



# Innovations in public health surveillance: An overview of novel use of data and analytic methods

Heather Rilkoff<sup>1\*</sup>, Shannon Struck<sup>2</sup>, Chelsea Ziegler<sup>3</sup>, Laura Faye<sup>1</sup>, Dana Paquette<sup>4</sup>, David Buckeridge<sup>5,6</sup>

## Abstract

Innovative data sources and methods for public health surveillance (PHS) have evolved rapidly over the past 10 years, suggesting the need for a closer look at the scientific maturity, feasibility, and utility of use in real-world situations. This article provides an overview of recent innovations in PHS, including data from social media, internet search engines, the Internet of Things (IoT), wastewater surveillance, participatory surveillance, artificial intelligence (AI), and nowcasting.

Examples identified suggest that novel data sources and analytic methods have the potential to strengthen PHS by improving disease estimates, promoting early warning for disease outbreaks, and generating additional and/or more timely information for public health action. For example, wastewater surveillance has re-emerged as a practical tool for early detection of the coronavirus disease 2019 (COVID-19) and other pathogens, and AI is increasingly used to process large amounts of digital data. Challenges to implementing novel methods include lack of scientific maturity, limited examples of implementation in real-world public health settings, privacy and security risks, and health equity implications. Improving data governance, developing clear policies for the use of AI technologies, and public health workforce development are important next steps towards advancing the use of innovation in PHS.

**Suggested citation:** Rilkoff H, Struck S, Ziegler C, Faye L, Paquette D, Buckeridge D. Innovations in public health surveillance: An overview of novel use of data and analytic methods. *Can Commun Dis Rep* 2024;50(3/4):93–101. <https://doi.org/10.14745/ccdr.v50i34a02>

**Keywords:** public health surveillance, innovative methods, novel data, artificial intelligence, wastewater surveillance, nowcasting

## Introduction

Public health surveillance (PHS) is the ongoing, systematic collection, analysis, and interpretation of data, followed by the dissemination of information, for the purpose of guiding actions to prevent and control diseases or improve population health (1–3). Traditionally, PHS was conducted with a limited number of data sources from public health information systems, health care, and laboratory information systems, as well as questionnaire-based surveys, which often require substantial resources and time to process, analyze, and disseminate.

The digitization of health care and other sectors has reduced the time lag, cost and burden associated with conducting PHS, and enabled exploration of other sources of data to augment traditional sources (4). In addition, artificial intelligence (AI) has

seen major advances over the past decade. Artificial intelligence-enabled methodologies that efficiently process large amounts of structured and unstructured data are increasingly used in PHS (5–7).

Many of these data sources and AI methods were used during the coronavirus disease 2019 (COVID-19) pandemic, where timely and complete information was crucial to understanding and responding to evolving pandemic risks (4). The rapid development of these innovative surveillance methods and use of novel data sources suggests the need to take a closer look at the scientific maturity, as well as the feasibility and utility of their use, in real-world applications (5,6,8). The objective of this paper is to highlight examples of the application of innovative methods

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).



## Affiliations

<sup>1</sup> Data, Surveillance and Foresight Branch, Public Health Agency of Canada, Toronto, ON

<sup>2</sup> Data, Surveillance and Foresight Branch, Public Health Agency of Canada, Winnipeg, MB

<sup>3</sup> Data, Surveillance and Foresight Branch, Public Health Agency of Canada, Calgary, AB

<sup>4</sup> Data, Surveillance and Foresight Branch, Public Health Agency of Canada, Ottawa, ON

<sup>5</sup> Data, Surveillance and Foresight Branch, Public Health Agency of Canada, Montréal, QC

<sup>6</sup> School of Population and Global Health, Department of Epidemiology and Biostatistics, McGill University, Montréal, QC

## \*Correspondence:

[heather.rilkoff@phac-aspc.gc.ca](mailto:heather.rilkoff@phac-aspc.gc.ca)



to PHS and provide insights for public health authorities on the potential benefits, risks, and challenges of using non-traditional data sources and methods in PHS.

This article provides an overview of PHS innovations in data and analytic methods published in the past five years, including any evidence of their application to real-world settings, ethical issues, and known health equity implications. Each innovation is described, including its level of scientific maturity and, where available, any evidence of its impact on surveillance practice or public health action. The results section starts by exploring novel data sources that have been applied to PHS, highlighting successful examples of their application to provide timely, accurate and reliable information to support public health action. It then focuses on innovative methods that have been developed to analyze surveillance data, including the development of AI to support the integration and analysis of large and/or non-traditional data sources and the application of advanced analytic methods to improve nowcasting of information.

## Methods

### Approach

This overview defines the term “innovative surveillance” broadly as the use of non-traditional data sources and/or analytic methods to detect and understand health events and determinants. The primary focus was on data sources and analytic methods; this overview does not provide detailed discussion of other components of the surveillance process (e.g., dissemination or evaluation strategies).

Relevant topic areas were identified for inclusion in this article by searching PubMed, Embase, Global Health, and Scopus in the spring of 2023. A detailed search strategy, developed with the support of a librarian, was restricted to peer reviewed articles published between January 1, 2013, and February 23, 2023, from member countries of the Organisation for Economic Co-operation and Development (OECD) and China, in English language only. Hand searching provided additional sources.

