Check for updates

Emerging data inputs for infectious diseases surveillance and decision making

Aminath Shausan[1,2]*, Yoni Nazarathy[2] and Amalie Dyda[1]

[1]School of Public Health, The University of Queensland, Brisbane, QLD, Australia, [2]School of Mathematics and Physics, The University of Queensland, Brisbane, QLD, Australia

Infectious diseases create a significant health and social burden globally and can lead to outbreaks and epidemics. Timely surveillance for infectious diseases is required to inform both short and long term public responses and health policies. Novel data inputs for infectious disease surveillance and public health decision making are emerging, accelerated by the COVID-19 pandemic. These include the use of technology-enabled physiological measurements, crowd sourcing, field experiments, and artificial intelligence (AI). These technologies may provide benefits in relation to improved timeliness and reduced resource requirements in comparison to traditional methods. In this review paper, we describe current and emerging data inputs being used for infectious disease surveillance and summarize key benefits and limitations.

Introduction

Infectious diseases create significant health and social burden globally and can lead to outbreaks and epidemics, highlighted by COVID-19. As of November 29, 2022, there have been almost 638 million confirmed cases and over 6 million deaths of COVID-19 globally.[1] Prior to this, many other infectious diseases contributed significantly to global mortality and morbidity. Infectious disease threats have arisen in the form of the 2009 Swine Flu pandemic, Middle East respiratory syndrome coronavirus and the West Africa Ebola virus epidemic in West Africa. Cases of newly emerging infectious disease are likely to rise in coming years due to changes in population demographics, climate change, increases in international travel and changes in agriculture, land use, and sanitation (1).

Timely surveillance of infectious diseases is essential to plan targeted long and short term public health response and prevention efforts. Traditionally, infectious disease data are gathered by public health authorities in the form of positive laboratory tests and hospitalization and death data. Various non-government organizations, such as the World Health Organization (WHO) collect and report these data on a global level.[1] Additional global reporting for COVID-19 specifically has also been implemented by a number of organizations,[2] (2). However, such data requires significant effort from public health officials and may not be consistently collected on a national or regional scale. There can

---

[1]https://covid19.who.int
[2]www.ncov2019.live

also be delays in reporting and inconsistencies in reporting of deaths and positive cases. Hospitalization and deaths capture primarily severe cases and may not identify those who have mild symptoms or are asymptomatic. Due to these limitations, these data may not provide the true incidence of the disease.

Over recent years, novel data inputs have emerged to complement traditional surveillance of infectious diseases. This includes the use of sensors and mobile apps to collect symptoms and disease data, crowd sourcing data using Internet based surveys and mobile apps, field experiment data based on scenario analysis, as well as social media and web search data in conjunction with artificial intelligence methods to predict disease outbreaks. Surveillance using these types of data inputs may provide benefits over traditional data sources in that they can provide additional timeliness and are often less resource intensive. We aim to provide an overview of different types of novel data inputs for infectious diseases.

Online physiological measures

During the early phase of the pandemic, many governments implemented temperature checks, using classical thermometers, as a surveillance method at specific locations. However, this was limited by a single time point for each individual and may not identify all positive cases (3). Real-time temperature reading tools such as the smart thermometer and wearable devices have been tested as potential tools for tracking COVID-19. These tools measure physiological matrices of an individual to identify deviations from a baseline level, which may indicate onset of an illness.

Prior to COVID-19, the smart thermometer has been popularly applied in the US for monitoring flu (Influenza) (4) and personal health risk indicators. During early 2020, data from these thermometers was utilized to forecast COVID-19 hot spots in various US states by correlating anomalously high influenza levels with confirmed COVID-19 cases (5). Similarly, continuous wearable devices were originally applied to track influenza but have been translated to detect COVID-19 (6). Data from smartwatches and Fitbit trackers has been applied to detect pre-symptomatic cases of COVID-19 (7). The WHOOP fitness tracker has been used to identify pre-symptomatic COVID-19 cases by analyzing changes in respiratory rate and heart rate data gathered from users (8).

