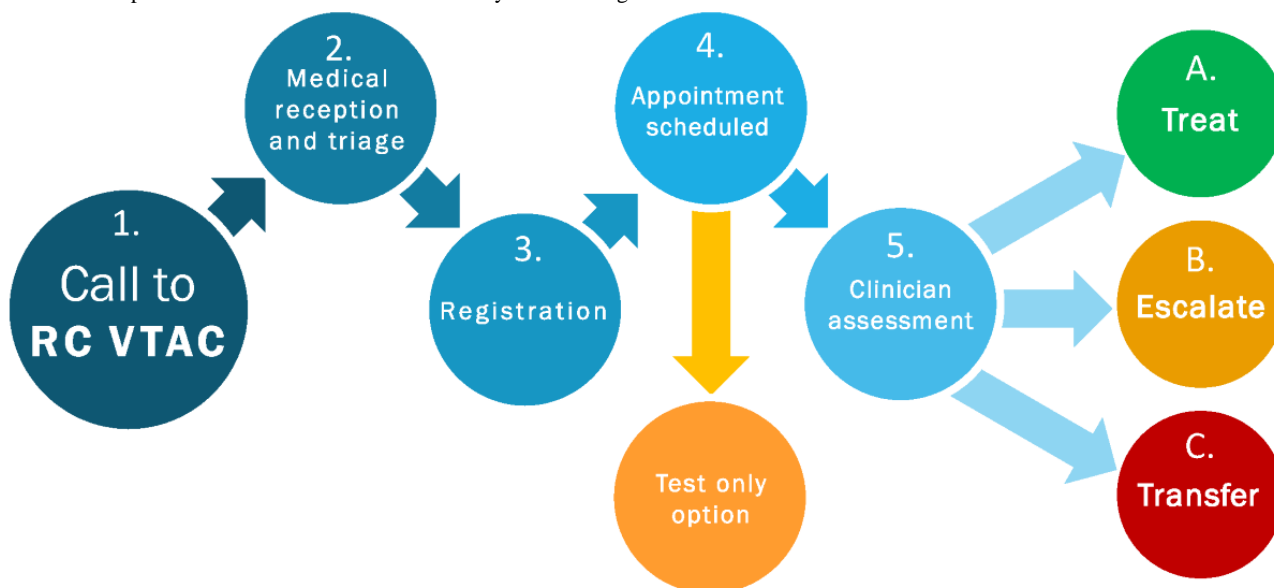
Renfrew County is the largest geographic county in Ontario, encompassing almost 7500 km<sup>2</sup>, with a population of approximately 107,000. Five larger towns in the county have community hospitals (including ED). A chronic shortage of family physicians has resulted in gross inequalities in access to primary care across the county. There are no walk-in clinics, and there is an enormous overreliance on ED as a means of accessing any form of health care.

### VTAC Development

Close integration of primary care and paramedic services in the region allowed for a model where paramedics would undertake all COVID-19 testing. Initially carried out entirely in patients' homes, this testing model gradually expanded to include pop-up sites for small groups and eventually various prescheduled drive-through sites across the county. Vulnerable housebound patients and those unable to attend a drive-through site were still able to be tested at home by community paramedics. The key message given to the public was, "If you have a health concern, call your family physician first. If you do not have a family physician or cannot access your family physician, then call VTAC."

### VTAC Model

Patients accessed VTAC by calling a toll-free phone number (Figure 1). A medical receptionist recorded the patient's demographic and other relevant information, including whether the patient had an existing family physician. For first time callers to VTAC, the receptionist created a basic medical chart in the electronic medical records (EMRs). Previous medical information could be retrieved electronically through Ontario's existing eHealth systems, Clinical Viewer and the Ontario Laboratories Information system.

**Figure 1.** VTAC patient flow. RC VTAC: Renfrew County Virtual Triage and Assessment Centre

Digital home monitoring equipment could be left in the patient's home, with community paramedics completing in-person patient teaching on how to use the equipment. Paramedics and a VTAC physician could then monitor the patients remotely and be available for check-ins at short notice. Clear referral pathways were created so that VTAC physicians could refer patients to the existing community health care resources such as mental health support and community palliative care as well as specialist services via electronic or telephone referral processes. Once all of these options are exhausted, VTAC physicians can advise patients to go to the ED or call 911 (Figure 1).

Physicians are compensated using Ontario's COVID-19 billing codes, created specifically for designated COVID-19 assessment centers [5]. The costs for paramedics, receptionists, and other administration are covered through Ontario's assessment center funding and managed by one of the hospital chief executive officers.

Thirty physicians provided care through VTAC. Most were physicians with an existing practice based in Renfrew County. Others worked in nearby Ottawa but had experience of working in Renfrew County through previous residency rotations or locum positions. Time commitments ranged from a few hours each week to regular, multiple shifts each week.

### Data Collection

Despite multiple attempts over many years, it has not been possible to obtain accurate or complete data regarding the number of residents who do not have a dedicated primary care provider (PCP). Therefore, family physicians and managers from primary care and hospital settings gathered data on the number of family physicians currently practicing in Renfrew County and their roster size. Primary care groups reported the number of patients with a nurse practitioner as their PCP. Of the 77 family physicians currently practicing in Renfrew County, 70 (90%) returned roster sizes from their EMR (individually or through their group EMR administrator). Moreover, 4 (5%) family physicians estimated their roster size as they currently

use paper charts, and 3 (3%) family physicians did not respond, in which case, local colleagues estimated their roster size.

We used routine utilization data for VTAC services in Renfrew County, which included the following elements: numbers of physician assessments; paramedic home visits; COVID-19 tests performed; physician assessments for patients without a family physician or alternate PCP; and gender, age, and postal code. Data were collected from the EMRs of VTAC patients by running a series of reports using built-in EMR capabilities, then transferred to Excel spreadsheets to identify repeat users, as well as the total number of events and unique patient users. For patients with an existing family physician, a copy of the VTAC encounter note was sent to their family physician to ensure integration with their regular EMR. We also cross-referenced postal codes with data from the 2016 Census [6] and IntelliHealth Ontario [7]. This allowed the evaluation of VTAC usage from each census subdivision within Renfrew County.

### Analysis

The analysis of data on the population of Renfrew County showed a small difference in the published figures from the 2016 census (102,394) [6] and IntelliHealth's 2020 report (107,286) [7]. We used the IntelliHealth data as the more recent review for results and cost analysis. We further analyzed the VTAC data sets to establish a more detailed demographic breakdown of age and gender.

We evaluated cost by reviewing VTAC budget reports of administrative costs and nonphysician staff costs, as well as billings submitted by VTAC physicians to the Ministry of Health. At the start of the pandemic, there was a willingness and availability of clinical, administrative, and management staff from hospitals, primary care teams, and other health care services to provide support in-kind to the efforts to combat the pandemic and provide a COVID-19 assessment center for the community. In October 2020, a new funding model for assessment centers was approved by the Government of Ontario [8]. Our budget planning moving forward includes the costs for support and services that have been partially or fully offered

in-kind during the initial pandemic period. Currently, there is a more stable prediction of future costs based on observation of trends, volumes, and usage of the service over 6 months of operation since inception.

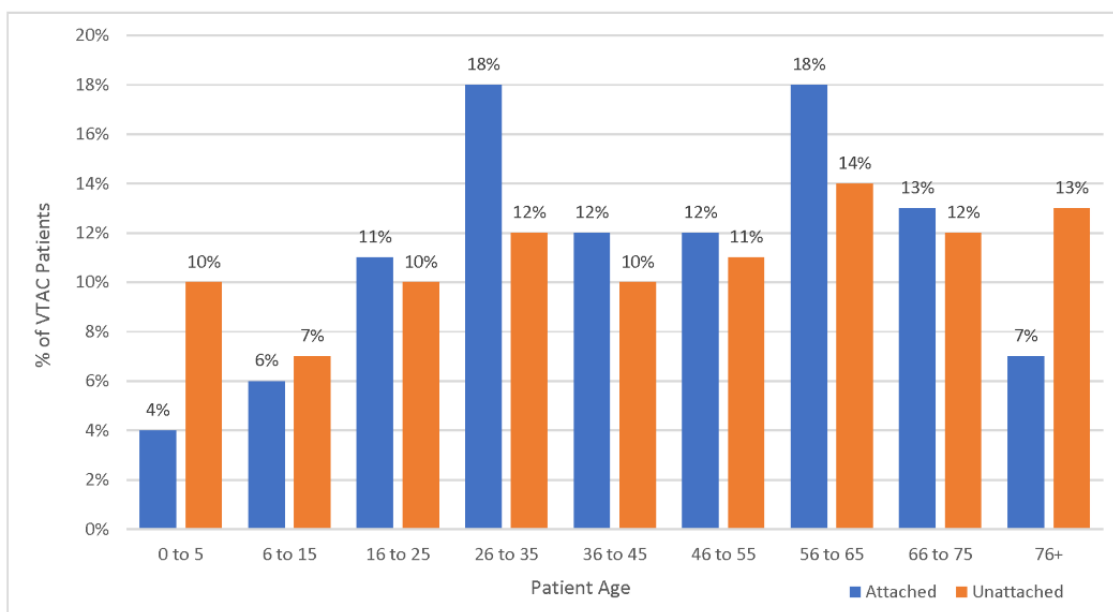
### Ethics Approval

Based on criteria provide by the Winchester District Memorial Hospital Research Ethics Board, this project was undertaken as a quality improvement initiative; thus, review by a research ethics board was not needed.

## Results

There were 82,450 patients registered with a practicing family physician and 2070 with a nurse practitioner, indicating that 22,766 residents (21.2% of Renfrew County residents) were not registered with a PCP. Patients with a family physician were slightly older on average than patients without a family physician (Figure 2), with 25% (n=1223) of attached patients being over the age of 65 versus 20% (n=728) of unattached patients. The proportion of male and female patients was similar between groups (62% [n=3034] of attached patients were female vs 58% [n=2111] of unattached patients).

**Figure 2.** Distribution of ages of VTAC patients unattached (n=3641) and attached (n=4894) to a primary care provider. VTAC: Virtual Triage and Assessment Centre.



### VTAC Usage

Residents from all census subdivisions and municipalities used VTAC. A total of 10,086 family physician assessments were completed, an average of 64 assessments per day. Of these, 8535 (84.6%) assessments were to unique patient users, of whom 3641 (42.7%) did not have a PCP (Figure 2).

Approximately one-third of the encounters were specifically COVID-19-related, and two-thirds were not directly COVID-19-related. This reduced unnecessary ED visits and the associated risks of infection and spread of COVID-19 within ED. Only 3% (302 of 10,086) of VTAC assessments resulted in a transfer to ED or 911.

Community paramedics supported by a VTAC physician remotely monitored 31 patients. A total of 1058 remote assessment days were completed, with many of these patients being monitored closely to prevent transfer to ED or admission to hospital. VTAC community paramedics completed 14,378 COVID-19 tests, including 3875 home visits.

### The Cost of VTAC

The total cost of establishing and operating VTAC up to September 30, 2020, was \$2,695,725. The actual costs

comprised community paramedics (\$1,262,042), physician fees (\$904,030), medical receptionists (\$266,364), information technology and communications (\$40,775), medical equipment (\$31,257), and other administrative support (\$11,256). In-kind support had an estimated value of \$180,000, which included management and administrative support (\$120,000); internet-based telephone system and tech support from a partnership with Ontario 211 community and social services (\$28,000); and infrastructure support from Telus, such as the use of the Practice Solutions Suite EMR situated at the Arnprior and District Family Health Team and reduction or waiving of various usual fees (\$32,000). A currency exchange rate of CAD \$1=US \$0.78 is applicable.

By using this EMR from the beginning, VTAC was able to establish several basic tracking forms to enable subsequent data collection. VTAC was used by 22,944 unique patients and was available to all residents. Overall, the total cost was equivalent to \$25.13 for every resident of Renfrew County for the setup and first 6 months of operating costs.

## Discussion

### Principal Findings

The primary aim of this manuscript was to describe and quantify the output of the Renfrew County VTAC in the first 6 months of operation. During that period, VTAC was accessed by nearly 23,000 unique patients for COVID-19 testing, virtual family physician assessment, community paramedic in-home assessment, or remote assessment using in-home monitoring equipment. Residents in rural, remote, urban, and First Nation communities who have used VTAC reported a high level of patient acceptability and satisfaction with the service. In addition to providing access to COVID-19 testing and assessment, VTAC has provided a vital and potentially cost-effective alternative to the ED for residents who do not have a primary care provider or are unable to access their PCP.

VTAC has attracted significant amounts of media attention. Local political leaders have fully endorsed VTAC due to the contribution it has made to the population of Renfrew County during the pandemic and the overwhelming desire of the community for the service to continue [9]. Moreover, VTAC has been shown to be highly accepted by its users. In a recent study among 219 VTAC patients, 86% reported that their concern was dealt with at the first virtual encounter, 93% reported being happy or very happy with the service, and 98% would recommend VTAC to family and friends. In addition, 46% of the 219 VTAC users reported that without VTAC, they would have attended an ED [10]. Therefore, since its implementation, VTAC has proven to be an effective vehicle for delivering high-quality acute, episodic care in rural, remote, and underserved communities.

Previous articles have described aspects of virtual care during the COVID-19 pandemic [11-17]. Some Canadian provinces have implemented virtual walk-in clinics, providing assessment by a physician through a mobile phone app [18]. One study described the expansion of a United States virtual telehealth program, which provides 24/7 access to a virtual acute care physician [16]. Another article described the triage and assessment of suspected COVID-19 patients originating from ED, hospital discharges, and PCPs. Patients determined to be safe for discharge were admitted to a "Virtual Ward" (ie, their home) for remote oximetry monitoring. Local paramedics could be called to arrange transport to hospital at any time, based on preestablished criteria suggestive of deterioration. The study concluded that this was a safe method for managing COVID-19 patients, and there were overall significant cost savings [19]. Although these virtual programs have some similarities to VTAC, VTAC is unique in its combination of many different new and innovative approaches to create a comprehensive virtual triage, assessment, and monitoring program for patients who do not have access to a PCP. These approaches include a tight partnership producing tremendous synergy between primary care and community paramedicine, the implementation of virtual care technologies (including remote monitoring), a close collaboration with hospitals and public health, and a partnership with a wide array of preexisting health care and community resources.

However, VTAC is not a replacement for comprehensive primary care. Complex health issues, such as chronic pain management, are poorly suited to virtual, episodic care [18]. Some complaints, including some musculoskeletal and neurological symptoms, are generally not amenable to virtual care [20]. Moreover, VTAC does not replace the continuity, depth, and breadth of care that is offered by a regular family physician and a strong, long-term doctor-patient relationship. However, our results showing that 21.2% of Renfrew County residents do not have a PCP are in keeping with the 2020 Canadian survey data, which found that 20% of Canadians have no family doctor [21]. The problem is exacerbated in Renfrew County by the absence of any walk-in clinics. The Canada-wide lack of access to PCPs strengthens the potential benefits that a program similar to VTAC could have in other communities. In addition to providing an alternative to the ED for those who do not have access to a PCP, virtual care has been shown to be cost-effective from both health and societal perspectives [11,22].

Notably, the VTAC model is well suited to assist with a COVID-19 community mass vaccination program. Community paramedics have implemented highly effective and efficient drive-through sites where they currently perform COVID-19 testing. It is within the scope of practice of the community paramedics to administer vaccinations. The physical sites are easily adaptable to include in-vehicle waiting areas (to monitor patients and be immediately available to deal with any postvaccination side effects). From the outset, the sites were "winterized" and can be capable of functioning in a range of weather conditions. These drive-through sites can accommodate large numbers of people in a timely manner and have the capacity to expand quickly to meet surges in demand.

The VTAC model could also provide a basis of longer-term health care support to the underserved population by providing an alternative to the ED for patients who do not have a family physician but require regular, chronic disease management. Urgent care can also be provided to patients without a family physician or when a patient's regular family physician is not available. One of the key benefits of having a family physician make these assessments is that they are ideally placed to identify high risk signs and symptoms that do warrant emergency care. This reassures patients that their need for emergency care is justified and warranted while helping to protect emergency services for genuine emergencies. Family physicians are also well versed in managing risk and being able to explain clear safety netting steps to patients so that they understand the circumstances where additional or repeat assessment would be required.