Results of the literature search were screened for relevance via title and abstract search and grouped into topic areas. Final selection of articles within each topic area was restricted to the past five years (January 1, 2018, to February 23, 2023) to ensure that articles were more reflective of current technological and methodological innovations. As the search yielded a large number of articles on analytic methods, decisions were made by the research team to exclude certain broad analytic topic areas (such as innovations in biostatistics, laboratory, or geospatial analytic methods), and focus on nowcasting and artificial intelligence, two areas that have been adopted by public health from other disciplines.

The authors focused this overview on a subset of articles that met the definition of “innovative surveillance”, discussed steps taken to evaluate or validate the method or data source(s), described potential or actual improvements to the PHS system, and, where possible, showed application to real-world public health practice.

## Results

### Novel data sources and their applications

#### Overview of novel data sources

The rise of digital technologies has made new data sources available for disease surveillance. Commonly used digital data sources include social media and aggregate search query data, where initial surveillance applications date from the early 2000s, as well as participatory surveillance methods, such as repeated cross-sectional online surveys and crowdsourcing of photos or sample submissions (9). More recently, PHS applications of other digital technologies are being explored, such as mobility data and the Internet of Things (IoT), which includes wearable devices and other physical objects that connect and exchange data via the Internet (8). Digital data sources may have the potential to provide more timely information and capture populations that may not seek health care; although possible to use as an independent source of information, they are generally considered to be complementary to traditional surveillance data (9).

#### Social media and web-search data

Social media (e.g., Twitter/X) and web search (e.g., Google Trends) data have been used to support disease surveillance as a source of data for nowcasting, situational awareness, and outbreak detection (9). A recent systematic review focusing on communicable disease surveillance noted that the majority of included studies used data from Twitter/X, and that studies that used Twitter/X data showed higher overall reliability and validity than studies using data from other social media platforms (10). The review also noted that the majority of studies focused on influenza surveillance, and that additional research was needed to assess the effectiveness of social media for other disease areas (10). Other examples of the use of social media and/or web search data included retrospective analyses to evaluate the potential of these sources to predict cases of sexually transmitted and blood-borne infections (STBBIs) (11), prioritizing restaurant inspections based on foodborne outbreak information (12), drug utilization estimates (13), and early warning systems for e-cigarette/vaping-related lung injuries (14) and COVID-19 outbreaks (15).



One of the challenges with the use of digital media is the need to collect and process large quantities of information, either through manual monitoring or automation (16). The European Centre for Disease Prevention and Control (ECDC) released *epitweetr*, an R-based software library that collects, aggregates, detects, and disseminates information for early detection of public health threats using Twitter/X. An evaluation of the tool noted greater timeliness when compared to manual review (16). Artificial intelligence methods such as natural language processing, described later in this paper, are also increasingly being used to process and analyze digital information sources.

While the utility of social media and web search data for disease surveillance has been explored for nearly two decades, the validity, reliability, and stability of these data continue to present challenges to developing standardized approaches to using this information (9). For example, changes to the query algorithms of search engines, the use of different language styles, confounding search terms, and demographic biases in terms of who uses digital technologies, may impact the quality of information from these sources for PHS (9,17). A recent systematic scoping review also noted that most studies on digital surveillance did not utilize their results for public health action, and that more rigorous methods were needed to operationalize this information for public health decision-making (17). Surveillance platforms that combine social media, web search, and healthcare data may improve the accuracy of results (9,18).

#### **Participatory surveillance data**

Participatory surveillance involves the voluntary recruitment and engagement of members of the public to participate in repeated surveys or other crowdsourcing methods (9). This approach is sometimes used as an augment to traditional disease surveillance, to capture information in a timelier way, and to capture populations that may not seek health care for testing and diagnosis (8). Examples include Flu Near You in the United States, InfluenzaNet in Europe (9), and FluWatchers in Canada (19). Community surveillance using self-collected specimens has also been implemented and has enabled rapid assessment of community-level burden of influenza (20). Additionally, studies have explored participatory syndromic surveillance using social media and newspaper reports as a source of information during the COVID-19 pandemic that may be timelier and more accessible than official public health case reports (21,22).

Outside of respiratory pathogens, recent studies suggest current use of participatory approaches to support surveillance of potential disease carrying vectors or vector-borne disease. For example, platforms such as iNaturalist, eTick.ca, and Mosquito Alert use crowdsourced photos to identify the distribution and seasonal trends of specific species of ticks and mosquitos (23–26), and initiatives such as tickMAP in New York state used community-submitted tick specimens to track the emergence of tick-borne pathogens in near real time (27).

Participatory surveillance may be applied in a way that enables participation from equity-deserving populations that may otherwise be excluded from traditional surveillance systems. For example, in a rural Appalachian community, participatory surveillance via an online or phone-based symptom self-checking tool was used to identify at-risk individuals who may otherwise have not sought health care and link them to resources from the local health department (28). However, certain populations may be less likely to participate in participatory surveillance, including males, younger and older age groups (29), and those with lower income and education (9). This may introduce bias and potential health equity issues, particularly if groups that are more likely to experience illness are excluded.

#### **New digital data sources**

The use of digital data sources, such as mobile technologies, IoT and wearables, represent emerging areas for further exploration. For example, mobility data was used to explore the impact of COVID-19 and government policy on travel patterns. Health inequities were also noted, as socially disadvantaged populations were often unable to benefit from stay-at-home orders (30,31).