Data inputs from wearables combined with additional sources have also shown value in surveillance. Prior to the pandemic, the TempTraq skin patch,[3] a disposable skin patch with wireless temperature monitoring, was used for diagnosis and early detection of health risks. Data from the skin patch with data from Fitbit smartwatches and bio specimens have been used to assist health care workers to self-monitor COVID-19 (9). The

TemPredict (10) study investigated whether physiological data such as temperature and heart rate collected by Oura Ring, with self-reported daily symptoms, could predict COVID-19 symptoms. Results showed a correlation between elevated temperature and self-reported fever (11) and algorithms have been developed to identify onset of (12) and recovery from COVID-19 (13). The Corona Data Donation project uses data from wearable devices[4] with data, such as age and postcode, to provide a daily interactive fever map for COVID-19 hotspots (14). The Digital Engagement and Tracking for Early Control and Treatment (DETECT) study[5] involved collection of data from volunteers across the US via a smartwatch or activity tracker with self-reported symptoms and diagnostic test results. By evaluating changes in heart rate, sleep, activity levels, and symptoms, these data can identify cases with greater success than symptoms alone (15).

Crowd sourcing-surveillance

Crowd sourcing involves the collection of information from a large group of people. With the increasing use of the internet these methods are now utilizing digital platforms. For infectious disease surveillance, studies often capture data on self-reported infection via apps or online surveys.

Online surveys for surveillance began in 2003 with the Great Influenza Survey (de Grote Griepmeting) which collects influenza symptoms from participants on a weekly basis. This shows similarity between seasonal influenza measured in the study compared to data reported by general practitioners (16). Tracking influenza through online surveys expanded quickly, with influenza systems now set up in the US (17), Portugal (18), Europe (16), and Australia (19). Since COVID-19 emerged, many of these systems have been adapted to include surveillance for COVID-19 (20,21). The CoronaSurveys (22) and the Global COVID-19 Trends and Impact Survey (23) provide global and country specific daily incidence rates and risk areas for COVID-19. Both surveys are available in multiple languages, attracting participants globally. These types of data collection have shown to provide more timely outbreak signals than testing, hospitalization, and death counts (24). Some systems also incorporate an information feedback loop. The Flu Near You (17) surveillance program tracks influenza symptoms in the US and Canada. Users are able to compare these data with influenza data from the Centers for Disease Control and Prevention sentinel influenza network, and Google Flu Trends. Participants can also identify local sources for influenza vaccination (17).

Numerous mobile phone apps have also been developed to track COVID-19. The COVID Symptom Study app records user

---

reported symptoms.[6] This data has been used to track population incidence of COVID-19 and to identify hot spots and regions at high risk for COVID-19 (25). The data has been also applied to reveal long-lasting and short-term COVID-19 symptoms (26).

Other studies have focused on monitoring contact patterns and evaluating adherence to social distancing policies. A UK contact tracing survey documented participants age, contact location and level of physical contact (27). Comparing contact patterns during lockdown and social contacts made during a non-epidemic period showed physical distancing measures substantially reduced contact levels. Citizens in China are required to scan a QR code when accessing public spaces to verify their infection status. The QR code provides individual infection risk, based on locations and medical information(28). Contact tracing via QR coding system has been implemented in various other locations, such as the UK (29) and Australia (30).

Crowd sourcing methods are expanding to collect data for a wider range of diseases. The Mo-Buzz system identifies dengue hot spots in the Colombo, Sri Lanka (31). Participants provide dengue symptoms and post pictures of potential dengue mosquito–breeding sites. The system provides data as visual maps and also recommends educational materials. In Puerto Rico, a similar surveillance program called SaludBoricua monitors a number of acute illnesses including influenza, dengue, and leptospirosis (32).

Novel field experiments

Asymptomatic individuals can also contribute significant risk to the spread of diseases. As discussed previously, various efforts have been made to identify asymptotic individuals via contact tracing. A number of field experiments have also been performed to automatically monitor contact patterns. Data inputs are gathered using applications designed for cellular devices utilizing GPS location tracking.

The UK FluPhone study (33) investigated how social encounters shape the spread of infectious diseases. Participants reported influenza-like symptoms through the Fluphone app. The app also collects proximity of participants' devices through Bluetooth and location via GPS. Participants were informed about an estimate of the number of encounters they made and a real-time visualization of the spread of virtual SEIR type diseases in the contact network (34). The Safe Blues experiment[7] is a similar study, which took place at the University of Auckland during 2021–2022 (35,36), examining how physical interactions affect the spread of diseases. The Safe Blues Android app spreads multiple safe virtual virus strands via Bluetooth based on individual's physical proximity, with strands mimicking the behavior of biological viruses. Data is presented as a real-time and historic dashboard.