### Strengths and Limitations

Our study has several limitations. It was conducted in 1 region (Renfrew County), and its findings might not be generalizable to other settings. There are limitations which are inherent to virtual care, including challenges obtaining vital signs, limited ability to perform a physical exam and subsequent risk of misdiagnosis [15]. Despite these limitations, VTAC's overwhelmingly positive patient feedback is comparable to the positive satisfaction rates noted in other telemedicine studies [11,12,23,24]. As COVID-19 risk stratification scores such as

RECAP-V0 (Remote COVID-19 Assessment in Primary Care) and DDC19 (dynamic risk assessment decision support system for COVID-19) are further validated, we hope to incorporate these into our assessments, helping to more objectively determine which patients are at the highest risk of deterioration [25,26].

Detailed information regarding VTAC video appointments were not specifically collected. Therefore, we cannot comment on the proportion of video visits during the study period, or what patient issues or conditions led to video or telephone appointments. This was, in part, due to the overwhelming majority of VTAC encounters being sufficiently dealt with and completed by phone. However, VTAC physicians were able to use a secure video platform if required. Furthermore, although VTAC was accessible for both COVID-19 and non-COVID-19 related issues, no specific insights were available for a detailed analysis of the suitability of VTAC by health concern. Further research is underway to follow up with VTAC physicians and analyze these issues. Lastly, while we were able to estimate the total cost equivalent for VTAC for the setup and first 6 months of operating (cost for every resident of Renfrew County), a detailed evaluation of the economic impact on overall health care costs was beyond the scope of this paper. Future research aiming to evaluate both the clinical and economic impacts of VTAC in Renfrew County is warranted.

## Conclusion

The introduction of VTAC in Renfrew County has greatly enhanced access to COVID-19 assessment and testing in rural communities during the pandemic, especially for individuals unattached to a primary care provider. Virtual care has proven to be overwhelmingly acceptable to VTAC patients and has improved their experience of health care and health outcomes. Physicians have quickly adapted to different techniques and benefitted from the advantages of being able to provide care despite the restrictions of the pandemic. When compared to the costs of 911 transfer, ED attendance, and hospital admissions, VTAC may provide a highly cost-effective improvement to the overall health care system. The VTAC structure is also well suited to assist with a COVID-19 community mass vaccination program; once the COVID-19 pandemic eventually subsides, the need for COVID-19 assessment centers will diminish. Furthermore, the VTAC model could provide a basis of health care support to the underserved population; it can also be a stepping-stone toward a situation whereby all Canadians have access to a family physician and comprehensive primary care. Future research should be aimed at assessing the clinical and economic impacts of the implementation and ongoing use of VTAC in Renfrew County.

## Conflicts of Interest

None declared.

## References

1. Coronavirus disease (COVID-19) pandemic. World Health Organization. URL: <https://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/novel-coronavirus-2019-ncov> [accessed 2020-12-05]
2. Responding to community spread of COVID-19: interim guidance. World Health Organization. 2020 Mar 7. URL: <https://www.who.int/publications/i/item/responding-to-community-spread-of-covid-19>
3. Recommendations for COVID-19 Assessment Centres. Ontario Health. URL: <https://www.ontariohealth.ca/sites/ontariohealth/files/2020-05/Recommendations%20for%20COVID-19%20Assessment%20Centres.pdf>
4. Patey C, Asghari S, Norman P, Hurley O. Redesign of a rural emergency department to prepare for the COVID-19 pandemic. CMAJ 2020 May 11;192(19):E518-E520 [FREE Full text] [doi: [10.1503/cmaj.200509](https://doi.org/10.1503/cmaj.200509)] [Medline: [32317277](https://pubmed.ncbi.nlm.nih.gov/32317277/)]
5. Ontario Health Insurance Plan INFOBulletin. Ontario Ministry of Health. URL: <https://www.seamo.ca/sites/seamo.ca/files/page/files/Update%20Temporary%20COVID-19%20Vaccine%20Fee%20Codes.pdf>
6. Census Profile, 2016 Census: Renfrew, County [Census division], Ontario and Saskatchewan [Province]. Statistics Canada. URL: <https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/page.cfm?Lang=E&Geo1=CD&Code1=3547&Geo2=PR&Code2=47&Data=Count&SearchText=Renfrew&SearchType=Begin&SearchPR=01&B1=All> [accessed 2020-12-05]
7. Welcome to IntelliHealth. IntelliHealth. URL: <https://intellihealth.moh.gov.on.ca/> [accessed 2022-01-28]
8. North American COVID-19 Policy Response Monitor: Ontario. North American Observatory on Health Systems and Policies. URL: [https://ihpme.utoronto.ca/wp-content/uploads/2021/02/ON-COVID19-Response-Monitor\\_20210221.pdf](https://ihpme.utoronto.ca/wp-content/uploads/2021/02/ON-COVID19-Response-Monitor_20210221.pdf) [accessed 2022-01-28]
9. Renfrew County Wins Municipal Innovation Award for Online Health Assessment Centre. Association of Municipalities Ontario. 2021 Aug 17. URL: <https://www.amo.on.ca/about-us/news-releases/renfrew-county-wins-municipal-innovation-award-online-health-assessment-centre> [accessed 2022-01-28]
10. Charbonneau V, Volyeva A. Effectiveness of Renfrew County Virtual Triage and Assessment Centre in Reducing Emergency Department Visits. County of Renfrew Health Committee 2020.
11. Ferguson JM, Jacobs J, Yefimova M, Greene L, Heyworth L, Zulman DM. Virtual care expansion in the Veterans Health Administration during the COVID-19 pandemic: clinical services and patient characteristics associated with utilization. J Am Med Inform Assoc 2021 Mar 01;28(3):453-462 [FREE Full text] [doi: [10.1093/jamia/ocaa284](https://doi.org/10.1093/jamia/ocaa284)] [Medline: [33125032](https://pubmed.ncbi.nlm.nih.gov/33125032/)]

12. Tan N, Yang LW, Tan MZ, Chng J, Tan MH, Tan C. Virtual care to increase military medical centre capacity in the primary health care setting: A prospective self-controlled pilot study of symptoms collection and telemedicine. *J Telemed Telecare* 2020 Oct 04;1357633X2095957. [doi: [10.1177/1357633x20959579](https://doi.org/10.1177/1357633x20959579)]
13. Sinha S, Kern LM, Gingras LF, Reshetnyak E, Tung J, Pelzman F, et al. Implementation of Video Visits During COVID-19: Lessons Learned From a Primary Care Practice in New York City. *Front Public Health* 2020 Sep 17;8:514 [FREE Full text] [doi: [10.3389/fpubh.2020.00514](https://doi.org/10.3389/fpubh.2020.00514)] [Medline: [33042950](https://pubmed.ncbi.nlm.nih.gov/33042950/)]
14. Blazey-Martin D, Barnhart E, Gillis J, Vazquez GA. Primary Care Population Management for COVID-19 Patients. *J Gen Intern Med* 2020 Oct 27;35(10):3077-3080 [FREE Full text] [doi: [10.1007/s11606-020-05981-1](https://doi.org/10.1007/s11606-020-05981-1)] [Medline: [32720239](https://pubmed.ncbi.nlm.nih.gov/32720239/)]
15. Spelman JF, Brienza R, Walsh RF, Drost P, Schwartz AR, Kravetz JD, et al. A Model for Rapid Transition to Virtual Care, VA Connecticut Primary Care Response to COVID-19. *J Gen Intern Med* 2020 Oct 23;35(10):3073-3076 [FREE Full text] [doi: [10.1007/s11606-020-06041-4](https://doi.org/10.1007/s11606-020-06041-4)] [Medline: [32705471](https://pubmed.ncbi.nlm.nih.gov/32705471/)]
16. Joshi AU, Lewiss RE. Telehealth in the time of COVID-19. *Emerg Med J* 2020 Oct 04;37(10):637-638. [doi: [10.1136/emered-2020-209846](https://doi.org/10.1136/emered-2020-209846)] [Medline: [32753392](https://pubmed.ncbi.nlm.nih.gov/32753392/)]
17. Greenhalgh T, Koh GCH, Car J. Covid-19: a remote assessment in primary care. *BMJ* 2020 Mar 25;368:m1182. [doi: [10.1136/bmj.m1182](https://doi.org/10.1136/bmj.m1182)] [Medline: [32213507](https://pubmed.ncbi.nlm.nih.gov/32213507/)]
18. Hardcastle L, Ogbogu U. Virtual care: Enhancing access or harming care? *Healthc Manage Forum* 2020 Nov 20;33(6):288-292 [FREE Full text] [doi: [10.1177/0840470420938818](https://doi.org/10.1177/0840470420938818)] [Medline: [32686506](https://pubmed.ncbi.nlm.nih.gov/32686506/)]
19. Nunan J, Clarke D, Malakouti A, Tannetta D, Calthrop A, Xu XH, et al. Triage Into the Community for COVID-19 (TICC-19) Patients Pathway - Service evaluation of the virtual monitoring of patients with COVID pneumonia. *Acute Med* 2020;19(4):183-191. [Medline: [33215171](https://pubmed.ncbi.nlm.nih.gov/33215171/)]
20. Dermer M. Virtual Care Playbook. Canadian Medical Association. URL: [https://www.cma.ca/sites/default/files/pdf/Virtual-Care-Playbook\\_mar2020\\_E.pdf](https://www.cma.ca/sites/default/files/pdf/Virtual-Care-Playbook_mar2020_E.pdf) [accessed 2020-12-05]
21. What Canadians Think About Virtual Health Care: Nationwide Survey Results. Canadian Medical Association. 2020 May. URL: <https://www.cma.ca/sites/default/files/pdf/virtual-care/cma-virtual-care-public-poll-june-2020-e.pdf> [accessed 2020-12-05]
22. Garattini L, Badinella Martini M, Zanetti M. More room for telemedicine after COVID-19: lessons for primary care? *Eur J Health Econ* 2021 Mar 24;22(2):183-186 [FREE Full text] [doi: [10.1007/s10198-020-01248-y](https://doi.org/10.1007/s10198-020-01248-y)] [Medline: [33231825](https://pubmed.ncbi.nlm.nih.gov/33231825/)]
23. Olayiwola JN, Magaña C, Harmon A, Nair S, Esposito E, Harsh C, et al. Telehealth as a Bright Spot of the COVID-19 Pandemic: Recommendations From the Virtual Frontlines ("Frontweb"). *JMIR Public Health Surveill* 2020 Jun 25;6(2):e19045 [FREE Full text] [doi: [10.2196/19045](https://doi.org/10.2196/19045)] [Medline: [32479413](https://pubmed.ncbi.nlm.nih.gov/32479413/)]
24. Holtz BE. Patients Perceptions of Telemedicine Visits Before and After the Coronavirus Disease 2019 Pandemic. *Telemed J E Health* 2021 Jan 01;27(1):107-112. [doi: [10.1089/tmj.2020.0168](https://doi.org/10.1089/tmj.2020.0168)] [Medline: [32614689](https://pubmed.ncbi.nlm.nih.gov/32614689/)]
25. Liu Y, Wang Z, Ren J, Tian Y, Zhou M, Zhou T, et al. A COVID-19 Risk Assessment Decision Support System for General Practitioners: Design and Development Study. *J Med Internet Res* 2020 Jun 29;22(6):e19786 [FREE Full text] [doi: [10.2196/19786](https://doi.org/10.2196/19786)] [Medline: [32540845](https://pubmed.ncbi.nlm.nih.gov/32540845/)]
26. Greenhalgh T, Thompson P, Weiriga S, Neves AL, Husain L, Dunlop M, et al. What items should be included in an early warning score for remote assessment of suspected COVID-19? qualitative and Delphi study. *BMJ Open* 2020 Nov 12;10(11):e042626 [FREE Full text] [doi: [10.1136/bmjopen-2020-042626](https://doi.org/10.1136/bmjopen-2020-042626)] [Medline: [33184088](https://pubmed.ncbi.nlm.nih.gov/33184088/)]

## Abbreviations

**DDC19:** dynamic risk assessment decision support system for COVID-19

**ED:** emergency department

**EMR:** electronic medical record

**PCP:** primary care provider

**RECAP:** Remote COVID-19 Assessment in Primary Care

**VTAC:** Virtual Triage and Assessment Centre

*Edited by T Sanchez; submitted 29.04.21; peer-reviewed by J Oldenburg, AM Lopez; comments to author 06.06.21; revised version received 24.06.21; accepted 06.01.22; published 08.02.22.*

*Please cite as:*

*Fitzsimon J, Gervais O, Lanos C*

*COVID-19 Assessment and Testing in Rural Communities During the Pandemic: Cross-sectional Analysis*

*JMIR Public Health Surveill* 2022;8(2):e30063

URL: <https://publichealth.jmir.org/2022/2/e30063>

doi: [10.2196/30063](https://doi.org/10.2196/30063)

PMID: [35022158](https://pubmed.ncbi.nlm.nih.gov/35022158/)

©Jonathan Fitzsimon, Oliver Gervais, Chelsea Lanos. Originally published in JMIR Public Health and Surveillance (<https://publichealth.jmir.org>), 08.02.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.



Original Paper

# COVID-19 Surveillance Updates in US Metropolitan Areas: Dynamic Panel Data Modeling

Theresa B Oehmke<sup>1</sup>, PhD; Charles B Moss<sup>2</sup>, PhD; James F Oehmke<sup>3</sup>, PhD

<sup>1</sup>Department of Civil and Environmental Engineering, University of California, Berkeley, Berkeley, CA, United States

<sup>2</sup>Institute of Food and Agricultural Sciences, University of Florida, Gainesville, FL, United States

<sup>3</sup>Department of Emergency Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

**Corresponding Author:**

Theresa B Oehmke, PhD

Department of Civil and Environmental Engineering

University of California, Berkeley

202 O'Brien Hall

Berkeley, CA, 94720

United States

Phone: 1 5108986406

Email: [toehmke@berkeley.edu](mailto:toehmke@berkeley.edu)

## Abstract

**Background:** Despite the availability of vaccines, the US incidence of new COVID-19 cases per day nearly doubled from the beginning of July to the end of August 2021, fueled largely by the rapid spread of the Delta variant. While the “Delta wave” appears to have peaked nationally, some states and municipalities continue to see elevated numbers of new cases. Vigilant surveillance including at a metropolitan level can help identify any reignition and validate continued and strong public health policy responses in problem localities.

**Objective:** This surveillance report aimed to provide up-to-date information for the 25 largest US metropolitan areas about the rapidity of descent in the number of new cases following the Delta wave peak, as well as any potential reignition of the pandemic associated with declining vaccine effectiveness over time, new variants, or other factors.

**Methods:** COVID-19 pandemic dynamics for the 25 largest US metropolitan areas were analyzed through September 19, 2021, using novel metrics of speed, acceleration, jerk, and 7-day persistence, calculated from the observed data on the cumulative number of cases as reported by USAFacts. Statistical analysis was conducted using dynamic panel data models estimated with the Arellano-Bond regression techniques. The results are presented in tabular and graphic forms for visual interpretation.

**Results:** On average, speed in the 25 largest US metropolitan areas declined from 34 new cases per day per 100,000 population, during the week ending August 15, 2021, to 29 new cases per day per 100,000 population, during the week ending September 19, 2021. This average masks important differences across metropolitan areas. For example, Miami’s speed decreased from 105 for the week ending August 15, 2021, to 40 for the week ending September 19, 2021. Los Angeles, San Francisco, Riverside, and San Diego had decreasing speed over the sample period and ended with single-digit speeds for the week ending September 19, 2021. However, Boston, Washington DC, Detroit, Minneapolis, Denver, and Charlotte all had their highest speed of the sample during the week ending September 19, 2021. These cities, as well as Houston and Baltimore, had positive acceleration for the week ending September 19, 2021.