Wearable devices, such as smartwatches, have been used to collect individual-level data on variables linked to viral infection, such as resting heart rate, sleep, and mobility (32,33). As an example, a study noted that wearable technologies may improve nowcasting of influenza-like illness (ILI) rates in the United States (33). Various applications of IoT have emerged in the past few years. In one study, researchers placed thermal sensors and microphones in hospital waiting rooms to monitor coughing, which was then used to support ILI surveillance (34).

New digital data sources from mobility, wearables, and IoT represent an emerging field that requires greater evaluation and assessment (8,32), including careful consideration of privacy and ethical concerns (35). Like other digital data sources, these sources involve self-selected populations and exclude groups who do not have access to digital technologies. Privacy issues have also emerged with the use of new digital technologies and social media data; data ownership and the right to share data and use the data for secondary purposes may differ among the public sector (e.g., government), private sector (e.g., Twitter/X), and geopolitical jurisdictions (9,25). The need for upgraded infrastructure and investment to support the integration and analysis of information generated from new technologies may also present substantial barriers (8,36).

#### **Wastewater**

Wastewater surveillance (WWS) has evolved as a data source that now supports global surveillance of infectious diseases in a manner that is independent of health-seeking behaviour and healthcare system access (37,38). When coupled with small area socio-demographic data, WWS has the potential to forewarn and confirm clinical trends, address health inequities, fill reporting gaps due to waning clinical testing, and provide purpose-built



sentinel surveillance of communities with higher-risk profiles for specific pathogens (38–42). The deluge of WWS data during the COVID-19 pandemic led to novel analytic methods to help inform public health action. These include sophisticated machine learning algorithms that were applied to estimate sewage flow rates to allow for data normalization (43), and the application of simple statistical methods that were then tested to identify early warning signals in a user-friendly manner (44,45). New methods developed for WWS during the pandemic were validated by comparing wastewater signals to clinical case data and COVID-like illness syndromic data (38,40,41,45,46). Innovations in WWS have also benefited from other novel data linkages. In a recent study in Iceland, wastewater signals were compared with driving under the influence records to help distinguish trends of recreational drug use from increased drug dependencies, the latter of which may require enhanced public health action (47).

Wastewater surveillance of COVID-19, other infectious pathogens, and illicit substances, has identified limitations of this approach including the inability to distinguish reasons for signal increases/decreases, the degradation of the pathogen/substance in the wastewater before testing is performed, changing population denominators, and non-standardized sampling methods (47–49). Wastewater surveillance is also limited by the epidemiological indicators it can provide (i.e., incidence and prevalence) and the population it can monitor (e.g., includes only those in the sewer shed of a wastewater treatment facility) from the WWS data alone (45–49).

## Innovative analytic methods

### Artificial intelligence

Artificial intelligence, which includes natural language processing (NLP), machine learning, and deep learning, can integrate, process, and interpret multiple sources of information more efficiently and more consistently than humans (50). The recent growth in the use of AI-based technologies that can process unstructured text data has enabled the use of novel data sources, including those discussed in the previous sections, to be leveraged more effectively (7). Artificial intelligence has enormous potential to improve PHS, as it is capable of processing large amounts of data to identify anomalies that may pose a threat to public health (7), however, it is still an emerging field in which more real-world evaluations are needed. Some of the published innovations using AI for PHS still reside within academic collaborations. One such study from the Yale School of Medicine used NLP, which applies AI methods to the interpretation of human language, to provide real-time monitoring of population health by identifying symptoms mentioned on social media platforms (51).

Machine learning identifies complex patterns in data for classification and prediction (50). In New York City (NYC), machine learning, in combination with NLP, was tested

to improve “pre-syndromic surveillance”, which seeks to identify rare or previously unseen threats to health from clinical information (52). In this study, multidimensional semantic scan (MUSES) is a machine learning and NLP-based method developed to improve early detection of illness by eliminating the need for predefined case definitions and automatically clustering information by small geographies and/or demographics. MUSES was applied to historical free-text complaint data from NYC emergency departments and was found to identify more events of public health interest and a lower false positive rate than the current approaches used by the New York City Department of Health and Mental Hygiene (52). Natural language processing-based PHS has also been tested to improve the timeliness of overdose mortality reporting by eliminating the need for manual coding of free-text death certificates (53). The above examples show the potential of AI in PHS, but it remains unclear how many AI methods have been implemented into PHS. One real-world application by the Department of Veterans Affairs in the United States showed successful adaptation of an existing NLP-based PHS method early in the COVID-19 pandemic to monitor travel history in clinical records for public health follow-up (54).

Deep learning is a specialized type of machine learning that incorporates sophisticated neural networks that support classification using large amounts of text and are designed to work in a manner similar to a human brain. It has been increasingly used to support disease surveillance (7,55). The Centers for Disease Control and Prevention (CDC) tested neural networks and found that deep learning can interpret physician records to accurately predict the chief complaint, and potentially improve the timeliness and accuracy of information available for syndromic surveillance (56). Deep learning has also been applied to internet-based surveillance systems to support early warning, situational awareness, and nowcasting of infectious diseases. For example, Sentinel, an American surveillance system, uses deep learning to identify and classify health-related social media posts, news media, and CDC data to detect possible outbreaks and provide situational awareness (55).