The 2018–2019 "Contagion! The BBC Four Pandemic" experiment (37) recruited nearly 30,000 volunteers to download the "BBC Pandemic app." The app records GPS location during daily activities, and self-reported contacts. Analysis showed that the contact network structure constructed had the potential for understanding and controlling real-world diseases such as the spread of influenza-like viruses (38). The data was also used to simulate various non-pharmaceutical intervention strategies, such as case isolation and social distancing to investigate their effectiveness on limiting the spread of the disease (39).

A number of these experiments have been conducted in the context of mass-gatherings. The RESTART-19 experiment conducted in Germany examined the transmission risks of SARS-CoV-2 in indoor mass-gathering events (40). Using contact tracing devices, participant movements were measured during a concert. This data was incorporated with a computational model of the arena simulating infectious aerosol distribution allowing the development of an individual-based model to estimate the excess burden of the COVID-19 epidemic caused by this event. The study concluded that with moderate hygiene practices and conditions of good ventilation, mass-gathering events contribute little to the spread of COVID-19.

A similar experiment in South Korea recruited healthcare workers involved in two protests (41) to self-report COVID-19 symptoms via an app. Following the rallies, PCR tests were conducted for 609 of the 646 (94%) of the attendees, with all tests negative. However, the PCR tests were performed on a sub sample of attendees and may not provide the true incidence rate. An indoor mass-gathering randomized clinical trial in Spain also assessed the risk of COVID-19 transmission during a concert (42). Contact tracing during the event was performed through a mobile phone app. After eight days, PCR tests showed that less than 1% of participants in the control group had a positive test result and no participant the intervention group tested positive. This provided preliminary evidence on the effectiveness of comprehensive measures to create safe indoor mass-gathering events.

To develop safe practices for mass-gatherings a series of trials called Fieldlab Events were carried out in The Netherlands during February and March 2021.[8] Four types of experiments investigated the risk of relaxing mask wearing and social distancing measures during mass-gathering events. Contact tracing devices and cameras were used to monitor participants' behavior during the events. Seven preventive measures were recommended for gatherings in addition to venues adhering to ventilation regulations for indoor events.

Artificial intelligence assisted data sourcing

Combining artificial intelligence (AI) and information available from informal sources is another surveillance method being applied to infectious diseases. These methods allow analysis and interpretation of large datasets sourced from various sources such

---

TABLE 1 Global surveillance systems incorporating artificial intelligence.

System	Type(s) of surveillance	AI methods	Data sources
GPHIN	Disease outbreaks, infectious diseases, contaminated food and water bioterrorism and exposure to chemicals, natural disasters, and issues related to the safety of products, drugs and medical devices and radioactive agents	KB, ML, NLP	Newspapers, online articles, emails, expert blogs, public health bodies and social media
HealthMap	Infectious disease outbreaks	ML, NLP	News media (e.g., Google News), expert-curated accounts (e.g., ProMED Mail), and validated official alerts World Health Organization announcements
BlueDot	Infectious disease outbreaks and risk assessments	ML, NLP	Online articles and airline data
Metabiota	Infectious disease outbreaks and risk assessments	ML, NLP	Data from various communities and regional clinics, online reports, social media data and airline data
WHO-EIOS	Outbreaks, risk assessments, pattern discovery and forecasting new outbreaks of all diseases	ML, NLP	Online media and specific social media sources, government and official web sites, news aggregators, blogs and expert groups, and collaborating initiatives (ProMED, GPHIN, HealthMap and the Europe Media Monitors
BioCaster	Infectious disease outbreaks	ML, NLP	Google News, ProMed Mail, MedISys, WHO, The United Nations, The European Centre for Disease Prevention and Control, Center for Infectious Disease Research and Policy, EurekAlert!, China News Service, The Food and Agriculture Organization of the United Nations
MedISys	Monitoring infectious diseases, bioterrorism, and chemical, biological, radiological and nuclear threats	KB, NLP, ML	Online articles

as online news and media, expert blogs, validated official alerts (e.g., WHO), clinical reports, social media, and airline data. The most common types of AI methods applied include knowledge-based (KB) techniques, natural language processing (NLP), and machine learning (ML). KB techniques use specific keywords and phrases determined by experts to quickly search through various types of information sources. NLP enables machines to process and interpret human language and perform tasks such as translation and disease classification. ML is the process of applying algorithms that allow machines to automatically learn and improve from experience (43). **Table 1** presents key global systems that incorporate AI methods for public health surveillance and their data sources.