**Conclusions:** There is great variation in epidemiological curves across US metropolitan areas, including increasing numbers of new cases in 8 of the largest 25 metropolitan areas for the week ending September 19, 2021. These trends, including the possibility of waning vaccine effectiveness and the emergence of resistant variants, strongly indicate the need for continued surveillance and perhaps a return to more restrictive public health guidelines for some areas.

(*JMIR Public Health Surveill* 2022;8(2):e28737) doi:[10.2196/28737](https://doi.org/10.2196/28737)

**KEYWORDS**

surveillance system; COVID-19; coronavirus; Sars-CoV-2; Houston; dynamic panel data model; speed; jerk; acceleration; 7-Day persistence; modeling; data; surveillance; monitoring; public health; United States; transmission; response

## Introduction

Almost a year and a half after the first case of the disease caused by the novel coronavirus SARS-CoV2 was recorded in the United States, the country has reported over 42 million cases, over 680,000 deaths, 88 million initial unemployment claims filed from March 21, 2020, to the present, and an estimated 30,000 additional deaths attributable to pandemic-related unemployment [1,2]. Additionally, up to 30% of COVID-19 survivors and some asymptomatic individuals may be subject to sequelae including a deterioration in health-related quality of life that may last 9 months or longer [3-6]. Despite the unprecedented pandemic, the US government has not implemented national COVID-19 restrictions, instead leaving it to states and metropolitan areas to determine what public health measures are appropriate in their context, based on unclear guidelines [7-12]. This has led to a patchwork of measures and sometimes diverging national, state, and local guidance [7-9]. Additionally, unclear and divergent public health recommendations, among other factors, have likely affected public acceptance of measures such as social distancing, face masking, and crowd avoidance [13-15]. For example, despite the evidence that masks help prevent the spread of the virus, on March 6, protestors held a mask-burning rally on the steps of the Idaho state capitol [16], the first conviction in a plot to kidnap Michigan Governor Whitmer for imposing a mask mandate was handed down in August of 2021 [17], and Florida Governor DeSantis's ban of local mask mandates has been subject to court challenges and likely appeals [18].

After enduring shutdowns for most of the 2020-21 winter, most US metropolitan areas began reopening in early- to mid-March, 2021 [16,19,20], despite contrary guidance and concerns of explosive growth after reopening [21,22]. Reopening is at least chronologically associated with public desensitization to COVID-19 news and lack of compliance with health guidance [23]. Since June 2021, the emergence of the delta variant has led to a global third wave of COVID-19 cases including in the United States, with 98.4% of new cases in the United States comprising lineage B.1.617.2 [24]. There is significant concern that reopening has led to an upsurge and possibly explosive growth in the pandemic, especially as new variants spread across the country, and if either health policy or social acceptance of policy reduces or eliminates social distancing. The ramifications of the Delta and other variants for the reemergence of explosive growth under reopening conditions and with limited vaccination are unclear [25]. Thus, it is imperative to have timely and objective surveillance at local levels to inform public health decisions, including metrics such as the 7-day persistence rate that inform questions around where on the epidemiological curve the pandemic currently lies—whether it is in the state of explosive growth, peaking and declining, or that it is peaking with a plateau [26,27]. A large 7-day persistence value is indicative of explosive growth.

This report provides surveillance data relevant to the question of whether the pandemic is exploding, peaking, or plateauing for the 25 largest US metropolitan areas. Specifically, it provides weekly data for the 6 weeks ending on September 19, 2021, on the novel metrics of speed, acceleration, and jerk, and the 7-day

persistence effect. This captures information relevant to the determination of a possible peak and reversal in the most recent Delta-fueled wave of COVID-19 in the United States.

## Methods

We calculated the speed, acceleration, jerk, and 7-day persistence rate by applying dynamic panel data modeling and additional methods introduced by [7,26]. These methods have been previously applied globally at the country level [26-31] and in the United States at the state level [28] and comprise the basis for the Global SARS-CoV-2 Surveillance Project [32]. This paper updates the findings by Oehmke et al [27], who previously applied these methods to US metropolitan areas.

The metrics used in the paper are speed, acceleration, jerk, and persistence. Speed is the number of new metropolitan area cases per day per 100,000 population and is a measure of how fast the pandemic is growing from day to day. To address differences in reporting across different days of the week, we calculate speed (and other indicators) on a weekly basis for 7-day ISO (International Organization for Standardization) weeks [33]. Acceleration is the week-over-week change in weekly average speed and is the primary indication of whether the pandemic is getting worse (positive acceleration) or better (negative acceleration or deceleration) from week to week. Jerk is the change in acceleration. A positive jerk signifies increasing acceleration—not only is the pandemic getting worse, but it is getting worse more rapidly this week than last. A large positive jerk can be associated with super-spreading events, emergence of a new variant, policy shift, or other change that affects the underlying infection rates. A negative jerk indicates a declining acceleration possibly including a shift from positive to negative acceleration, and if associated with a policy shift, may indicate policy success. The 7-day persistence rate is the number of cases per 100,000 today that are statistically attributable to cases 7 days ago. A positive 7-day persistence rate can signify the presence of linked super-spreader events, the emergence of a new variant, or continued policy ineffectiveness. During the first year of the pandemic, persistence was associated with mega-spreader events (eg, when on a weekend people were infected at a sports event, religious gathering, political rally, etc, they go to another event the next weekend and infect more people, who then go to another event, leading to a persistent “echo forward” effect of the original infection [27]). Mathematically, persistence case numbers that are close to or exceeding 100% of total cases lead to rapid or explosive growth in the number of cases. A small positive persistence may indicate a diminishing pandemic. A negative number indicates that the high number of cases last week is associated with a lower number of cases this week, suggesting that the high number of cases was an aberration rather than an indication of a persistent issue in the metropolis. A negative 7-day persistence rate is indicative of a slowing pandemic, whether via a natural progression or through public health control measures.

We collected data from USAFacts for the week of 2021 ending August 1, 2021, through the week ending September 19, 2021. Using the US Census Bureau definitions of metropolitan area, we determined the various counties comprising the 25 largest

US metropolitan areas and collected data on the total (cumulative) number of COVID-19 cases in each county for each day. We preprocessed the data only by creating a variable containing the number of new cases per county per day to show the change in the cumulative case count from the prior day to that day (first difference). We used data from the first 2 weeks, ie, the weeks ending August 1, 2021, and August 8, 2021, to create lagged values used in the analysis, resulting in a sample period of 6 weeks from the week ending August 15, 2021, to the week ending September 19, 2021, for which we had a complete data set. We analyzed the data using STATA/MP, version 17.0 (Stata Corp LLC); the dynamic panel data modeling is accomplished using STATA's "xtabond" procedure. We report the weekly average speed, acceleration, jerk, and 7-day persistence effect for the most recent 6 weeks of 2021. The results for the prior weeks of 2021 are available from the authors upon request.

We also report the vaccination rates in California and Florida, which are useful in interpreting the results. Vaccination data are taken from the USAFacts website. [Multimedia Appendix 1](#) contains additional methodological detail.

## Results

**Table 1** reports the population-weighted average results for the 25 metropolitan areas. Speed increased from 34.03 cases per day per 100,000 persons during the week ending August 15, 2021, to a peak of 36.84 during the week ending September 5, 2021, before declining to 29.01 for the week ending September 19, 2021. Acceleration started at a positive 3.96, then declined and turned negative during the weeks ending September 12, 2021, and September 19, 2021. Jerk was negative for the first 5 weeks of the period, before turning positive during the week ending September 19, 2021. The positive jerk coupled with a negative acceleration (deceleration) means that the deceleration is slowing, which is potentially indicative of a slow decline or upcoming plateau in the number of cases. Persistence started at 1.67 new cases per day per 100,000 that were statistically attributable to cases from the prior week, during the week ending August 15, 2021. Persistence increased to a peak of 2.22 the week ending August 29, 2021, before declining to a value of 1.06 the week ending September 19, 2021.

**Table 1.** Population-weighted averages of reported COVID-19 incidents for the 25 most populous metropolitan areas in the United States.

Week ending	Population-weighted averages			
	Speed (reported cases per day per 100,000 people)	Acceleration (reported cases per day per day per 100,000 people)	Jerk (reported cases per day per day per day per 100,000 people)	Persistence (number of cases during 7-day period due to cases in pervious 7-day period)
August 15, 2021	34.03	3.96	-2.89	1.67
August 20, 2021	36.27	2.24	-1.72	1.91
August 29, 2021	36.80	.53	-1.71	2.22
September 5, 2021	36.84	.04	-.49	2.17
September 12, 2021	31.89	-4.95	-4.99	1.26
September 19, 2021	29.01	-2.88	2.07	1.06

Disaggregating the results by metropolitan area reveals a variety of patterns over the 6 weeks. We present an illustrative selection of these patterns graphically; complete results in tabular form are available in [Multimedia Appendix 2](#) [34-38].

**Figure 1** shows illustrative temporal patterns for speed. The population-weighted average showed a slight increase in speed through the week ending September 5, 2021, then a slight decrease. Chicago followed the national pattern of a slight increase over the first 3 or 4 weeks of the sample period, and then a slight decrease in speed. Other cities following this pattern include New York, Houston, Atlanta, Philadelphia, Phoenix, Boston, San Francisco, Riverside, Orlando, Tampa, Seattle, San Diego, and Portland. Miami started off the sample period at its

peak with a high rate of speed, which declined significantly over the sample period especially in the last 2 weeks; Tampa and Orlando followed the national pattern of peaking in mid-August but had significant decreases in speed over the last 2 weeks, similarly to Miami. Los Angeles also saw a significant decline in speed over the period, but from a lower starting point. San Francisco and San Diego, while mirroring the national peak, declined rapidly in the end of the sample period to significantly decrease their speed, with San Francisco and San Diego reaching single-digit speeds. San Antonio experienced a small decrease in speed over the sample period. Dallas, Washington DC, Detroit, Minneapolis, Baltimore, Charlotte, St. Louis, and Portland saw slightly increasing speed over the period.

**Figure 1.** Illustrative patterns for speed; metro average, Miami, Chicago, Dallas, and San Francisco; weeks ending August 15, 2021, to September 19, 2021.

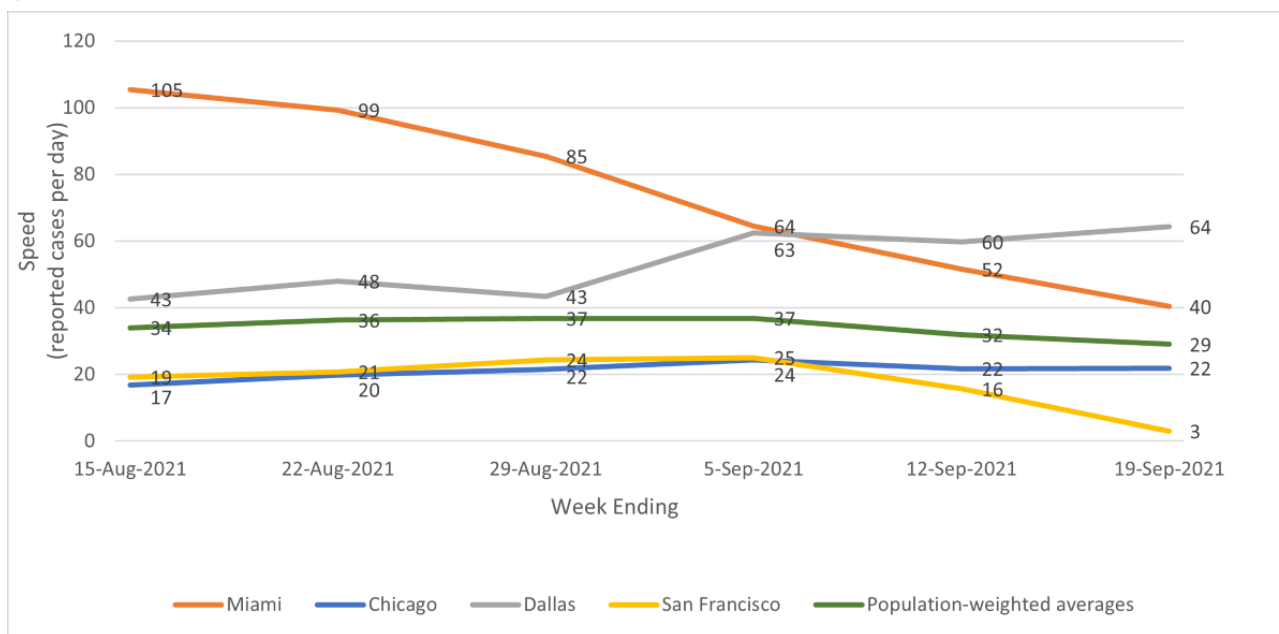


Figure 2 shows illustrative temporal patterns for acceleration. The population-weighted average acceleration started slightly positive and ended slightly negative. Chicago’s acceleration mirrored this pattern before ending at essentially zero for the week ending September 19, 2021. Other cities following this pattern include New York, Houston, Atlanta, Philadelphia, Phoenix, Boston, San Francisco, Riverside, Orlando, Tampa, Seattle, San Diego, and Portland. Miami started with positive acceleration, but then began a rapid deceleration (negative

acceleration) that is reflected in a still declining number of cases. Los Angeles, San Francisco, Orlando, Tampa, San Diego, Atlanta, and Riverside also ended the sample period with large decelerations, although the deceleration in San Francisco and the other California areas occurred primarily in the last 2 weeks of the sample period. Dallas, Washington DC, Detroit, Minneapolis, Baltimore, St. Louis, Charlotte, and Portland had some amount of fluctuation in acceleration, but ended the period with small positive accelerations.

**Figure 2.** Illustrative patterns for acceleration; metro average, Miami, Chicago, Dallas, and San Francisco; weeks ending August 15, 2021, to September 19, 2021.