The use of AI to support PHS is a new and emerging field that still needs evaluation of implementation into existing public health systems. Algorithms and machine learning models built with inaccurate, incomplete, or unrepresentative datasets, may both limit the accuracy of AI-based methods as well as bias results based on race, gender, or other characteristics (50,57). It is important to ensure that there is transparency in how AI models are built so that results are explainable, and that those who are interpreting the outputs of AI analyses are adequately skilled in PHS and can apply appropriate judgment. It is also important for public health professionals to understand AI methods, their applications, and their risks before applying it to public health practice (57).



### Nowcasting

Nowcasting uses recent surveillance data to model the current situation (e.g., case counts) when real-time data are unavailable (58). In one study, nowcasting using a Bayesian approach accurately estimated COVID-19 rates to inform resource allocation in NYC, successfully overcoming delays between testing and reporting (59). Advances in nowcasting have also been adopted in One Health surveillance systems to help fill data gaps and help anticipate zoonotic outbreaks. For example, the Norwegian Institute of Public Health successfully applied nowcasting principals to respond to gastrointestinal illness outbreaks using *Campylobacter* testing data from poultry farms and meteorological data (60). While nowcasting can be useful to estimate current situational awareness during rapidly changing public health emergencies, it is limited by the quality of data and the clarity of the interpretations provided to decision makers (59,61).

### Discussion

This review has explored innovations in PHS over the past decade and, where possible, described examples of their applications to PHS programs. Examples of the use of these novel sources to support PHS include providing novel information that improves estimates of disease, promoting early warning and identification of potential threats to health, and generating new information for public health action.

Despite these opportunities, there are substantial challenges to integrating innovations in PHS into practice. As new data sources and methods are added to the PHS toolbox, their risks and benefits should be considered with the goal of improving overall population health. Most of the areas explored in this paper are lacking in scientific maturity, and in many cases, are so novel that standard methods and best practices do not yet exist to help advance these fields reliably and responsibly (49,50,57). Many of the novel methods identified in this paper were tested in academic environments with no clear real-life implementation strategy (51,55). More evaluations of these interventions in real-world settings, which assess their utility in improving PHS and implications for public health action, are needed. These evaluations could be used to develop and disseminate guidance and standardized approaches to support public health organizations in implementing novel methods.

The use of digital technologies and AI in PHS also introduces challenges for privacy and security, data governance, and ethical considerations. For example, there is a need to balance between the benefits of having large quantities of granular information for analysis and the need to ensure individuals cannot be (re)identified. This is particularly true with AI methods, given the large quantity of information that is usually required to train the model (54,57,62,63). In the case of digital data, which may be

publicly available, but where permission to use for surveillance purposes has not been acquired, it is not clear how/whether informed consent can or needs to be obtained. Particular care needs to be taken to ensure that data are anonymized and confidential information is not revealed (63). Protection of digital data and transparency in how and what data is acquired, stored, and used are key to maintaining public trust and ensuring the sustainability of these systems (57,64), and thus progress towards digital data governance is needed to fully operationalize these data sources. Ethical frameworks for the use of AI and social media data in research (63), and guidelines for the use of AI more broadly (65–67), have been developed to support responsible conduct and protection of individuals from whom data is collected.

Health equity is an important consideration in implementing new surveillance methods. This overview identified several examples of approaches that could be used to support health equity, as they include populations that may be missed in traditional surveillance. However, a recent review article noted that there were no studies that specifically focused on vulnerable populations in the use of digital PHS, and thus substantial work is needed to explore the health equity implications of its use (17). Furthermore, greater work is needed to explore, identify, and address biases in AI algorithms and in the data used to train AI algorithms to ensure that these methods are not perpetuating harmful outputs as a consequence of biased inputs (57).

### Limitations

Limitations of this overview should be noted. This article was intended to provide a snapshot of recent innovations in PHS and explore examples of real-world application. As such, it is not intended to be an exhaustive list, and cannot provide detailed appraisal of the effectiveness of these innovations. The article focused on peer-reviewed literature only, and thus may have omitted articles from applied public health settings that were published as grey literature. The use of peer-reviewed literature may also have produced a positive publication bias, with studies noting negative results or unintended consequences potentially being under-represented. This is an important consideration given that non-traditional data sources may also be a source of public health misinformation (68), and thus require careful consideration and evaluation prior to use.

### Conclusion

Novel data and methods for PHS have the potential to improve the quantity, accuracy, completeness, timeliness, and accessibility of information available for public health response; however, the evidence base to support their utility in the real-world, as opposed to academic, settings appears to be lacking. Substantial barriers prevent the implementation of novel data and methods in PHS, ranging from health equity, privacy, and ethical concerns to training and availability of data and technologies. Improving data governance mechanisms, developing clear policies



for ethical use of AI technologies in PHS, and training the public health workforce on the responsible use of innovative technologies are important next steps towards advancing greater use of novel methods and data sources.

## Authors' statement

HR — Formal analysis, data analysis, writing—original draft, writing—review & editing

SS — Formal analysis, data analysis, writing—original draft, writing—review & editing

CZ — Formal analysis, data analysis, writing—original draft

LF — Scientific direction, supervision, writing—review

DP — Scientific direction, supervision, writing—review

DB — Scientific direction, supervision, writing—review

HR and SS contributed equally and are considered co-lead authors. All authors read and approved the final manuscript.

## Competing interests

None.