The Global Public Health Intelligence Network (GPHIN[9]) was the first digital tool providing early warnings for diseases worldwide (44), and provided early detection of SARS in 2002 (45). The system automatically extracts data from media aggregators such as Google News, other online news, public health organizations and expert blogs, which are analyzed using KB, NLP, and ML methods. These findings are validated and fine tuned by experts and displayed as maps, graphs and listed documents (45). Initially the system monitored human diseases, however the platform has been expanded to include food safety surveillance.

The HealthMap,[10] BlueDot[11] and Metabiota[12] platforms integrate data from a range of formal and informal sources including news media such as Google News, expert curated

accounts such as ProMed-Mail, social media and validated official alerts such as WHO (46). These systems use NLP and ML to analyse these data for public health use. HealthMap was among the first systems to issue an early warning of COVID-19 outbreak 11 days prior to WHO confirming the disease (46). BlueDot (47) reported COVID-19 outbreaks 10 days before the disease was confirmed by WHO, and has successfully identified urban areas with high risk of spreading the disease (48).

The Epidemic Intelligence from Open Sources (EIOS[13]) is an initiative of WHO which integrates existing and new surveillance initiatives, networks, and systems globally (49). Data is gathered from a range of sources including traditional online media, social media, government and official web sites, blogs, and collaborating initiatives, such as the HealthMap and ProMed-Mail. Using NLP and ML the system sorts and categorizes articles by topics, country, language, source and contextual indices. Data from both EIOS and BlueDot were used to enhance event-based surveillance in Japan during the 2020 Olympic and Paralympic Summer Games, which helped to identify events, characterize risk and reduced the surveillance workload of Japan (50).

BioCaster[14] analyzes data from local and international news, sourced from news aggregators such as Google News and RSS feeds. Using NLP and ML the system automatically provides information of disease prevalence and outbreaks specific to a given location, as visual maps and graphs (51). Similarly, the Medical Information System (MedISys[15]) monitors events occurring across the globe using information from news articles

(52). The system curates data from official government websites, health news and the general media, and performs categorization of a range of public health threats by applying KB, ML and NLP techniques.

Discussion

These data inputs have several strengths in comparison to traditional methods. A key strength is timeliness as they often rely on automatic data collection. They may also reach to a wider population, more localized areas and population groups, and thus data collected from these systems can be more comprehensive. They can capture important aspects of a disease or transmission, such as changing behavior and perceptions regarding public health policies. These systems also require less human capital to run, and hence are less costly. When all aspects of digital systems are integrated, including data collection, analysis and presentation, this increases flexibility. Despite these benefits, there are notable limitations.

Tools measuring physiological indicators have limited specificity as they often rely on syndromic definitions of diseases, and there is difficulty in adjusting for confounders. Gathering more information in relation to demographics and risk factors to address this may lead to decreased in participation (53). Some systems using physiological measurements, as well as crowd sourcing, phone data, and mobile apps depend on user generated data. This method of data collection has inherent challenges, including introducing bias as those who choose to participate may not be representative of the general population. Data may also be biased due to user demography. Crowd sourcing can also require a participation over a long period time with recruitment of participants and ongoing engagement challenging.

Similarly, field experiments require a large number of participants making such experiments challenging to implement. This is highlighted by the lengthy disruption in the Safe Blues experiment due to the lockdown started during August 2021 in New Zealand (36). Furthermore, the results of these studies are restricted to specific populations, and may not be useful for different population groups or locations.

While powerful, mobile phone apps based on contact tracing, QR codes, and GPS based location trackers can create privacy and ethical challenges (54), and thus can be inefficient due to low user uptake (55,56). Privacy and ethical concerns can also arise from social media and web search data, due to the fact that such data are not primarily health related, often include personal characteristics, and owned by private online platforms (57). Such data can incorrectly predict prevalence due to having irrelevant information in search outputs or changing user behavior (58).

AI enabled systems collect and interpret data from a broad range of sources in real time and can potentially monitor disease trends and detect emerging diseases. However, this technology requires a large volume of data to make useful insights, and specialist computational infrastructure is needed to manage and store such data. Additionally, many of these data sources are unstructured and contain missing information and errors which again requires specific skills to analyze. As many of the AI enabled system integrate data from social media, they may not be representative of the population, with younger people over-represented in the samples (59).

In summary, new digital technologies provide benefits and opportunities to improve disease surveillance globally. However, the limitations of these data inputs and associated methods must be understood and addressed when incorporated into public health practice.