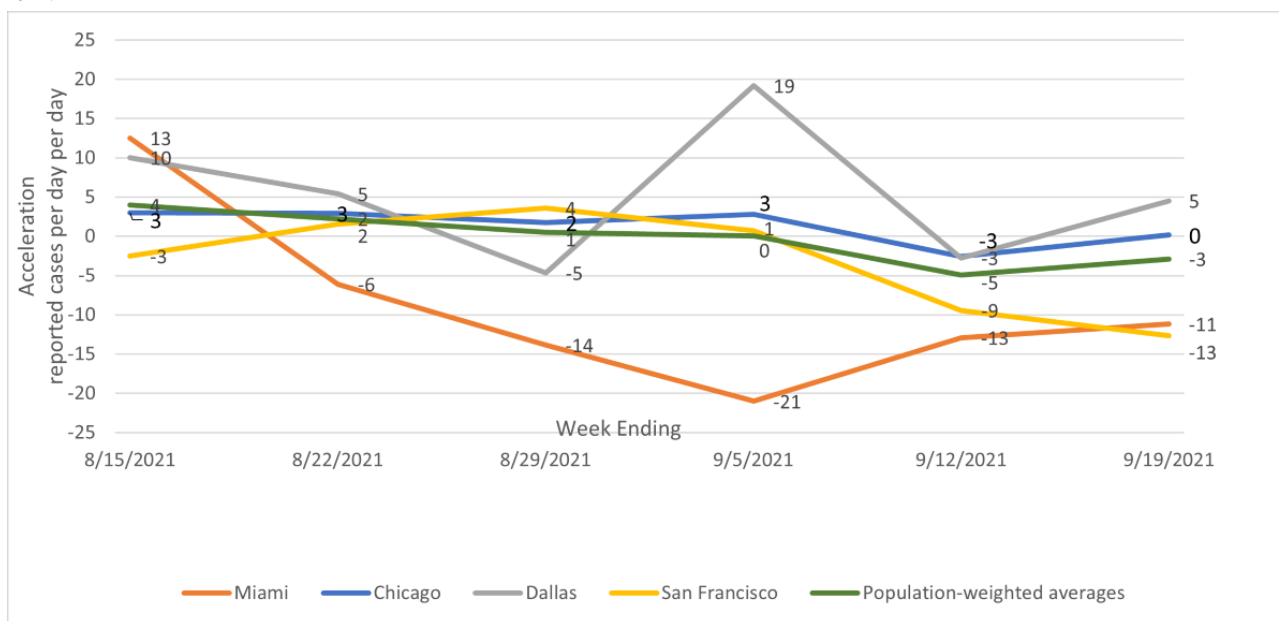


Figure 3 shows illustrative temporal patterns for jerk—the change in acceleration. The population-weighted average jerk changes little over the sample period: it is very slightly negative the week ending August 15, 2021, and very slightly positive the week ending September 19, 2021. Since acceleration was

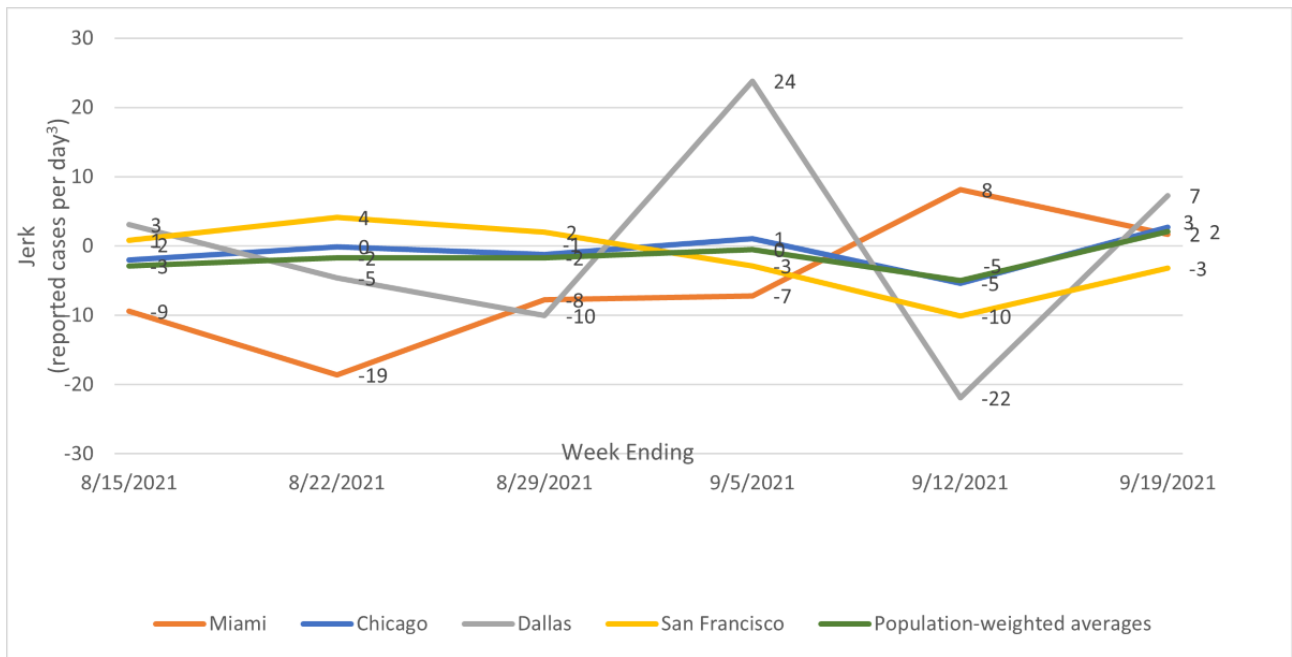
negative (deceleration) that week, the positive jerk means that the speed is still decreasing but less rapidly. In other words, the latest wave of the pandemic has peaked, but at a national level, it is not decreasing rapidly. Chicago follows almost exactly the population-weighted average. New York, Boston, Detroit,

Seattle, Minneapolis, Denver, and Baltimore were mostly similar. St. Louis inverted the average pattern, starting with small negative jerk, moving into a positive jerk, but ending the sample period with a small negative jerk. Dallas, Houston, Washington DC., Philadelphia, Phoenix, Charlotte, San Antonio, and Portland showed large fluctuations in jerk. Miami had negative jerk for the weeks ending August 15, 2021, through September 5, 2021, but jerk turned positive for the weeks ending

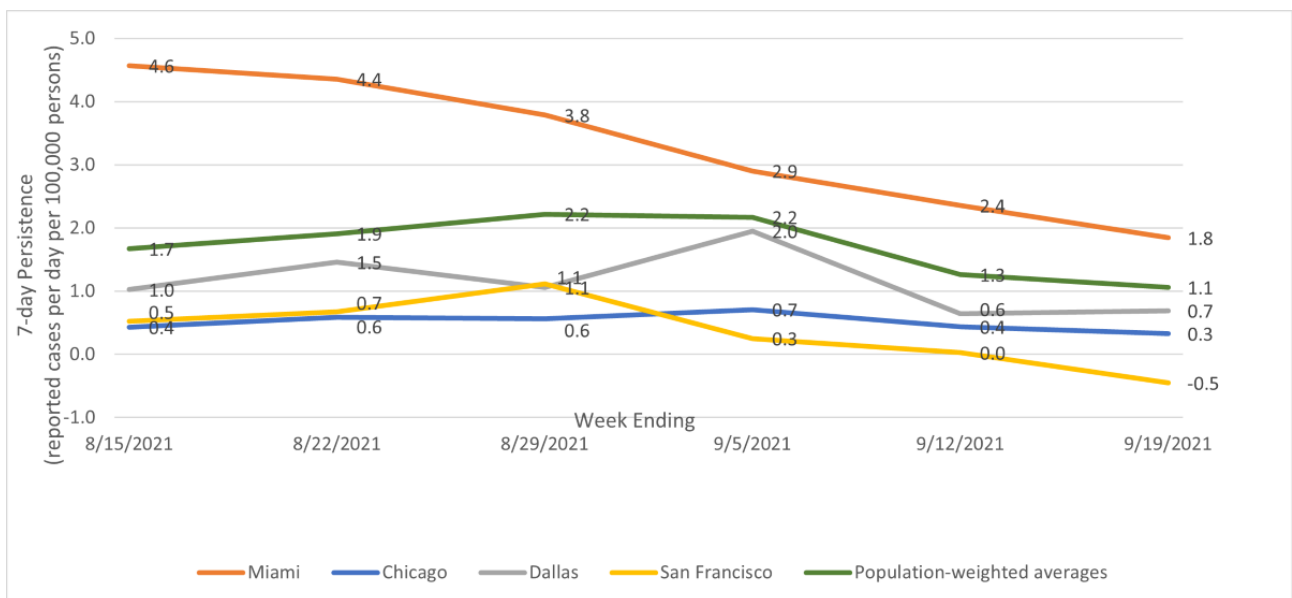
September 12, 2021, and September 19, 2021. Los Angeles, Atlanta, San Francisco, Riverside, San Diego, Tampa, and Orlando moved from positive or near-zero jerk to large negative jerk.

Figure 4 shows illustrative temporal patterns for persistence, the number of daily new cases statistically attributable to new cases 7 days previously.

**Figure 3.** Illustrative patterns for jerk; metro average, Miami, Chicago, Dallas, and San Francisco; weeks ending August 15, 2021, to September 19, 2021.



**Figure 4.** Illustrative patterns for persistence; metro average, Miami, Chicago, Dallas, and San Francisco; weeks ending August 15, 2021, to September 19, 2021.



The metropolitan average persistence started out at 1.7 new cases per day per 100,000 statistically attributable to new cases 7 days prior, peaked at 2.2 the week ending August 29, 2021, and then declined at 1.1 the week ending September 19, 2021. As a percentage of speed, metro average persistence peaked at 6.0% the week ending August 29, 2021, and then declined to

3.7% the week ending September 19, 2021. Of the 25 largest metropolitan areas, 22 (88%) followed a similar pattern.

Exceptions to the pattern were Washington DC, Charlotte, and St. Louis. Charlotte had a definite peak the week ending September 5, 2021, and a decline the following 2 weeks, ending

the sample period with a slightly higher persistence than at the start. Washington DC was similar to Charlotte but with lower numbers and a muted peak. St. Louis had a relatively high and flat persistence rate. As a percentage of total new cases, St. Louis had values from 20% to 26% over the sample period, which was notably higher than most other cities.

**Table 2** provides comparisons across the 25 largest US metropolitan areas for the week ending September 19, 2021. We report the surveillance metrics of speed, acceleration, jerk, and persistence for each metropolitan area, ranked by speed. Charlotte, NC has the fastest speed at 73 new cases per day per 100,000 population, positive acceleration and jerk, and the second-highest persistence at 6.94 new cases the week ending September 19, 2021 (meaning that 6.94 new cases per day are statistically attributable to new cases in the prior week). Tampa, Orlando, and Miami have negative acceleration, consistent with

their falling numbers of new cases over the entire sample period. St. Louis has the highest persistence at 7.01, meaning that over 21% of new cases during the week ending September 12, 2021, were statistically attributable to new cases the prior week. This is nearly 6 times the population-weighted average rate of 3.6%. At the bottom of the table, Riverside, San Diego, Los Angeles, and San Francisco all had single-digit speed and negative acceleration, jerk, and persistence.

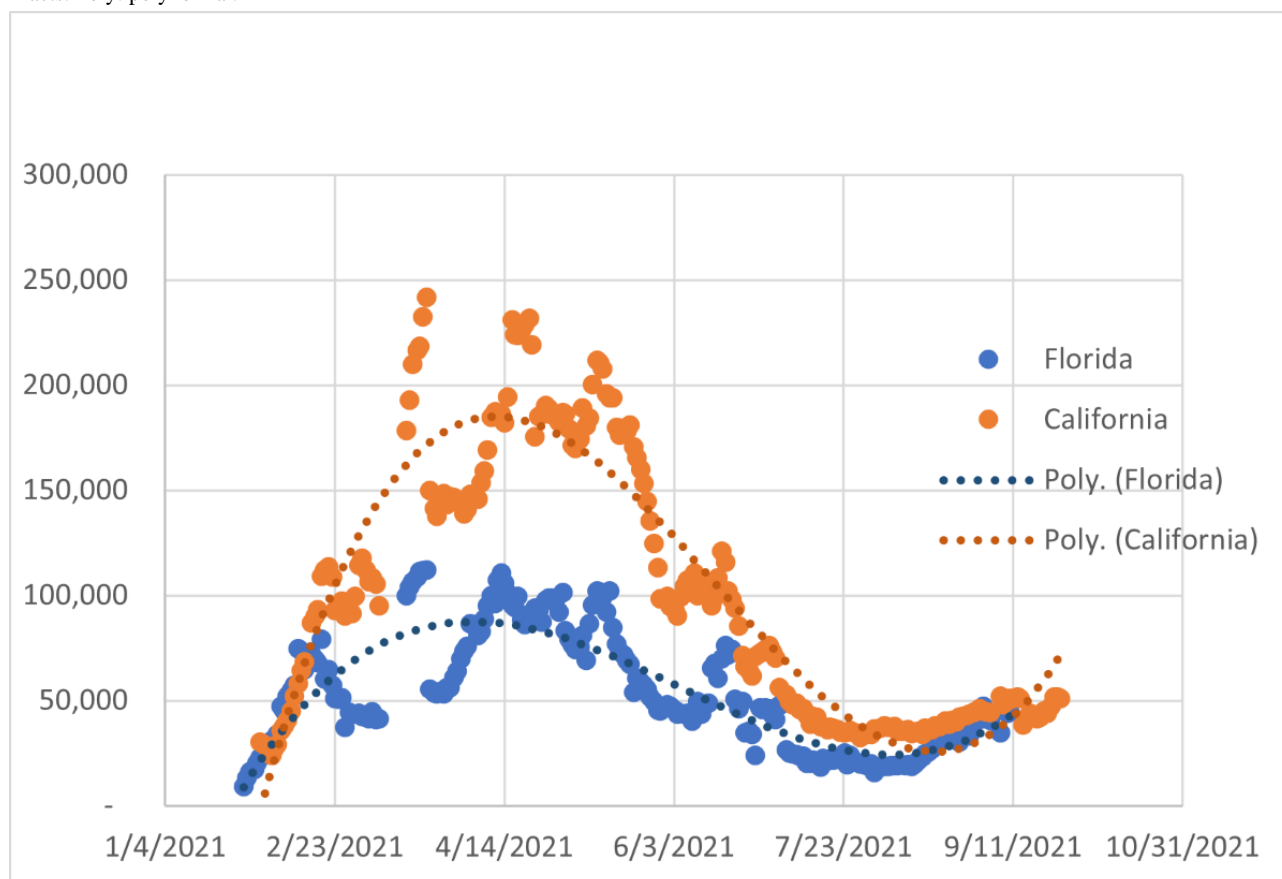
The correlation coefficient between the prior week's persistence number and the current week's speed is 0.577 ( $P < .001$ ).

For additional context in interpreting these results, we present vaccination rate data from California and Florida in [Figure 5](#). Both California and Florida saw the number of new vaccinations per day peak in mid-April and decline through the end of July. Both states had a smaller "second wave" of vaccinations that began about mid-August.

**Table 2.** Speed, acceleration, jerk, and 7-Day persistence for the 25 most populous metropolitan areas in the United States, September 13-19, 2021.

Metropolitan area	Population (number of people in each metropolitan area)	Speed (reported cases per day per 100,000 people)	Acceleration (reported cases per day per 100,000 people)	Jerk (reported cases per day per 100,000 people)	Persistence (number of cases during 7-day period due to cases in previous 7-day period)
Charlotte	2,636,883	73	11	22	6.94
Dallas	7,573,136	64	5	7	0.69
San Antonio	2,550,960	61	1	2	0.07
Tampa	3,194,831	56	-19	-2	1.10
Houston	7,066,141	56	6	24	0.66
Orlando	2,608,147	49	-12	7	5.18
Miami	6,166,488	40	-11	2	1.84
Phoenix	4,948,203	40	1	8	0.14
Atlanta	6,020,364	36	-8	-2	1.86
St. Louis	2,803,228	33	0	-3	7.01
Minneapolis	3,640,043	32	7	8	2.01
Portland	2,492,412	31	2	8	0.13
Seattle	3,979,845	31	0	3	0.27
Denver	2,967,239	28	1	1	1.06
Washington DC	5,371,160	24	5	6	3.39
Philadelphia	5,378,441	23	2	4	0.68
Boston	4,873,019	23	4	5	0.83
Detroit	4,319,629	23	2	2	0.80
Chicago	9,458,539	22	0	3	0.33
New York City	19,216,182	21	-1	1	0.70
Baltimore	2,800,053	16	1	2	1.11
Riverside	4,650,631	7	-20	-16	-0.41
San Diego	3,338,330	6	-17	-9	-0.25
Los Angeles	13,214,799	4	-13	-10	-0.07
San Francisco	4,731,803	3	-13	-3	-0.45

**Figure 5.** Change in fully vaccinated populations, Florida and California; 7-day moving averages and fitted cubic polynomial. Data sourced from USAFacts. Poly: polynomial.



## Discussion

### Principal Findings

The 25 largest metropolitan areas are exhibiting 4 or 5 distinct patterns of COVID-19. This is not surprising as areas differ in vaccination rates, health guidelines, especially around mask mandates, and public compliance with both mandates and voluntary guidelines. Overall, the third wave of COVID-19 in the largest US metropolitan areas, associated with the Delta variant, seems to have peaked. However, how and even whether the number of new cases declines will vary by metropolitan area.

The majority of the metropolitan areas are exhibiting indications that the Delta wave peaked in mid-August, and current indications are positive. However, declines in new cases since the peak have been slow, and the epidemiological curve seems to be flattening more than rapidly declining.

Charlotte, NC exhibited the highest speed among the metropolitan areas. This high speed is consistent with low vaccination rates: less than half of North Carolina residents are fully vaccinated, and there are 272 outbreaks in nursing homes [39,40]. The next 6 areas—Dallas, San Antonio, Tampa, Houston, Orlando, Miami—are in Texas and Florida, 2 states with legislation prohibiting mask mandates [39,40].