## Acknowledgements

We are thankful to our colleagues from the National Microbiology Lab and the Global Public Health Intelligence Network within the Public Health Agency of Canada for their subject matter review. We also acknowledge the skilled assistance from the Federal Science Libraries Network's Health Library for supporting our literature search.

## Funding

This work was supported by the Public Health Agency of Canada.

## References

- Centers for Disease Control and Prevention. Introduction to Public Health Surveillance. [Accessed 2024 Jan 19]. <https://www.cdc.gov/training/publichealth101/surveillance.html>
- Nsubuga P, White ME, Thacker SB, Anderson MA, Blount SB, Broome CV, Chiller TM, Espitia V, Imtiaz R, Sosin D, Stroup DF, Tauxe RV, Vijayaraghavan M, Trostle M. Public Health Surveillance: A Tool for Targeting and Monitoring Interventions. *Dis Control Priorities Dev Ctries*. 2006. [Accessed 2024 Jan 19]. <https://www.ncbi.nlm.nih.gov/books/NBK11770/>
- Public Health Agency of Canada. Surveillance. Ottawa, ON: PHAC; 2024. [Accessed 2024 Jan 19]. <https://www.canada.ca/en/public-health/services/public-health-practice/surveillance.html>
- Sahu KS, Majowicz SE, Dubin JA, Morita PP. NextGen Public Health Surveillance and the Internet of Things (IoT). *Front Public Health* 2021;9:756675. DOI PubMed
- Lavigne M, Mussa F, Creatore MI, Hoffman SJ, Buckeridge DL. A population health perspective on artificial intelligence. *Healthc Manage Forum* 2019;32(4):173–7. DOI PubMed
- Morgenstern JD, Rosella LC, Daley MJ, Goel V, Schünemann HJ, Piggott T. "AI's gonna have an impact on everything in society, so it has to have an impact on public health": a fundamental qualitative descriptive study of the implications of artificial intelligence for public health. *BMC Public Health* 2021;21(1):40. DOI PubMed
- Zeng D, Cao Z, Neill DB. Artificial intelligence-enabled public health surveillance—from local detection to global epidemic monitoring and control. In: *Artificial Intelligence in Medicine: Technical Basis and Clinical Applications*. Elsevier Applied Science 2020;437–53. DOI PubMed
- Francombe J, Ali GC, Gloinson ER, Feijao C, Morley KI, Gunashekar S, de Carvalho Gomes H. Assessing the Implementation of Digital Innovations in Response to the COVID-19 Pandemic to Address Key Public Health Functions: Scoping Review of Academic and Nonacademic Literature. *JMIR Public Health Surveill* 2022;8(7):e34605. DOI PubMed
- Aiello AE, Renson A, Zivich PN. Social media- and internet-based disease surveillance for public health. *Annual Rev Public Health* 2020;41:101–18. DOI PubMed
- Pilipiec P, Samsten I, Bota A. Surveillance of communicable diseases using social media: A systematic review. *PLoS One* 2023;18(2):e0282101. DOI PubMed
- Johnson AK, Bhaumik R, Tabidze I, Mehta SD. Nowcasting Sexually Transmitted Infections in Chicago: Predictive Modeling and Evaluation Study Using Google Trends. *JMIR Public Health Surveill* 2020;6(4):e20588. DOI PubMed
- Sadilek A, Caty S, DiPrete L, Mansour R, Schenk T, Bergtholdt M, Adam Jha A, Ramaswami P, Gabrilovich E. Machine-learned epidemiology: real-time detection of foodborne illness at scale. *npj Digit Med* 2018;1(1). DOI