Author contributions

All authors contributed to the concept development of this review. AS conducted the initial review and data extraction. All authors contributed to the article and approved the submitted version.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Baker RE, Mahmud AS, Miller IF, Rajeev M, Fidisoa R, Rice BL, et al. Infectious disease in an era of global change. *Nat Rev Microbiol*. (2022) 20:193–205. doi: 10.1038/s41579-021-00639-z

2. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet*. (2020) 20:553–34. doi: 10.1016/S1473-3099(20)30120-1

3. Normile D. Airport screening is largely futile, research shows. *Science*. (2020) 367:1177–8. doi: 10.1126/science.367.6483.1177

4. Miller AC, Singh I, Koehler E, Polgreen PM. A smartphone-driven thermometer application for real-time population-and individual-level influenza surveillance. *Clin Infect Dis*. (2018) 67:388–97. doi: 10.1093/cid/ciy073

5. Chamberlain SD, Singh I, Ariza CA, Daitch AL, Philips PB, Dalziel BD. Real-time detection of COVID-19 epicenters within the United States using a network of smart thermometers. *MedRxiv* (2020). Available from: https://doi.org/10.1101/2020.04.06.20039909

6. Seshadri DR, Davies EV, Harlow ER, Hsu JJ, Knighton SC, Walker TA, et al. Wearable sensors for COVID-19: a call to action to harness our digital infrastructure for remote patient monitoring, virtual assessments. *Front Digit Health*. (2020) 2:8. doi: 10.3389/fdgth.2020.00008

7. Mishra T, Wang M, Metwally AA, Bogu GK, Brooks AW, Bahmani A, et al. Pre-symptomatic detection of influenza from smartwatch data. *Nat Biomed Eng*. (2020) 4:1208–20. doi: 10.1038/s41551-020-00640-6

8. Miller DJ, Capodilupo JV, Lastella M, Sargent C, Roach GD, Lee VH, et al. Analyzing changes in respiratory rate to predict the risk of COVID-19 infection. *PLoS ONE*. (2020) 15:e0243693. doi: 10.1371/journal.pone.0243693

9. Clingan CA, Dittakavi M, Rozwadowski M, Gilley KN, Cislo CR, Barabas J, et al. Monitoring health care workers at risk for COVID-19 using wearable sensors, smartphone technology: Protocol for an observational mHealth study. *JMIR Res Protoc*. (2021) 10:e29562. doi: 10.2196/29562

10. Purawat S, Dasgupta S, Song J, Davis S, Claypool KT, Chandra S, et al. TemPredict: a big data analytical platform for scalable exploration and monitoring of personalized multimodal data for COVID-19. *2021 IEEE International Conference on Big Data (Big Data)*. IEEE (2021). p. 4411–4420. Available from: https://doi.org/10.1109/BigData52589.2021.9671441.

11. Smarr BL, Aschbacher K, Fisher SM, Chowdhary A, Dilchert S, Puldon K, et al. Feasibility of continuous fever monitoring using wearable devices. *Sci Rep*. (2020) 10:1–11. doi: 10.1038/s41598-020-78355-6

12. Mason AE, Hecht FM, Davis SK, Natale JL, Hartogensis W, Damaso N, et al. Detection of COVID-19 using multimodal data from a wearable device: results from the first TemPredict Study. *Sci Rep*. (2022) 12:1–15. doi: 10.1038/s41598-022-07314-0

13. Mason AE, Kasl P, Hartogensis W, Natale JL, Dilchert S, Dasgupta S, et al. Metrics from wearable devices as candidate predictors of antibody response following vaccination against COVID-19: data from the second tempredict study. *Vaccines*. (2022) 10:264. doi: 10.3390/vaccines10020264

14. Wiedermann M, Rose AH, Maier BF, Kolb JJ, Hinrichs D, Brockmann D. Evidence for positive long- and short-term effects of vaccinations against COVID-19 in wearable sensor metrics–Insights from the German Corona Data Donation Project [preprint] (2022). Available at: https://doi.org/10.48550/arXiv.2204.02846

15. Quer G, Radin JM, Gadaleta M, Baca-Motes K, Ariniello L, Ramos E, et al. Wearable sensor data and self–reported symptoms for COVID-19 detection. *Nat Med*. (2021) 27:73–7. doi: 10.1038/s41591-020-1123-x

16. Koppeschaar CE, Colizza V, Guerrisi C, Turbelin C, Duggan J, Edmunds WJ, et al. Influenzanet: citizens among 10 countries collaborating to monitor influenza in Europe. *JMIR Public Health Surveill*. (2017) 3:e7429. doi: 10.2196/publichealth.7429