Miami, Tampa, and Orlando are exceptions and have exhibited the biggest declines in new cases per day per 100,000 population. However, they started from very high numbers of

new cases, so the current pandemic speed is significantly above average. California's metropolitan areas also experienced significant decelerations in September, although not as large as Florida's. However, California started at lower levels than did Florida; after deceleration, Los Angeles, San Francisco, Riverside, and San Diego all had single-digit speed during the week ending September 19, 2021. These metropolitan areas are likely close to defeating COVID-19.

The Florida pattern—the highest initial speeds during the week ending August 15, 2021, followed by the largest decline in speeds throughout the sample period for Miami and during the weeks ending September 12 and September 9, 2021, for Orlando and Tampa—is atypical among metropolitan areas. Anecdotally, the rapid decline is associated with an increase in vaccinations as the public responded to the increasing speed [41]. Figure 5 shows the daily change in the number of fully vaccinated persons in Florida as a 7-day moving average and a cubic polynomial curve. The daily increase (7-day moving average) in the number of fully vaccinated persons in Florida bottomed on August 12, 2021, and then began increasing, which is chronologically consistent with the declines in speed later in August and September. California's metropolitan areas also experienced significant decelerations in September, although not as large as Florida's. California also experienced an upswing in the rate of vaccination starting in mid-August, although proportionately the upswing was smaller than in Florida. Hence there is prima facie evidence that is consistent with the vaccination hypothesis. However, further research is necessary to draw any firm conclusions.

Throughout the analysis, St. Louis has presented uniquely, exhibiting a very mild trough instead of a peak, and a high persistence. We found a high and statistically significant correlation between last week's persistence and this week's speed; in the case of St. Louis, the high persistence number may indicate that the metropolitan area remains at risk of the pandemic worsening in the next few weeks.

However, examination of the novel surveillance metric of persistence revealed that no metropolitan area showed indications of explosive growth. The relatively small persistence numbers for the majority of metropolitan areas further suggest that the infection pattern for the Delta variant, compared with that of the Alpha or Beta variants, is less reliant on mega-spreader events and more likely to be spread by an infected person coming into casual contact with a susceptible person, or perhaps by encountering multiple susceptible persons sequentially rather than at a single event. This is heuristically consistent with the findings that Delta has a longer infectious period and higher viral shedding loads than earlier variants [42].

### Conclusions

The novel surveillance metrics in this paper help paint a more complete picture of the COVID-19 pandemic in the largest US

metropolitan areas. Acceleration and jerk provide additional information as to the direction of the pandemic (ie, where a metropolitan area lies on the epidemiological curve). Perhaps most importantly, last week's 7-day persistence is highly and significantly correlated with this week's speed, which suggests that 7-day persistence may add significant predictive power to the existing set of surveillance metrics. This also indicates the need for further research to clarify and refine the predictive power of the 7-day persistence metric.

In mid-August, the third or Delta wave of COVID-19 peaked in most, but not all, of the 25 largest US metropolitan areas. Washington DC and St. Louis are exceptions to this peak. Following the peak, new cases declined significantly in metropolitan areas in California and Florida, although from high starting values in Florida. In California, the Los Angeles, San Francisco, Riverside, and San Diego metropolitan areas have single-digit speed (new cases per day per 100,000), which, at least for the moment, establishes control over the pandemic. However, most metropolitan areas are showing only small decreases in the number of new cases after the peak. This leads to the conclusion that more needs to be done in these areas to control the pandemic, and additional monitoring, surveillance, and improved prediction are necessary.

---

### Conflicts of Interest

None declared.

---

#### Multimedia Appendix 1

Weekly surveillance summary 2021 for 25 largest U.S. metropolitan areas.

[[XLS File \(Microsoft Excel File\), 231 KB - publichealth\\_v8i2e28737\\_app1.xls](#) ]

---

#### Multimedia Appendix 2

Extended Methodology Section.

[[DOCX File , 19 KB - publichealth\\_v8i2e28737\\_app2.docx](#) ]

---

### References

1. Matthay EC, Duchowny KA, Riley AR, Galea S. Projected All-Cause Deaths Attributable to COVID-19-Related Unemployment in the United States. *Am J Public Health* 2021 Apr;111(4):696-699. [doi: [10.2105/AJPH.2020.306095](https://doi.org/10.2105/AJPH.2020.306095)] [Medline: [33600244](https://pubmed.ncbi.nlm.nih.gov/33600244/)]
2. Weekly Initial Claims. Federal Reserve Bank of St. Louis. URL: <https://fred.stlouisfed.org/categories/32240> [accessed 2021-12-16]
3. Huang C, Huang L, Wang Y, Li X, Ren L, Gu X, et al. 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. *Lancet* 2021 Jan 16;397(10270):220-232 [FREE Full text] [doi: [10.1016/S0140-6736\(20\)32656-8](https://doi.org/10.1016/S0140-6736(20)32656-8)] [Medline: [33428867](https://pubmed.ncbi.nlm.nih.gov/33428867/)]
4. Carfi A, Bernabei R, Landi F. Persistent Symptoms in Patients After Acute COVID-19. *JAMA* 2020 Aug 11;324(6):603. [doi: [10.1001/jama.2020.12603](https://doi.org/10.1001/jama.2020.12603)]
5. Logue JK, Franko NM, McCulloch DJ, McDonald D, Magedson A, Wolf CR, et al. Sequelae in Adults at 6 Months After COVID-19 Infection. *JAMA Netw Open* 2021 Feb 01;4(2):e210830 [FREE Full text] [doi: [10.1001/jamanetworkopen.2021.0830](https://doi.org/10.1001/jamanetworkopen.2021.0830)] [Medline: [33606031](https://pubmed.ncbi.nlm.nih.gov/33606031/)]
6. Proal AD, VanElzakker MB. Long COVID or Post-acute Sequelae of COVID-19 (PASC): An Overview of Biological Factors That May Contribute to Persistent Symptoms. *Front Microbiol* 2021 Jun 23;12:698169 [FREE Full text] [doi: [10.3389/fmicb.2021.698169](https://doi.org/10.3389/fmicb.2021.698169)] [Medline: [34248921](https://pubmed.ncbi.nlm.nih.gov/34248921/)]
7. Oehmke JF, Moss CB, Singh LN, Oehmke TB, Post LA. Dynamic Panel Surveillance of COVID-19 Transmission in the United States to Inform Health Policy: Observational Statistical Study. *J Med Internet Res* 2020 Oct 5;22(10):e21955. [doi: [10.2196/21955](https://doi.org/10.2196/21955)]



8. Parodi SM, Liu VX. From Containment to Mitigation of COVID-19 in the US. *JAMA* 2020 Apr 21;323(15):1441-1442. [doi: [10.1001/jama.2020.3882](https://doi.org/10.1001/jama.2020.3882)] [Medline: [32167525](https://pubmed.ncbi.nlm.nih.gov/32167525/)]
9. Ren X. Pandemic and lockdown: a territorial approach to COVID-19 in China, Italy and the United States. *Eurasian Geography and Economics* 2020 May 08;61(4-5):423-434. [doi: [10.1080/15387216.2020.1762103](https://doi.org/10.1080/15387216.2020.1762103)]
10. Ganz S. Does your state have 14-days of declining COVID-19 cases? American Enterprise Institute. URL: <https://www.aei.org/economics/does-dc-have-14-days-of-declining-covid-19-cases-the-data-are-unclear/> [accessed 2021-12-16]
11. Spinelli A, Pellino G. COVID-19 pandemic: perspectives on an unfolding crisis. *Br J Surg* 2020 Jun;107(7):785-787 [FREE Full text] [doi: [10.1002/bjs.11627](https://doi.org/10.1002/bjs.11627)] [Medline: [32191340](https://pubmed.ncbi.nlm.nih.gov/32191340/)]
12. Hodge, JR JG. Federal vs. State Powers in Rush to Reopen Amid Coronavirus Pandemic. Just Security. URL: <https://www.justsecurity.org/69880/federal-vs-state-powers-in-rush-to-reopen-amid-coronavirus-pandemic/> [accessed 2021-12-16]
13. Barrios J, Benmelech E, Hochberg Y, Sapienza P, Zingales L. Civic capital and social distancing during the Covid-19 pandemic. National Bureau of Economic Research. URL: [https://www.nber.org/system/files/working\\_papers/w27320/w27320.pdf](https://www.nber.org/system/files/working_papers/w27320/w27320.pdf) [accessed 2021-12-16]
14. Bavel JJV, Baicker K, Boggio PS, Capraro V, Cichocka A, Cikara M, et al. Using social and behavioural science to support COVID-19 pandemic response. *Nat Hum Behav* 2020 May 30;4(5):460-471. [doi: [10.1038/s41562-020-0884-z](https://doi.org/10.1038/s41562-020-0884-z)] [Medline: [32355299](https://pubmed.ncbi.nlm.nih.gov/32355299/)]
15. Brzezinski A, Deiana G, Kecht V, Dijcke D. The COVID-19 Pandemic: Government vs. Community Action Across the United States. Institute for New Economic Thinking 2020 Apr 18:1-43 [FREE Full text]
16. Bogel-Burroughs N. Covid-19 News: Republican Governors Push for Schools to Reopen. *The New York Times*. URL: <https://www.nytimes.com/live/2021/03/06/world/covid-19-coronavirus> [accessed 2021-12-16]
17. New York Times. Defendant in Plot to Kidnap Michigan Governor is Sentenced to Six Years. URL: <https://www.nytimes.com/2021/08/25/us/gretchen-whitmer-kidnapping-plot-michigan.html> [accessed 2021-12-16]
18. Spencer T, Anderson C. Judge blocks Florida governor's order banning mask mandates. *APNews*. URL: <https://apnews.com/article/lifestyle-health-education-florida-coronavirus-pandemic-1908088a0b5c5b02d89fd7e007822408> [accessed 2021-12-16]
19. New York Times. Texas Drops Its Virus Restrictions as a Wave of Reopenings Takes Hold. URL: <https://www.nytimes.com/2021/03/02/us/coronavirus-reopening-texas.html> [accessed 2021-12-16]
20. Kelly J. The States That Are Now Reopening And Not Requiring Mask-Wearing. *Forbes*. URL: <https://www.forbes.com/sites/jackkelly/2021/03/07/the-states-that-are-now-reopening-and-not-requiring-mask-wearing/?sh=4590670e168a> [accessed 2021-12-16]
21. 'Now Is Not The Time To Stop Wearing A Mask,' Says CDC Director Rochelle Walensky. *National Public Radio*. URL: <https://www.npr.org/2021/03/03/973286247/now-is-not-the-time-to-stop-wearing-a-mask-says-cdc-director-rochelle-walensky> [accessed 2021-12-16]
22. Liu M, Thomadsen R, Yao S. Forecasting the Spread of COVID-19 under Different Reopening Strategies. *SSRN Journal* 2020 Jun 15:1-17. [doi: [10.2139/ssrn.3607977](https://doi.org/10.2139/ssrn.3607977)]
23. COVID Data Tracker. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/covid-data-tracker/#cases> [accessed 2021-12-16]
24. COVID data tracker: United States at a glance. Centers for Disease Control and Prevention. URL: <https://covid.cdc.gov/covid-data-tracker/#nowcasting> [accessed 2021-12-16]
25. Callaway E. Multitude of coronavirus variants found in the US - but the threat is unclear. *Nature* 2021 Mar 05;591(7849):190-190. [doi: [10.1038/d41586-021-00564-4](https://doi.org/10.1038/d41586-021-00564-4)] [Medline: [33674807](https://pubmed.ncbi.nlm.nih.gov/33674807/)]
26. Oehmke JF, Oehmke TB, Singh LN, Post LA. Dynamic Panel Estimate-Based Health Surveillance of SARS-CoV-2 Infection Rates to Inform Public Health Policy: Model Development and Validation. *J Med Internet Res* 2020 Sep 22;22(9):e20924 [FREE Full text] [doi: [10.2196/20924](https://doi.org/10.2196/20924)] [Medline: [32915762](https://pubmed.ncbi.nlm.nih.gov/32915762/)]
27. Oehmke TB, Post LA, Moss CB, Issa TZ, Boctor MJ, Welch SB, et al. Dynamic Panel Data Modeling and Surveillance of COVID-19 in Metropolitan Areas in the United States: Longitudinal Trend Analysis. *J Med Internet Res* 2021 Feb 09;23(2):e26081 [FREE Full text] [doi: [10.2196/26081](https://doi.org/10.2196/26081)] [Medline: [33481757](https://pubmed.ncbi.nlm.nih.gov/33481757/)]
28. Post LA, Issa TZ, Boctor MJ, Moss CB, Murphy RL, Ison MG, et al. Dynamic Public Health Surveillance to Track and Mitigate the US COVID-19 Epidemic: Longitudinal Trend Analysis Study. *J Med Internet Res* 2020 Dec 03;22(12):e24286 [FREE Full text] [doi: [10.2196/24286](https://doi.org/10.2196/24286)] [Medline: [33216726](https://pubmed.ncbi.nlm.nih.gov/33216726/)]
29. Post L, Marogi E, Moss CB, Murphy RL, Ison MG, Achenbach CJ, et al. SARS-CoV-2 Surveillance in the Middle East and North Africa: Longitudinal Trend Analysis. *J Med Internet Res* 2021 Jan 15;23(1):e25830 [FREE Full text] [doi: [10.2196/25830](https://doi.org/10.2196/25830)] [Medline: [33302252](https://pubmed.ncbi.nlm.nih.gov/33302252/)]
30. Post LA, Argaw ST, Jones C, Moss CB, Resnick D, Singh LN, et al. A SARS-CoV-2 Surveillance System in Sub-Saharan Africa: Modeling Study for Persistence and Transmission to Inform Policy. *J Med Internet Res* 2020 Nov 19;22(11):e24248 [FREE Full text] [doi: [10.2196/24248](https://doi.org/10.2196/24248)] [Medline: [33211026](https://pubmed.ncbi.nlm.nih.gov/33211026/)]
31. Post LA, Lin JS, Moss CB, Murphy RL, Ison MG, Achenbach CJ, et al. SARS-CoV-2 Wave Two Surveillance in East Asia and the Pacific: Longitudinal Trend Analysis. *J Med Internet Res* 2021 Feb 01;23(2):e25454 [FREE Full text] [doi: [10.2196/25454](https://doi.org/10.2196/25454)] [Medline: [33464207](https://pubmed.ncbi.nlm.nih.gov/33464207/)]

32. Global SARS-Cov-2 Surveillance Project Dashboard. Northwestern University Buehler Center for Health Policy and Economics. URL: <https://sites.northwestern.edu/covidglobalsurveillance/> [accessed 2021-12-16]
33. Week Numbers for 2021. Epoch Converter. URL: <https://www.epochconverter.com/weeks/2021> [accessed 2021-12-16]
34. Baltagi B. Econometric Analysis of Panel Data. In: Wiley. Chichester, UK: John Wiley & Sons Inc; 2009. [doi: [10.1007/978-3-030-53953-5](https://doi.org/10.1007/978-3-030-53953-5)]
35. Post LA, Benishay ET, Moss CB, Murphy RL, Achenbach CJ, Ison MG, et al. Surveillance Metrics of SARS-CoV-2 Transmission in Central Asia: Longitudinal Trend Analysis. *J Med Internet Res* 2021 Feb 03;23(2):e25799 [FREE Full text] [doi: [10.2196/25799](https://doi.org/10.2196/25799)] [Medline: [33475513](https://pubmed.ncbi.nlm.nih.gov/33475513/)]
36. Hansen LP. Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica* 1982 Jul;50(4):1029. [doi: [10.2307/1912775](https://doi.org/10.2307/1912775)]
37. Hurlin C. Panel Data Econometrics. Advanced Econometrics Internet University of Geneva. 2018. URL: <https://sites.google.com/view/christophe-hurlin/teaching-resources/panel-data-econometrics> [accessed 2021-12-16]
38. Bond SR. Dynamic panel data models: a guide to micro data methods and practice. *Portuguese Economic Journal* 2002 Aug 1;1(2):141-162. [doi: [10.1007/s10258-002-0009-9](https://doi.org/10.1007/s10258-002-0009-9)]
39. Vestal C. 10 States Have School Mask Mandates While 8 Forbid Them Internet. PEW Trusts. URL: <https://www.pewtrusts.org/en/research-and-analysis/blogs/stateline/2021/08/10/10-states-have-school-mask-mandates-while-8-forbid-them> [accessed 2021-12-16]
40. US Coronavirus Vaccine Tracker Internet. USAFacts. URL: <https://usafacts.org/visualizations/covid-vaccine-tracker-states/> [accessed 2021-12-16]
41. Musgrave J. COVID cases across Florida drop by 40% since last week; pandemic fatalities also down. *The Palm Beach Post*. URL: <https://www.palmbeachpost.com/story/news/2021/09/24/covid-florida-sees-decline-new-cases-deaths-number-vaccinated/5842798001/> [accessed 2021-12-16]
42. Delta Variant: What we know about the science. Centers for Disease Control. URL: <https://www.cdc.gov/coronavirus/2019-ncov/variants/delta-variant.html> [accessed 2021-12-16]