13. Keller R, Spanu A, Puhan MA, Flahault A, Lovis C, Mütsch M, Beau-Lejdstrom R. Social media and internet search data to inform drug utilization: A systematic scoping review. *Front Digit Health* 2023;5:1074961. [DOI PubMed](#)
14. Hartnett KP, Kite-Powell A, Patel MT, Haag BL, Sheppard MJ, Dias TP, King BA, Melstrom PC, Ritchey MD, Stein Z, Idaikkadar N, Vivolo-Kantor AM, Rose DA, Briss PA, Layden JE, Rodgers L, Adjemian J. Syndromic Surveillance for E-Cigarette, or Vaping, Product Use-Associated Lung Injury. *N Engl J Med* 2020;382(8):766–72. [DOI PubMed](#)
15. Stolerman LM, Clemente L, Poirier C, Parag KV, Majumder A, Masyn S, Resch B, Santillana M. Using digital traces to build prospective and real-time county-level early warning systems to anticipate COVID-19 outbreaks in the United States. *Sci Adv* 2023;18;9(3):eabq0199. [DOI](#)
16. Espinosa L, Wijermans A, Orchard F, Höhle M, Czernichow T, Coletti P, Hermans L, Faes C, Kissling E, Mollet T. Epi tweeter: Early warning of public health threats using Twitter data. *Euro Surveill* 2022;27(39):2200177. [DOI](#)
17. Shakeri Hossein Abad Z, Kline A, Sultana M, Noaen M, Nurmambetova E, Lucini F, Al-Jefri M, Lee J. Digital public health surveillance: a systematic scoping review. *NPJ Digit Med* 2021;4(1):41. [DOI PubMed](#)
18. Yang S, Santillana M, Brownstein JS, Gray J, Richardson S, Kou SC. Using electronic health records and Internet search information for accurate influenza forecasting. *BMC Infect Dis* 2017;17(1):332. [DOI PubMed](#)
19. Lee L, Desroches M, Mukhi S, Bancej C. FluWatchers: evaluation of a crowdsourced influenza-like illness surveillance application for Canadian influenza seasons 2015–2016 to 2018–2019. *Can Commun Dis Rep* 2021;47(9):357–63. [DOI PubMed](#)
20. Kim AE, Brandstetter E, Wilcox N, Heimonen J, Graham C, Han PD, Starita LM, McCulloch DJ, Casto AM, Nickerson DA, Van de Loo MM, Mooney J, Ilcisin M, Fay KA, Lee J, Sibley TR, Lyon V, Geyer RE, Thompson M, Lutz BR, Rieder MJ, Bedford T, Boeckh M, Englund JA, Chu HY. Evaluating Specimen Quality and Results from a Community-Wide, Home-Based Respiratory Surveillance Study. *J Clin Microbiol* 2021;20;59(5):e02934-20. [DOI](#)
21. Astley CM, Tuli G, Mc Cord KA, Cohn EL, Rader B, Varrelman TJ, Chiu SL, Deng X, Stewart K, Farag TH, Barkume KM, LaRocca S, Morris KA, Kreuter F, Brownstein JS. Global monitoring of the impact of the COVID-19 pandemic through online surveys sampled from the Facebook user base. *Proc Natl Acad Sci USA* 2021;118(51):e2111455118. [DOI PubMed](#)
22. Mitze T, Rode J. Early-stage spatial disease surveillance of novel SARS-CoV-2 variants of concern in Germany with crowdsourced data. *Sci Rep* 2022;12(1):899. [DOI PubMed](#)
23. Canada Communicable Disease Report. ID News: Infectious disease climate change fund. *Can Commun Dis Rep* 2020;46(7/8):252. <https://www.canada.ca/en/public-health/services/reports-publications/canada-communicable-disease-report-ccdr/monthly-issue/2020-46/issue-7-8-july-2-2020/id-news.html#a1>
24. Cull B. Monitoring Trends in Distribution and Seasonality of Medically Important Ticks in North America Using Online Crowdsourced Records from iNaturalist. *Insects* 2022;13(5):404. [DOI PubMed](#)
25. Eritja R, Delacour-Estrella S, Ruiz-Arrondo I, González MA, Barceló C, García-Pérez AL, Lucientes J, Miranda MÁ, Bartumeus F. At the tip of an iceberg: citizen science and active surveillance collaborating to broaden the known distribution of *Aedes japonicus* in Spain. *Parasit Vectors* 2021;14(1):375. [DOI PubMed](#)
26. Pataki BA, Garriga J, Eritja R, Palmer JR, Bartumeus F, Csabai I. Deep learning identification for citizen science surveillance of tiger mosquitoes. *Sci Rep* 2021;11(1):4718. [DOI PubMed](#)
27. Hart CE, Bhaskar JR, Reynolds E, Hermance M, Earl M, Mahoney M, Martinez A, Petzlova I, Esterly AT, Thangamani S. Community engaged tick surveillance and tickMAP as a public health tool to track the emergence of ticks and tick-borne diseases in New York. *PLOS Glob Public Health* 2022;2(6):e0000215. [DOI](#)
28. Runkle JD, Sugg MM, Graham G, Hodge B, March T, Mullendore J, Tove F, Salyers M, Valeika S, Vaughan E. Participatory COVID-19 Surveillance Tool in Rural Appalachia : Real-Time Disease Monitoring and Regional Response. *Public Health Rep* 2021;136(3):327–37. [DOI PubMed](#)
29. Desroches M, Lee L, Mukhi S, Bancej C. Representativeness of the FluWatchers Participatory Disease Surveillance Program 2015–2016 to 2018–2019: how do participants compare with the Canadian population? *Can Commun Dis Rep* 2021;47(9):364–72. [DOI](#)
30. Huang X, Lu J, Gao S, Wang S, Liu Z, Wei H. Staying at Home Is a Privilege: Evidence from Fine-Grained Mobile Phone Location Data in the United States during the COVID-19 Pandemic. *Ann Am Assoc Geogr* 2022;112(1):286–305. [DOI](#)



31. Jay J, Bor J, Nsoesie EO, Lipson SK, Jones DK, Galea S, Raifman J. Neighbourhood income and physical distancing during the COVID-19 pandemic in the United States. *Nat Hum Behav* 2020;4(12):1294–302. [DOI PubMed](#)
32. Cheong SH, Ng YJ, Lau Y, Lau ST. Wearable technology for early detection of COVID-19: A systematic scoping review. *Prev Med* 2022;162:107170. [DOI PubMed](#)
33. Radin JM, Wineinger NE, Topol EJ, Steinhubl SR. Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study. *Lancet Digit Health* 2020;2(2):e85–93. [DOI PubMed](#)
34. Al Hossain F, Lover AA, Corey GA, Reich NG, Rahman T. FluSense: A contactless syndromic surveillance platform for influenza-like illness in hospital waiting areas. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2020;4(1):1. [DOI](#)
35. Wirth FN, Johns M, Meurers T, Prasser F. Citizen-centered mobile health apps collecting individual-level spatial data for infectious disease management: Scoping review. *JMIR Mhealth Uhealth* 2020;8(11):e22594. [DOI PubMed](#)
36. Vannice K, Hood J, Yarid N, Kay M, Harruff R, Duchin J. Accuracy of Medical Examiner's Assessment for Near-Real-Time Surveillance of Fatal Drug Overdoses, King County, Washington, March 2017–February 2018. *Public Health Rep* 2022;137(3):463–70. [DOI PubMed](#)
37. Kilaru P, Hill D, Anderson K, Collins MB, Green H, Kmush BL, Larsen DA. Wastewater Surveillance for Infectious Disease: A Systematic Review. *Am J Epidemiol* 2023;192(2):305–22. [DOI PubMed](#)
38. Wu F, Xiao A, Zhang J, Moniz K, Endo N, Armas F, Bonneau R, Brown MA, Bushman M, Chai PR, Duvallet C, Erickson TB, Foppe K, Ghaeli N, Gu X, Hanage WP, Huang KH, Lee WL, Matus M, McElroy KA, Nagler J, Rhode SF, Santillana M, Tucker JA, Wuertz S, Zhao S, Thompson J, Alm EJ. SARS-CoV-2 RNA concentrations in wastewater foreshadow dynamics and clinical presentation of new COVID-19 cases. *Sci Total Environ* 2022;805:150121. [DOI PubMed](#)
39. Cheng L, Dhiyebi HA, Varia M, Atanas K, Srikanthan N, Hayat S, Ikert H, Fuzzen M, Sing-Judge C, Badlani Y, Zeeb E, Bragg LM, Delatolla R, Giesy JP, Gilliland E, Servos MR. Omicron COVID-19 Case Estimates Based on Previous SARS-CoV-2 Wastewater Load, Regional Municipality of Peel, Ontario, Canada. *Emerg Infect Dis* 2023;29(8):1580–8. [DOI PubMed](#)
40. Morvan M, Jacomo AL, Souque C, Wade MJ, Hoffmann T, Pouwels K, Lilley C, Singer AC, Porter J, Evens NP, Walker DI, Bunce JT, Engeli A, Grimsley J, O'Reilly KM, Danon L. An analysis of 45 large-scale wastewater sites in England to estimate SARS-CoV-2 community prevalence. *Nat Commun* 2022;13(1):4313. [DOI PubMed](#)
41. Peccia J, Zulli A, Brackney DE, Grubaugh ND, Kaplan EH, Casanovas-Massana A, Ko AI, Malik AA, Wang D, Wang M, Warren JL, Weinberger DM, Arnold W, Omer SB. Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. *Nat Biotechnol* 2020;38(10):1164–7. [DOI PubMed](#)
42. Rousis NI, Li Z, Bade R, McLachlan MS, Mueller JF, O'Brien JW, Samanipour S, Tschärke BJ, Thomaidis NS, Thomas KV. Socioeconomic status and public health in Australia: A wastewater-based study. *Environ Int* 2022;167:107436. [DOI PubMed](#)
43. Kanneganti D, Reinersman LE, Holm RH, Smith T. Estimating sewage flow rate in Jefferson County, Kentucky, using machine learning for wastewater-based epidemiology applications. *Water Supply* 2022;22(12):8434–9. [DOI](#)
44. Xiao A, Wu F, Bushman M, Zhang J, Imakaev M, Chai PR, Duvallet C, Endo N, Erickson TB, Armas F, Arnold B, Chen H, Chandra F, Ghaeli N, Gu X, Hanage WP, Lee WL, Matus M, McElroy KA, Moniz K, Rhode SF, Thompson J, Alm EJ. Metrics to relate COVID-19 wastewater data to clinical testing dynamics. *Water Res* 2022;212:118070. [DOI PubMed](#)
45. Zhao L, Zou Y, David RE, Withington S, McFarlane S, Faust RA, Norton J, Xagoraki I. Simple methods for early warnings of COVID-19 surges: lessons learned from 21 months of wastewater and clinical data collection in Detroit, Michigan, United States. *Sci Total Environ* 2023;864:161152. [DOI PubMed](#)
46. Kotlarz N, Holcomb DA, Tanvir Pasha AB, Reckling S, Kays J, Lai YC, Daly S, Palani S, Bailey E, Guidry VT, Christensen A, Berkowitz S, Hoppin JA, Mitsova H, Engel LS, Reyes FL 3<sup>rd</sup>, Harris A. Timing and Trends for Municipal Wastewater, Lab-Confirmed Case, and Syndromic Case Surveillance of COVID-19 in Raleigh, North Carolina. *Am J Public Health* 2023;113(1):79–88. [DOI PubMed](#)
47. Löve AS, Ásgrímsson V, Ólafsdóttir K. Illicit drug use in Reykjavik by wastewater-based epidemiology. *Sci Total Environ* 2022;803:149795. [DOI PubMed](#)
48. Keshaviah A, Hu XC, Henry M. Developing a Flexible National Wastewater Surveillance System for COVID-19 and Beyond. *Environ Health Perspect* 2021;129(4):45002. [DOI PubMed](#)