## Abbreviations

**ISO:** International Organization for Standardization

*Edited by T Sanchez; submitted 15.03.21; peer-reviewed by E Luellen, C Zhao; comments to author 08.07.21; revised version received 30.09.21; accepted 30.11.21; published 24.02.22.*

*Please cite as:*

*Oehmke TB, Moss CB, Oehmke JF*

*COVID-19 Surveillance Updates in US Metropolitan Areas: Dynamic Panel Data Modeling*

*JMIR Public Health Surveill* 2022;8(2):e28737

URL: <https://publichealth.jmir.org/2022/2/e28737>

doi: [10.2196/28737](https://doi.org/10.2196/28737)

PMID: [34882569](https://pubmed.ncbi.nlm.nih.gov/34882569/)

©Theresa B Oehmke, Charles B Moss, James F Oehmke. Originally published in *JMIR Public Health and Surveillance* (<https://publichealth.jmir.org>), 24.02.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

Original Paper

# Determining the Case Fatality Rate of COVID-19 in Italy: Novel Epidemiological Study

Mengqing Yan<sup>1,2</sup>; Wenjun Kang<sup>1,2</sup>; Zhifeng Guo<sup>1,2</sup>; Qi Wang<sup>1,3</sup>; Peizhong Peter Wang<sup>3,4</sup>; Yun Zhu<sup>5</sup>; Yongli Yang<sup>6</sup>; Wei Wang<sup>1,2</sup>

<sup>1</sup>Department of Occupational and Environmental Health, College of Public Health, Zhengzhou University, Zhengzhou, China

<sup>2</sup>The Key Laboratory of Nanomedicine and Health Inspection of Zhengzhou, Zhengzhou, China

<sup>3</sup>Center for New Immigrant Wellbeing, Markham, ON, Canada

<sup>4</sup>Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

<sup>5</sup>Department of Epidemiology and Biostatistics, Tianjin Medical University, Tianjin, China

<sup>6</sup>Department of Epidemiology and Biostatistics, College of Public Health, Zhengzhou University, Zhengzhou, China

**Corresponding Author:**

Wei Wang

Department of Occupational and Environmental Health

College of Public Health

Zhengzhou University

No.100, Science Avenue

Zhengzhou, 450001

China

Phone: 86 371 67781466

Email: [ww375@zzu.edu.cn](mailto:ww375@zzu.edu.cn)

## Abstract

**Background:** COVID-19, which emerged in December 2019, has spread rapidly around the world and has become a serious public health event endangering human life. With regard to COVID-19, there are still many unknowns, such as the exact case fatality rate (CFR).

**Objective:** The main objective of this study was to explore the value of the discharged CFR (DCFR) to make more accurate forecasts of epidemic trends of COVID-19 in Italy.

**Methods:** We retrieved the epidemiological data of COVID-19 in Italy published by the John Hopkins Coronavirus Resource Center. We then used the proportion of deaths to discharged cases (including deaths and recovered cases) to calculate the total DCFR (tDCFR), monthly DCFR (mDCFR), and stage DCFR (sDCFR). Furthermore, we analyzed the trend in the mDCFR between January and December 2020 using joinpoint regression analysis, used ArcGIS version 10.7 to visualize the spatial distribution of the epidemic CFR, and assigned different colors to each province based on the CFR or tDCFR.

**Results:** We calculated the numbers and obtained the new indices of the tDCFR and mDCFR for calculating the fatality rate. The results showed that the tDCFR and mDCFR fluctuated greatly from January to May. They first showed a rapid increase followed by a rapid decline after reaching the peak. The map showed that the provinces with a high tDCFR were Emilia-Romagna, Puglia, and Lombardia. The change trend of the mDCFR over time was divided into the following 2 stages: the first stage (from January to May) and the second stage (from June to December). With regard to worldwide COVID-19 statistics, among 6 selected countries, the United States had the highest tDCFR (4.26%), while the tDCFR of the remaining countries was between 0.98% and 2.72%.

**Conclusions:** We provide a new perspective for assessing the fatality of COVID-19 in Italy, which can use ever-changing data to calculate a more accurate CFR and scientifically predict the development trend of the epidemic.

(*JMIR Public Health Surveill* 2022;8(2):e32638) doi:[10.2196/32638](https://doi.org/10.2196/32638)

**KEYWORDS**

COVID-19; case fatality rate; discharged case fatality rate; new infectious diseases

## Introduction

COVID-19, which emerged in December 2019, has spread rapidly around the world and has become a serious public health event endangering human life [1,2]. The COVID-19 pandemic has strained or overwhelmed health systems across the world, with over 128 million COVID-19 cases and 2.8 million deaths as of March 31, 2021 [3]. Reports indicate that more than 200 countries have confirmed COVID-19 cases [4]. COVID-19 has high infection and mortality rates, and thus, it has become a pandemic [5,6]. Numerous studies have been conducted since the outbreak of COVID-19, and the reported findings have provided insights into the prevention and control of the disease [7,8]. With regard to COVID-19, there are still many unknowns, such as the exact case fatality rate (CFR) and the speed at which it spreads across communities. This lack of evidence complicates the design of appropriate response policies.

The European population was severely hit by the COVID-19 outbreak, particularly in Italy, which was the first European country to be affected by COVID-19 [9]. The first case of pneumonia due to SARS-CoV-2 in Italy, without a history of possible exposure abroad, was diagnosed in Lombardy (Northern Italy) on February 20, 2020. Within a few days, several COVID-19 cases were confirmed in the surrounding areas, and they included a substantial number of critically ill patients. Based on the number of cases and the advanced disease stage, it was estimated that community spread had been occurring since January 2020 [10]. Current statistics indicate that Italy is one of the countries severely affected by COVID-19-induced pneumonia [11]. By March 31, 2021, Italy had reported 3,584,899 confirmed COVID-19 cases and 109,346 deaths, ranking sixth worldwide in the number of deaths. Moreover, the CFR was approximately 3.05%, but it continues to fluctuate. A recent study analyzing over 70,000 COVID-19 patients in Italy revealed a wide variability in the CFR [12]. Therefore, the true incidence and CFR in Italy might be underestimated. There are some limitations in traditional mortality assessment methods. In order to solve these limitations and shortcomings, we introduce a new assessment method.

The purpose of our research was to determine how to make full use of the ever-changing authoritative data to make more accurate trend predictions of the epidemic. In order to more accurately assess the actual situation of the COVID-19 epidemic, we explored the use of the discharged CFR (DCFR) instead of the CFR to estimate the true situation and the use of the DCFR to make more accurate forecasts of epidemic trends. Public health institutions can use this method to calculate the dynamic fatality rate in different regions in real time, evaluate the medical conditions in different regions, and scientifically guide and reasonably arrange follow-up medical approaches. In addition, after entering the “turning point,” the overall situation of the entire epidemic can be predicted in advance based on the development data of the epidemic at that time and with reference to the real-time dynamic fatality rate data.

## Methods

### Data Collection and Characteristics

We obtained daily case reporting data from the Johns Hopkins University Center for Systems Science and Engineering (CSSE) [13]. The Center’s time-series data provided cumulative totals of COVID-19 cases and deaths by country [3]. The CSSE pools data from multiple sources, including the World Health Organization, the European Centre for Disease Prevention and Control, and the US Centers for Disease Control and Prevention, to produce daily country totals of confirmed cases and deaths. Data on COVID-19 in Italy involve daily updates from the Italian Ministry of Health managed by the Civil Protection Department.

### Statistical Analysis

#### Estimation of the DCFR

In this study, the DCFR and 95% CIs were estimated at the national and provincial levels of Italy. In addition, we selected several countries with a high number of confirmed cases, and the same method was performed to describe the death rate. ArcGIS software (version 10.7; Esri) was used to visualize the spatial distribution of the epidemic CFR and assign different colors to each province based on the CFR or total DCFR (tDCFR).

Notably, the DCFR includes the tDCFR, the monthly DCFR (mDCFR), and the stage DCFR (sDCFR). Discharged cases include deaths and recovered cases. The tDCFR is the proportion of deaths among discharged cases in the entire pandemic, the mDCFR is the proportion of deaths among discharged cases in each month, and the sDCFR is the proportion of deaths among discharged cases at each stage. The CFR, tDCFR, mDCFR, and sDCFR were calculated and analyzed as follows:

$$\text{CFR} = (\text{number of deaths attributed to COVID-19} / \text{number of total confirmed case of COVID-19}) \times 100\% \quad (1)$$

$$\text{tDCFR} = (\text{number of total deaths attributed to COVID-19} / [\text{number of total deaths attributed to COVID-19} + \text{number of total recovered cases}]) \times 100\% \quad (2)$$

$$\text{mDCFR} = (\text{number of monthly deaths attributed to COVID-19} / [\text{number of monthly deaths attributed to COVID-19} + \text{number of monthly recovered cases}]) \times 100\% \quad (3)$$

$$\text{sDCFR} = (\text{number of total deaths at each stage attributed to COVID-19} / [\text{number of total deaths at each stage attributed to COVID-19} + \text{number of total recovered cases at each stage}]) \times 100\% \quad (4)$$

The CFR, tDCFR, mDCFR, and sDCFR were estimated with 95% CIs. CI is an interval range containing population parameters constructed under a certain degree of confidence, and is widely used to estimate the range of population parameters [14,15]. We calculated the 95% CIs through a normal approximate method. The following formula was used to calculate the CIs [16,17]:

$$95\% \text{ CI} = (P - Z_{\alpha/2}S_p, P + Z_{\alpha/2}S_p) \text{ (5)}$$

where  $P = n/N$ ,  $Z_{\alpha/2} = 1.96$  for a 95% CI, and  $S_p = \sqrt{\frac{P(1-P)}{n}}$ .

### Trend Analysis

Furthermore, we analyzed the trend in mDCFR between January and December 2020 using joinpoint regression analysis. Joinpoint analysis used the Joinpoint Regression Program (version 3.4.3). The joinpoint is a point in the trend curve where a statistically significant change in trend over time is observed. The analysis requires a minimum number of 3 observations from a joinpoint to either end of the data and a minimum number of 4 observations between 2 joinpoints. From the joinpoint regression model, we extracted the monthly percentage change (MPC) and the average monthly percentage change (AMPC). The MPC is calculated for each significant trend from a piecewise log-linear model on the logarithm of the age-standardized rate versus the year, while the AMPC represents the average of MPC estimates per significant trend weighted by the corresponding trend length (number of years in the trend). The trend analysis using the joinpoint regression model was performed by the SEER\*Stat software [18] (Joinpoint Trend Analysis software from the Surveillance Research

Program of the US National Cancer Institute; version 4.8.0.1 [19]).

## Results

### COVID-19 Situation in Italy

By December 31, 2020, in Italy, the incidence rate in 2020 was 3.49%, the annual CFR was 3.52%, and the mortality rate was 123.11 per 100,000 cases. The first confirmed case of COVID-19 in Italy was reported on January 31, 2020, and the numbers showed a slight increase each month after February. The first peak was observed in March, but the number decreased from April to a minimum in July. In August 2020, the number of newly diagnosed patients rose to 21,677, followed by a sharp rise in November to reach the second peak of 922,124 new cases. With regard to COVID-19-related deaths, the number was relatively small in February, but the number increased rapidly from March, exceeding 10,000 deaths per month. The highest number of deaths (15,549) was recorded in April. However, the number declined in the following months, with the lowest (374) being recorded in August. Moreover, the number of new deaths in September and October remained low, but the number rose sharply in November (Table 1).

**Table 1.** The characteristics of COVID-19 in Italy in 2020.

Month	Total confirmed cases, n	Monthly confirmed new cases, n	Total recovered, n	Monthly recovered, n	Total deaths, n	Monthly deaths, n	CFR <sup>a</sup> (%), value (95% CI)	tDCFR <sup>b</sup> (%), value (95% CI)	mDCFR <sup>c</sup> (%), value (95% CI)
1	2	2	0	0	0	0	0	0	0
2	1128	1126	46	46	29	29	2.57 (-19.36 to 24.51)	38.67 (27.65 to 49.69)	38.67 (27.65 to 49.69)
3	105,792	104,664	15,729	15,683	12,428	12,399	11.75 (9.87 to 13.63)	44.14 (43.56 to 44.72)	44.15 (43.57 to 44.73)
4	205,463	99,671	75,945	60,216	27,967	15,539	13.61 (13.41 to 13.82)	26.91 (26.64 to 27.18)	20.51 (20.22 to 20.80)
5	232,997	27,534	157,507	81,562	33,415	5448	14.34 (14.19 to 14.49)	17.50 (17.33 to 17.67)	6.26 (6.10 to 6.42)
6	240,578	7581	190,248	32,741	34,767	1352	14.45 (14.31 to 14.59)	15.45 (15.30 to 15.60)	3.97 (3.76 to 4.17)
7	247,537	6959	199,974	9726	35,141	374	14.20 (14.06 to 14.34)	14.95 (14.8 to 15.09)	3.70 (3.33 to 4.07)
8	269,214	21,677	207,653	7679	35,483	342	13.18 (13.05 to 13.31)	14.59 (14.45 to 14.73)	4.26 (3.82 to 4.71)
9	314,861	45,647	227,704	20,051	35,894	411	11.40 (11.28 to 11.52)	13.62 (13.49 to 13.75)	2.01 (1.82 to 2.20)
10	679,430	364,569	289,426	61,722	38,618	2724	5.68 (5.60 to 5.76)	11.77 (11.66 to 11.88)	4.23 (4.07 to 4.38)
11	1,601,554	922,124	757,507	468,081	55,576	16,958	3.47 (3.43 to 3.51)	6.84 (6.78 to 6.89)	3.50 (3.44 to 3.55)
12	2,107,166	505,612	1,463,111	705,604	74,159	18,583	3.52 (3.49 to 3.55)	4.82 (4.79 to 4.86)	2.57 (2.53 to 2.60)

<sup>a</sup>CFR: case fatality rate.

<sup>b</sup>tDCFR: total discharged case fatality rate.

<sup>c</sup>mDCFR: monthly discharged case fatality rate.

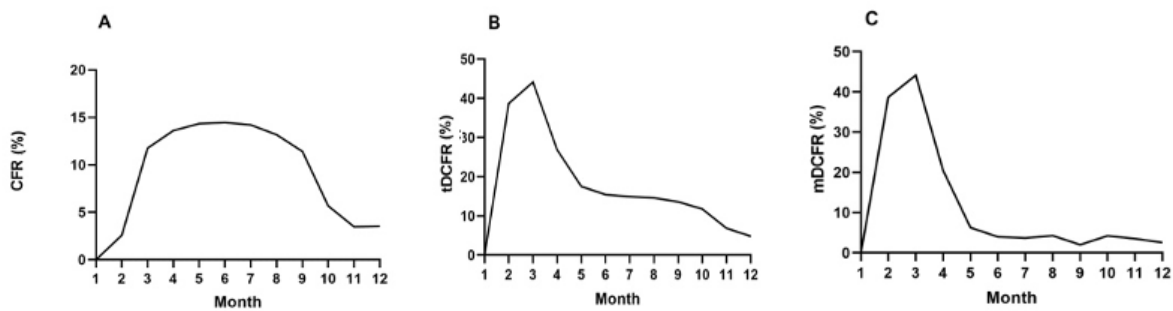
**CFR, tDCFR, and mDCFR**

Based on the daily report of COVID-19 cases in Italy, we used the new definition method described above to calculate the number of deaths and the number of recovered patients to obtain the new indices tDCFR and mDCFR for calculating the fatality rate. By December 31, 2020, the CFR in Italy was 3.52% and the tDCFR was 4.82% (Table 1).