49. National Academies of Sciences Engineering and Medicine. Wastewater-based Disease Surveillance for Public Health Action. Washington, DC: The National Academies Press; 2023. [PubMed](#)
50. Scardoni A, Balzarini F, Signorelli C, Cabitza F, Odone A. Artificial intelligence-based tools to control healthcare associated infections: A systematic review of the literature. *J Infect Public Health* 2020;13(8):1061–77. [DOI](#)
51. Ding Q, Massey D, Huang C, Grady CB, Lu Y, Cohen A, Matzner P, Mahajan S, Caraballo C, Kumar N, Xue Y, Dreyer R, Roy B, Krumholz HM. Tracking Self-reported Symptoms and Medical Conditions on Social Media During the COVID-19 Pandemic: infodemiological Study. *JMIR Public Health Surveill* 2021;7(9):e29413. [DOI](#) [PubMed](#)
52. Nobles M, Lall R, Mathes RW, Neill DB. Presyndromic surveillance for improved detection of emerging public health threats. *Sci Adv* 2022;8(44):eabm4920. [DOI](#) [PubMed](#)
53. Ward PJ, Rock PJ, Slavova S, Young AM, Bunn TL, Kavuluru R. Enhancing timeliness of drug overdose mortality surveillance: A machine learning approach. *PLoS One* 2019;14(10):e0223318. [DOI](#)
54. Peterson KS, Lewis J, Patterson OV, Chapman AB, Denhalter DW, Lye PA, Stevens VW, Gamage SD, Roselle GA, Wallace KS, Jones M. Automated Travel History Extraction From Clinical Notes for Informing the Detection of Emergent Infectious Disease Events: Algorithm Development and Validation. *JMIR Public Health Surveill* 2021;7(3):e26719. [DOI](#) [PubMed](#)
55. Şerban O, Thapen N, Maginnis B, Hankin C, Foot V. Real-time processing of social media with SENTINEL: A syndromic surveillance system incorporating deep learning for health classification. *Inf Process Manage* 2019;56(3):1166–84. [DOI](#)
56. Lee SH, Levin D, Finley PD, Heilig CM. Chief complaint classification with recurrent neural networks. *J Biomed Inform* 2019;93:103158. [DOI](#) [PubMed](#)
57. Murphy K, Di Ruggiero E, Upshur R, Willison DJ, Malhotra N, Cai JC, Malhotra N, Lui V, Gibson J. Artificial intelligence for good health: a scoping review of the ethics literature. *BMC Med Ethics* 2021;22(1):14. [DOI](#)
58. Wu JT, Leung K, Lam TT, Ni MY, Wong CK, Peiris JS, Leung GM. Nowcasting epidemics of novel pathogens: lessons from COVID-19. *Nat Med* 2021;27(3):388–95. [DOI](#) [PubMed](#)
59. Greene SK, McGough SF, Culp GM, Graf LE, Lipsitch M, Menzies NA, Kahn R. Nowcasting for Real-Time COVID-19 Tracking in New York City: An Evaluation Using Reportable Disease Data From Early in the Pandemic. *JMIR Public Health Surveill* 2021;7(1):e25538. [DOI](#) [PubMed](#)
60. Swanson D, Koren C, Hopp P, Jonsson ME, Rø GI, White RA, Grøneng GM. A One Health real-time surveillance system for nowcasting *Campylobacter* gastrointestinal illness outbreaks, Norway, week 30 2010 to week 11 2022. *Euro Surveill* 2022;27(43):2101121. [PubMed](#)
61. Charniga K, Madewell ZJ, Masters NB, Asher J, Nakazawa Y, Spicknall IH. Nowcasting and Forecasting the 2022 U.S. Mpox Outbreak: Support for Public Health Decision Making and Lessons Learned. *medRxiv*. 2023. [DOI](#)
62. Thorpe LE, Chunara R, Roberts T, Pantaleo N, Irvine C, Conderino S, Li Y, Hsieh PY, Gourevitch MN, Levine S, Ofrane R, Spoer B. Building Public Health Surveillance 3.0: Emerging Timely Measures of Physical, Economic, and Social Environmental Conditions Affecting Health. *Am J Public Health* 2022;112(10):1436–45. [DOI](#)
63. Gilbert JP, Ng V, Niu J, Rees EE. A call for an ethical framework when using social media data for artificial intelligence applications in public health research. *Can Commun Dis Rep* 2020;46(6):169–73. [DOI](#) [PubMed](#)
64. Smolinski MS, Crawley AW, Baltrusaitis K, Chunara R, Olsen JM, Wójcik O, Santillana M, Nguyen A, Brownstein JS. Flu near you: crowdsourced symptom reporting spanning 2 influenza seasons. *Am J Public Health* 2015;105(10):2124–30. [DOI](#) [PubMed](#)
65. Bosa K, Statistics Canada. Responsible use of machine learning at Statistics Canada. Ottawa, ON: StatCan; 2022. [Accessed 2023 Aug 28]. <https://www.statcan.gc.ca/en/data-science/network/machine-learning>
66. Government of Canada. Directive on Automated Decision Making. 2023. Ottawa, ON: Government of Canada; 2023. [Accessed 2023 Aug 28]. <https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592&section=html>
67. World Health Organization. Ethics and governance of artificial intelligence for health. 2021. Geneva, CH: WHO; 2021. <https://www.who.int/publications/item/9789240029200>
68. Knudsen J, Perlman-Gabel M, Uccelli IG, Jeavons J, Chokshi DA. Combating Misinformation as a Core Function of Public Health. *NEJM Catal* 2023;4(2). [DOI](#) [PubMed](#)