Figure 1 shows the trends of the CFR, tDCFR, and mDCFR of COVID-19 in Italy from January 31, 2020, to December 31, 2020. The results indicated that the CFR first rose, then fell, and finally stabilized. The highest value of 14.45% was observed

in June, and it stabilized at about 3% after November (Figure 1A). On the other hand, the tDCFR and mDCFR fluctuated greatly. The results indicated that the tDCFR and mDCFR fluctuated greatly from January to May. They first showed a rapid increase followed by a rapid decline after reaching the peak (Figure 1B). The highest values of the tDCFR and mDCFR were observed in March (44.14% and 44.15%, respectively). After May, the tDCFR gradually decreased and finally stabilized at about 4%, while the mDCFR first declined after May and then showed a slight upward trend in August, and after falling to the lowest value of 2.01% in September, there was a slight increase in October, but it finally stabilized at around 2.5%.

**Figure 1.** The trends of the (A) case fatality rate (CFR), (B) total discharged case fatality rate (tDCFR), and (C) monthly discharged case fatality rate (mDCFR) of COVID-19 in Italy.



**Calculation of the CFR and tDCFR for 20 Provinces in Italy**

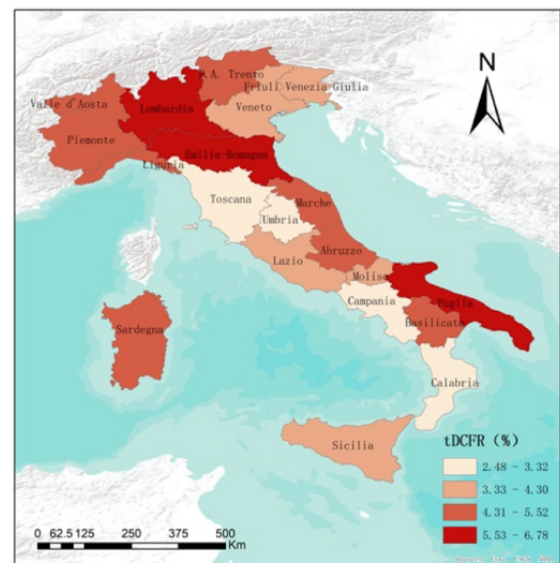
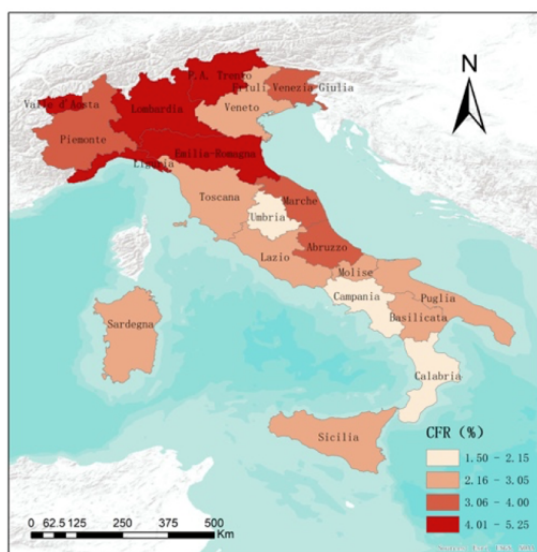
The provinces are divided into 4 levels according to the value of the CFR or tDCFR, which are represented by 4 colors. Among them, red indicates areas with high CFR or tDCFR values, while white indicates areas with low values. The map shows that the provinces with high CFR values were Lombardia, Valle d'Aosta, Liguria, Emilia-Romagna, P.A. Trento, and Piemonte. In

addition, there were 3 regions with high tDCFR values, including Emilia-Romagna, Puglia, and Lombardia. The results showed that the fatality rate of the epidemic, whether the CFR or tDCFR, was significantly higher in Northern Italy than in the Southern regions. However, there were some differences between the 2 evaluation indicators. For example, in the Puglia region, the tDCFR was 6.51% but the CFR was 2.72%, and in the Basilicata region, the tDCFR was 5.36% but the CFR was 2.36% (Figure 2).

**Figure 2.** Map showing the (A) case fatality rate (CFR) and (B) total discharged case fatality rate (tDCFR) in Italian provinces.

A

B

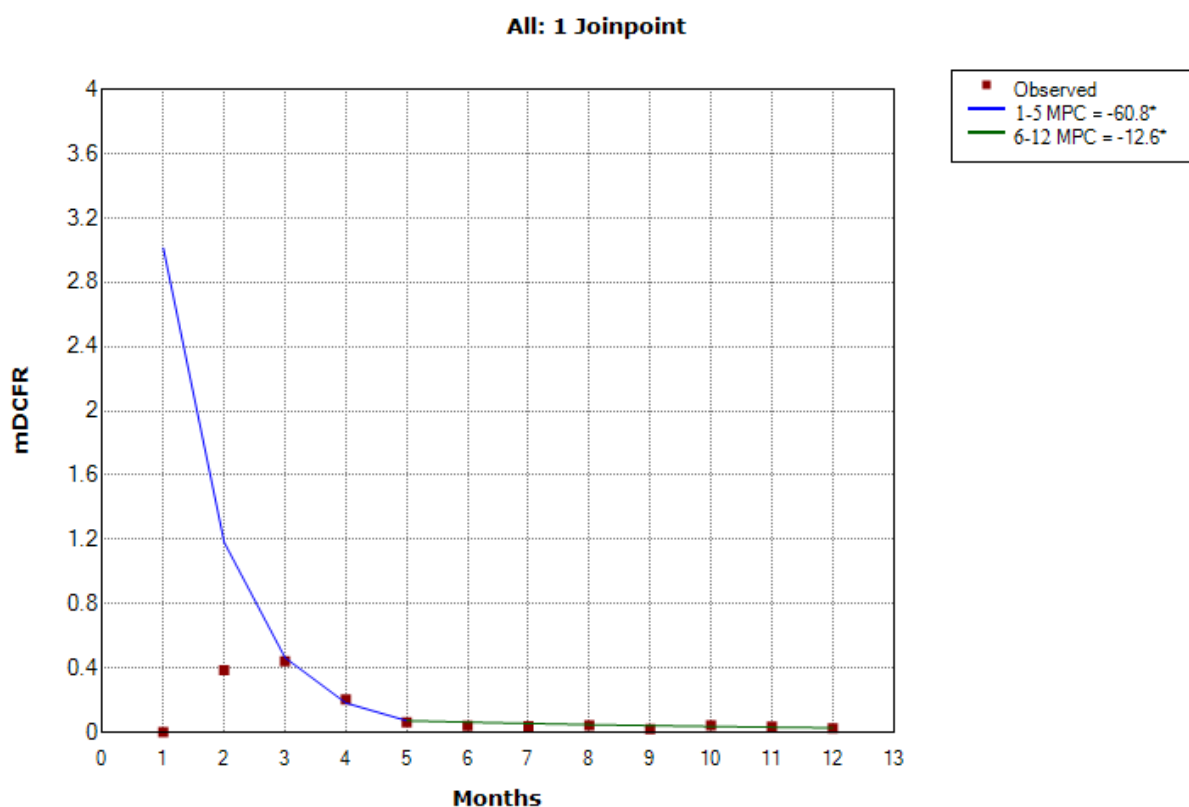


### Estimation of Pandemic Stages From the Calculated mDCFR and sDCFR

We used the Joinpoint Regression Program to analyze the change trend of the mDCFR over time, and determine a segmentation point, which was divided into 2 stages (Figure 3). The first stage was a sharp decline from January to May 2020, with the MPC being  $-60.8$  (95% CI  $-69.6$  to  $-49.5$ ;  $P < .001$ ). The second stage ranged from June to December 2020, with a sharp decrease in the declining stage, and the MPC was  $-12.6$

(95% CI  $-21.0$  to  $-3.5$ ;  $P = .02$ ). In comparison, the global AMPC for months 1 to 12 was  $-34.7$  (95% CI  $-40.6$  to  $-28.3$ ;  $P < .001$ ). The first phase represented the outbreak period, while the second phase represented the stable period. In addition, we calculated the sDCFR of each stage based on the number of deaths and the number of recovered cases in the different stages in Italy (Table 2). The results showed that the sDCFR values of the first and second stages were 17.50 (95% CI 17.33-17.67) and 3.03 (95% CI 3.00-3.06), respectively.

**Figure 3.** Estimation of pandemic stages in Italy using the monthly discharged case fatality rate (mDCFR). \*The monthly percentage change (MPC) is significantly different from 0 at the alpha level of .05.



Final Selected Model: 1 Joinpoint

**Table 2.** The stage discharged case fatality rate of COVID-19 at different stages in Italy in 2020.

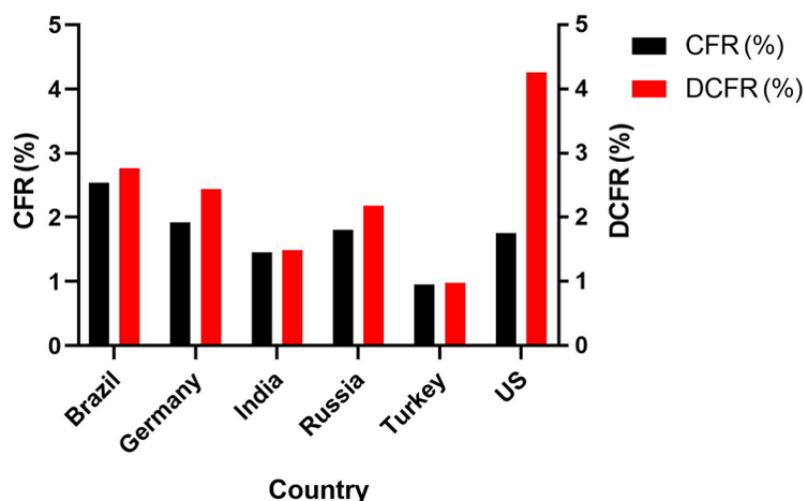
Stage	Time period	Deaths, n	Recovered cases, n	sDCFR <sup>a</sup> (%), value (95% CI)
The first stage	January 31 to May 31	33,415	157,507	17.50 (17.33-17.67)
The second stage	June 1 to December 31	40,744	1,305,604	3.03 (3.00-3.06)

<sup>a</sup>sDCFR: stage discharged case fatality rate.

### Worldwide COVID-19 Statistics

By December 31, 2020, the total number of confirmed COVID-19 cases in the world reached 83,521,859; the total number of deaths reached 1,824,666; and the total number of recovered cases reached 47,032,627, with a CFR of 2.18% and tDCFR of 3.73%. We also analyzed COVID-19 data from the

top 10 countries with the highest number of total confirmed cases. However, we were not able to calculate the DCFR for Spain, France, and the United Kingdom due to inaccurate data on the number of recovering patients. Among the other countries, the United States had the highest tDCFR (4.26%), while the tDCFR in the remaining countries ranged between 0.98% and 2.72% (Figure 4; Table 3).

**Figure 4.** Epidemic situation in different countries. CFR: case fatality rate; tDCFR: total discharged case fatality rate.**Table 3.** Worldwide COVID-19 statistics in 2020.

Country	Total confirmed cases, n	Total deaths, n	Total recovered cases, n	CFR <sup>a</sup> (%), value (95% CI)	tDCFR <sup>b</sup> (%), value (95% CI)
Brazil	7,675,973	194,949	6,848,844	2.54 (2.53-2.55)	2.77 (2.76-2.78)
Germany	1,760,520	33,791	1,350,708	1.92 (1.90-1.94)	2.44 (2.41-2.47)
India	10,266,674	148,738	9,860,280	1.45 (1.44-1.46)	1.49 (1.48-1.49)
Russia	3,127,347	56,271	2,527,722	1.80 (1.78-1.81)	2.18 (2.16-2.20)
Turkey	2,208,652	20,881	2,100,650	0.95 (0.93-0.96)	0.98 (0.97-1.00)
United States	20,099,363	352,078	7,919,313	1.75 (1.75-1.76)	4.26 (4.24-4.27)

<sup>a</sup>CFR: case fatality rate.

<sup>b</sup>tDCFR: total discharged case fatality rate.

## Discussion

### Reason for Proposing the Concept of the DCFR

The CFR is very crucial in the prediction of the epidemic trend because it reflects the degree of the danger that the epidemic poses [20]. The CFR is usually expressed in terms of the number of deaths per 100 treated patients [21]. However, in actual calculations, it is necessary to wait until the end of the epidemic when the total number of patients and total number of deaths can be counted, thereby providing accurate results. Therefore, it is difficult to use this formula to calculate the final CFR during the epidemic because the total number of patients and deaths is not conclusive. Moreover, during the epidemic period, the relevant departments sometimes replace the total death toll with the current number of deaths and replace the final confirmed number with the current confirmed number to get an approximate value of the CFR. However, this calculation method does not take into account the “number of patients cured and discharged.” Thus, it is not an appropriate measure of the medical level of hospitals. In addition, when the epidemic virus is unclear or the diagnosis ability is insufficient, the situation is more complicated and the number of patients is not accurate. Not only does the number of patients change daily, but also patients are divided into confirmed and suspected groups, and sometimes they will transform each other. Therefore, we propose a new method for calculating the CFR of ongoing infectious

diseases, which is referred to as the DCFR (including the tDCFR and mDCFR). The DCFR refers to the ratio of the cumulative number of deaths to the sum of the cumulative number of deaths and the cumulative number of cured patients at a certain time point. Calculating the DCFR has several advantages. It can show the dynamic fatality rate in different regions in real time, it can evaluate the medical conditions in different regions, and it can scientifically guide and help arrange follow-up medical matters reasonably. In addition, after entering the “turning point,” the overall situation of the entire epidemic can be predicted in advance based on the development data of the epidemic at that time and with reference to the real-time dynamic fatality rate data.

### Comparing the CFR and DCFR With Real Data From Italy

Italy is one of the countries severely affected by the current COVID-19 pandemic [22,23], and thus, researchers have focused on assessing the evolution of the Italian epidemic. Italy was the first European country to face a massive outbreak of COVID-19 cases [24,25], where the epidemic started in the northern region in February 2020 and quickly spread to all regions of the country [26]. The Italian COVID-19 epidemic has been unique from several points of view, with more than 2,000,000 confirmed cases and a fatality rate estimated to be one of the highest in the world [27]. The CFR is an indicator that reflects the severity of the disease. It should have a



relatively stable value under natural circumstances without other influencing factors. However, the CFR is affected by many factors in the early stage, where it first rises and then maintains a higher level for a period of time before falling. On the other hand, the CFR is stable in the later stage of a pandemic. Generally, the mortality rate in the early stage is relatively high due to an inadequate understanding of new infectious diseases in this stage, inadequate prevention and control measures, and unclear diagnosis. Therefore, the CFR cannot truly reflect the actual situation of the pandemic. In this study, we have used the trends of the DCFR and mDCFR to truly reflect the serious situation in the early stage of the COVID-19 epidemic in Italy. The obtained DCFR and mDCFR values showed that the COVID-19 fatality rate in Italy rose rapidly in February and reached a peak in March. Our results are consistent with the results of De Natale et al [22] who reported that COVID-19 was very serious in Italy in the early stage. After May, the fluctuation in the number of deaths gradually reduced the impact on the DCFR due to the large increase in the number of people discharged from the hospital, and finally, the DCFR remained stable. In a natural state without other influencing factors, the overall CFR should be relatively stable. In this study, we listed the numbers of confirmed cases, deaths, and recovered patients in each province in Italy, and used the numbers to calculate the CFR and tDCFR of each region. The results are presented on a map as shown in Figure 2. Some findings deserve attention from relevant departments of epidemic prevention and control. For example, the CFR and tDCFR values were both high in the Valle d'Aosta area despite the number of confirmed cases being low. In addition, the calculated DCFR was high in the Puglia area despite the CFR being very low. Therefore, the severity of COVID-19 in these areas cannot be ignored. This nonassociated result indicates that researchers should consider the CFR and tDCFR together with the determined number when assessing the severity of COVID-19. The DCFR may have a large error at the beginning of the pandemic, when the number of discharged cases is small, but the error will decrease as the number of discharged cases increases.

### Practical Applications of the DCFR

The role of public health institutions is to dynamically analyze the epidemic trend from the massive data after an outbreak, and provide government departments with both forward-looking and accurate professional prevention and control recommendations. Before the end of the epidemic, the epidemic management department and the general public are usually more concerned about the CFR of the epidemic, how many people

are diagnosed with the infection, and how long the epidemic lasts. Only by clarifying these issues as early as possible can we accurately grasp the trend of the epidemic, and take targeted control measures to curb the rapid spread of the epidemic and avoid the generation of rumors and panic.

Prediction of the COVID-19 trend and control effects in the earlier stage is very important. This study has shown that the DCFR can be used as a predictor for the trend, and it can provide additional information for controlling diseases. Briefly, we proposed the value of the DCFR in describing emerging infectious diseases, divided the pandemic stages based on the mDCFR, and used the mDCFR to evaluate the control effect at different stages. We collected data from January 22 to December 31, and used joinpoint regression analysis to divide the data into 2 stages (Figure 3). Our DCFR results showed a gradually decreasing trend in the 2 stages (Table 2), which can be attributed to the characteristics of early cases [28], the improvement of diagnosis and treatment measures [3], government interventions, and the increase in the proportion of less dangerous viruses [18]. This suggests that it is reasonable to use the DCFR to predict the pandemic trend.

### Limitations

Although this study proposed the DCFR index first, some limitations should be addressed. We were not able to conduct some critical analyses due to unavailability of full data access in some countries. In our next study, we will try to find complete data from some key countries and use the data to calculate the indicators of the DCFR, such as age and gender. In addition, we will compare them with the findings in this study, classify the DCFR in detail in a larger database, and conduct long-term analysis.

### Conclusions

The DCFR can use ever-changing data to calculate a more accurate CFR, divide the pandemic stages of new infectious diseases, and analyze the dynamic trend. Our results suggest that the DCFR can be used as one of the pandemic control indicators. The results showed that the DCFR was high in the early stage of the COVID-19 outbreak in Italy, but it then decreased to a stable level. This suggests that other countries may also adopt the DCFR as one of the indicators of pandemic control. Furthermore, the DCFR may have potential application value in many emerging infectious diseases, such as Middle East respiratory syndrome and severe acute respiratory syndrome.

---

### Conflicts of Interest

None declared.

---

### References

1. Zhou L, Liu J, Dong X, McGoogan JM, Wu Z. COVID-19 seeding time and doubling time model: an early epidemic risk assessment tool. *Infect Dis Poverty* 2020 Jun 23;9(1):76 [FREE Full text] [doi: [10.1186/s40249-020-00685-4](https://doi.org/10.1186/s40249-020-00685-4)] [Medline: [32576256](https://pubmed.ncbi.nlm.nih.gov/32576256/)]

2. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, China Novel Coronavirus Investigating Research Team. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* 2020 Feb 20;382(8):727-733 [FREE Full text] [doi: [10.1056/NEJMoa2001017](https://doi.org/10.1056/NEJMoa2001017)] [Medline: [31978945](https://pubmed.ncbi.nlm.nih.gov/31978945/)]
3. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* 2020 May 19;20(5):533-534. [doi: [10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1)] [Medline: [32087114](https://pubmed.ncbi.nlm.nih.gov/32087114/)]
4. Zhang G, Pan H, Hu X, He S, Chen Y, Wei C, et al. The role of isolation rooms, facemasks and intensified hand hygiene in the prevention of nosocomial COVID-19 transmission in a pulmonary clinical setting. *Infect Dis Poverty* 2020 Jul 23;9(1):104 [FREE Full text] [doi: [10.1186/s40249-020-00725-z](https://doi.org/10.1186/s40249-020-00725-z)] [Medline: [32703281](https://pubmed.ncbi.nlm.nih.gov/32703281/)]
5. Mori H, Schol J, Geeraerts A, Huang I, Jandee S, Gonlachanvit S, et al. The Impact of COVID-19 on Gastrointestinal Motility Testing in Asia and Europe. *J Clin Med* 2020 Oct 01;9(10):3189 [FREE Full text] [doi: [10.3390/jcm9103189](https://doi.org/10.3390/jcm9103189)] [Medline: [33019626](https://pubmed.ncbi.nlm.nih.gov/33019626/)]
6. Tian W, Zhang N, Jin R, Feng Y, Wang S, Gao S, et al. Immune suppression in the early stage of COVID-19 disease. *Nat Commun* 2020 Nov 17;11(1):5859 [FREE Full text] [doi: [10.1038/s41467-020-19706-9](https://doi.org/10.1038/s41467-020-19706-9)] [Medline: [33203833](https://pubmed.ncbi.nlm.nih.gov/33203833/)]
7. Barda N, Riesel D, Akriv A, Levy J, Finkel U, Yona G, et al. Developing a COVID-19 mortality risk prediction model when individual-level data are not available. *Nat Commun* 2020 Sep 07;11(1):4439 [FREE Full text] [doi: [10.1038/s41467-020-18297-9](https://doi.org/10.1038/s41467-020-18297-9)] [Medline: [32895375](https://pubmed.ncbi.nlm.nih.gov/32895375/)]
8. Soo RJJ, Chiew CJ, Ma S, Pung R, Lee V. Decreased Influenza Incidence under COVID-19 Control Measures, Singapore. *Emerg Infect Dis* 2020 Aug;26(8):1933-1935 [FREE Full text] [doi: [10.3201/eid2608.201229](https://doi.org/10.3201/eid2608.201229)] [Medline: [32339092](https://pubmed.ncbi.nlm.nih.gov/32339092/)]
9. Buonomo B, Della Marca R. Effects of information-induced behavioural changes during the COVID-19 lockdowns: the case of Italy. *R Soc Open Sci* 2020 Oct;7(10):201635 [FREE Full text] [doi: [10.1098/rsos.201635](https://doi.org/10.1098/rsos.201635)] [Medline: [33204488](https://pubmed.ncbi.nlm.nih.gov/33204488/)]
10. Torri E, Sbrogiò LG, Rosa ED, Cinquetti S, Francia F, Ferro A. Italian Public Health Response to the COVID-19 Pandemic: Case Report from the Field, Insights and Challenges for the Department of Prevention. *Int J Environ Res Public Health* 2020 May 22;17(10):3666 [FREE Full text] [doi: [10.3390/ijerph17103666](https://doi.org/10.3390/ijerph17103666)] [Medline: [32456072](https://pubmed.ncbi.nlm.nih.gov/32456072/)]
11. de Girolamo G, Bellelli G, Bianchetti A, Starace F, Zanetti O, Zarbo C, et al. Older People Living in Long-Term Care Facilities and Mortality Rates During the COVID-19 Pandemic in Italy: Preliminary Epidemiological Data and Lessons to Learn. *Front Psychiatry* 2020;11:586524 [FREE Full text] [doi: [10.3389/fpsy.2020.586524](https://doi.org/10.3389/fpsy.2020.586524)] [Medline: [33173526](https://pubmed.ncbi.nlm.nih.gov/33173526/)]
12. Gadi N, Wu SC, Spihlman AP, Moulton VR. What's Sex Got to Do With COVID-19? Gender-Based Differences in the Host Immune Response to Coronaviruses. *Front Immunol* 2020;11:2147 [FREE Full text] [doi: [10.3389/fimmu.2020.02147](https://doi.org/10.3389/fimmu.2020.02147)] [Medline: [32983176](https://pubmed.ncbi.nlm.nih.gov/32983176/)]
13. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). Johns Hopkins University. URL: <https://coronavirus.jhu.edu/map.html> [accessed 2022-01-25]
14. Sim J, Reid N. Statistical inference by confidence intervals: issues of interpretation and utilization. *Phys Ther* 1999 Feb;79(2):186-195. [Medline: [10029058](https://pubmed.ncbi.nlm.nih.gov/10029058/)]
15. Nakagawa S, Cuthill IC. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol Rev Camb Philos Soc* 2007 Nov;82(4):591-605. [doi: [10.1111/j.1469-185X.2007.00027.x](https://doi.org/10.1111/j.1469-185X.2007.00027.x)] [Medline: [17944619](https://pubmed.ncbi.nlm.nih.gov/17944619/)]
16. Simon R. Confidence intervals for reporting results of clinical trials. *Ann Intern Med* 1986 Sep;105(3):429-435. [doi: [10.7326/0003-4819-105-3-429](https://doi.org/10.7326/0003-4819-105-3-429)] [Medline: [3740683](https://pubmed.ncbi.nlm.nih.gov/3740683/)]
17. Moreno-Küstner B, Del Campo-Ávila J, Ruíz-Ibáñez A, Martínez-García AI, Castro-Zamudio S, Ramos-Jiménez G, et al. Epidemiology of Suicidal Behavior in Malaga (Spain): An Approach From the Prehospital Emergency Service. *Front Psychiatry* 2019;10:111 [FREE Full text] [doi: [10.3389/fpsy.2019.00111](https://doi.org/10.3389/fpsy.2019.00111)] [Medline: [30918488](https://pubmed.ncbi.nlm.nih.gov/30918488/)]
18. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. *Stat Med* 2000 Feb 15;19(3):335-351. [doi: [10.1002/\(sici\)1097-0258\(20000215\)19:3<335::aid-sim336>3.0.co;2-z](https://doi.org/10.1002/(sici)1097-0258(20000215)19:3<335::aid-sim336>3.0.co;2-z)] [Medline: [10649300](https://pubmed.ncbi.nlm.nih.gov/10649300/)]
19. Joinpoint Trend Analysis Software. National Cancer Institute. URL: <https://surveillance.cancer.gov/joinpoint/> [accessed 2022-02-04]
20. Yang S, Cao P, Du P, Wu Z, Zhuang Z, Yang L, et al. Early estimation of the case fatality rate of COVID-19 in mainland China: a data-driven analysis. *Ann Transl Med* 2020 Feb;8(4):128 [FREE Full text] [doi: [10.21037/atm.2020.02.66](https://doi.org/10.21037/atm.2020.02.66)] [Medline: [32175421](https://pubmed.ncbi.nlm.nih.gov/32175421/)]
21. Giangreco G. Case fatality rate analysis of Italian COVID-19 outbreak. *J Med Virol* 2020 Jul;92(7):919-923 [FREE Full text] [doi: [10.1002/jmv.25894](https://doi.org/10.1002/jmv.25894)] [Medline: [32297983](https://pubmed.ncbi.nlm.nih.gov/32297983/)]
22. De Natale G, Ricciardi V, De Luca G, De Natale D, Di Meglio G, Ferragamo A, et al. The COVID-19 Infection in Italy: A Statistical Study of an Abnormally Severe Disease. *J Clin Med* 2020 May 21;9(5):1564 [FREE Full text] [doi: [10.3390/jcm9051564](https://doi.org/10.3390/jcm9051564)] [Medline: [32455807](https://pubmed.ncbi.nlm.nih.gov/32455807/)]
23. Scafetta N. Distribution of the SARS-CoV-2 Pandemic and Its Monthly Forecast Based on Seasonal Climate Patterns. *Int J Environ Res Public Health* 2020 May 17;17(10):3493 [FREE Full text] [doi: [10.3390/ijerph17103493](https://doi.org/10.3390/ijerph17103493)] [Medline: [32429517](https://pubmed.ncbi.nlm.nih.gov/32429517/)]
24. Remuzzi A, Remuzzi G. COVID-19 and Italy: what next? *The Lancet* 2020 Apr 11;395(10231):1225-1228 [FREE Full text] [doi: [10.1016/S0140-6736\(20\)30627-9](https://doi.org/10.1016/S0140-6736(20)30627-9)] [Medline: [32178769](https://pubmed.ncbi.nlm.nih.gov/32178769/)]

25. Dowd JB, Andriano L, Brazel DM, Rotondi V, Block P, Ding X, et al. Demographic science aids in understanding the spread and fatality rates of COVID-19. *Proc Natl Acad Sci U S A* 2020 May 05;117(18):9696-9698 [FREE Full text] [doi: [10.1073/pnas.2004911117](https://doi.org/10.1073/pnas.2004911117)] [Medline: [32300018](https://pubmed.ncbi.nlm.nih.gov/32300018/)]
26. Milano M, Cannataro M. Statistical and Network-Based Analysis of Italian COVID-19 Data: Communities Detection and Temporal Evolution. *Int J Environ Res Public Health* 2020 Jun 12;17(12):4182 [FREE Full text] [doi: [10.3390/ijerph17124182](https://doi.org/10.3390/ijerph17124182)] [Medline: [32545441](https://pubmed.ncbi.nlm.nih.gov/32545441/)]
27. Michelozzi P, de'Donato F, Scortichini M, Pezzotti P, Stafoggia M, De Sario M, et al. Temporal dynamics in total excess mortality and COVID-19 deaths in Italian cities. *BMC Public Health* 2020 Aug 14;20(1):1238 [FREE Full text] [doi: [10.1186/s12889-020-09335-8](https://doi.org/10.1186/s12889-020-09335-8)] [Medline: [32795276](https://pubmed.ncbi.nlm.nih.gov/32795276/)]
28. Rubino S, Kelvin N, Bermejo-Martin JF, Kelvin D. As COVID-19 cases, deaths and fatality rates surge in Italy, underlying causes require investigation. *J Infect Dev Ctries* 2020 Mar 31;14(3):265-267 [FREE Full text] [doi: [10.3855/jidc.12734](https://doi.org/10.3855/jidc.12734)] [Medline: [32235086](https://pubmed.ncbi.nlm.nih.gov/32235086/)]

## Abbreviations

**AMPC:** average monthly percentage change  
**CFR:** case fatality rate  
**CSSE:** Center for Systems Science and Engineering  
**DCFR:** discharged case fatality rate  
**mDCFR:** monthly discharged case fatality rate  
**MPC:** monthly percentage change  
**sDCFR:** stage discharged case fatality rate  
**tDCFR:** total discharged case fatality rate

*Edited by T Sanchez; submitted 05.08.21; peer-reviewed by H Akram, S Pesälä; comments to author 28.10.21; revised version received 18.11.21; accepted 28.12.21; published 10.02.22.*

*Please cite as:*

*Yan M, Kang W, Guo Z, Wang Q, Wang PP, Zhu Y, Yang Y, Wang W  
Determining the Case Fatality Rate of COVID-19 in Italy: Novel Epidemiological Study  
JMIR Public Health Surveill 2022;8(2):e32638  
URL: <https://publichealth.jmir.org/2022/2/e32638>  
doi: [10.2196/32638](https://doi.org/10.2196/32638)  
PMID: [34963659](https://pubmed.ncbi.nlm.nih.gov/34963659/)*

©Mengqing Yan, Wenjun Kang, Zhifeng Guo, Qi Wang, Peizhong Peter Wang, Yun Zhu, Yongli Yang, Wei Wang. Originally published in *JMIR Public Health and Surveillance* (<https://publichealth.jmir.org>), 10.02.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

---

Publisher:  
JMIR Publications  
130 Queens Quay East.  
Toronto, ON, M5A 3Y5  
Phone: (+1) 416-583-2040  
Email: [support@jmir.org](mailto:support@jmir.org)

---

<https://www.jmirpublications.com/>