17. Chunara R, Aman S, Smolinski M, Brownstein JS. Flu near you: an online self-reported influenza surveillance system in the USA. *Online J Public Health Inform*. (2013) 5:e133. doi: 10.5210/ojphi.v5i1.4456

18. van Noort SP, Muehlen M, De Andrade HR, Koppeschaar C, Lourenço JL, Gomes MG. Gripenet: an internet-based system to monitor influenza-like illness uniformly across Europe. *Eurosurveillance*. (2007) 12:5–6. doi: 10.2807/esm.12.07.00722-en

19. Wójcik OP, Brownstein JS, Chunara R, Johansson MA. Public health for the people: participatory infectious disease surveillance in the digital age. *Emerg Themes Epidemiol*. (2014) 11:1–7. doi: 10.1186/1742-7622-11-7

20. Mccoll K, Debin M, Souty C, Guerrisi C, Turbelin C, Falchi A, et al. Are people optimistically biased about the risk of COVID-19 infection? Lessons from the first wave of the pandemic in Europe. *Int J Environ Res Public Health*. (2021) 19:436. doi: 10.3390/ijerph19010436

21. Sullivan SG, Carlson S, Cheng AC, Chilver MB, Dwyer DE, Irwin M, et al. Where has all the influenza gone? The impact of COVID-19 on the circulation of influenza, other respiratory viruses, Australia, March to September 2020. *Eurosurveillance*. (2020) 25:2001847. doi: 10.2807/1560-7917.ES.2020.25.47.2001847

22. Baquero C, Casari P, Fernandez Anta A, García-García A, Frey D, Garcia-Agundez A, et al. The CoronaSurveys system for COVID-19 incidence data collection and processing. *Front Comput Sci*. (2021) 3:641237. doi: 10.3389/fcomp.2021.641237

23. Astley CM, Tuli G, Mc Cord-De Iaco KA, Cohn EL, Rader B, Chiu S, et al. Global Monitoring of the Impact of COVID-19 Pandemic through Online Surveys Sampled from the Facebook User Base. medRxiv [Preprint] (2021). Available at: https://doi.org/10.1101/2021.07.05.21259989.

24. Alwashmi MF. The use of digital health in the detection and management of COVID-19. *Int J Environ Res Public Health*. (2020) 17:2906. doi: 10.3390/ijerph17082906

25. Menni C, Valdes AM, Freidin MB, Sudre CH, Nguyen LH, Drew DA, et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. *Nat Med*. (2020) 26:1037–40. doi: 10.1038/s41591-020-0916-2

26. Sudre CH, Murray B, Varsavsky T, Graham MS, Penfold RS, Bowyer RC, et al. Attributes and predictors of long COVID. *Nat Med*. (2021) 27:626–31. doi: 10.1038/s41591-021-01292-y

27. Jarvis CI, Van Zandvoort K, Gimma A, Prem K, Klepac P, Rubin GJ, et al. Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK. *BMC Med*. (2020) 18:1–10. doi: 10.1186/s12916-020-01597-8

28. Ye Q, Zhou J, Wu H. Using information technology to manage the COVID-19 pandemic: development of a technical framework based on practical experience in China. *JMIR Med Inform*. (2020) 8:e19515. doi: 10.1016/j.jinf.2020.07.024

29. Wymant C, Ferretti L, Tsallis D, Charalambides M, Abeler-Dörner L, Bonsall D, et al. The epidemiological impact of the NHS COVID-19 App. *Nature*. (2021) 594:408–12. doi: 10.1038/s41586-021-03606-z

30. [Dataset] Queensland Government. Check in Qld app data! (2021). Available from: https://www.data.qld.gov.au/article/news-and-events/check-in-qld-app-data [Accessed November 30, 2022]

31. Lwin MO, Jayasundar K, Sheldenkar A, Wijayamuni R, Wimalaratne P, Ernst KC, et al. Lessons from the implementation of Mo-Buzz, a mobile pandemic surveillance system for dengue. *JMIR Public Health Surveill*. (2017) 3:e7376. doi: 10.2196/publichealth.7376

32. Johansson M, Wojcik O, Chunara R, Smolinski M, Brownstein J. Participatory disease surveillance in Latin America. *Proceedings of the 22nd International Conference on World Wide Web*. New York: Association for Computing Machinery (2013). p. 695–696. Available from: https://doi.org/10.1145/2487788.2488025.

33. Yoneki E, Crowcroft J. EpiMap: towards quantifying contact networks for understanding epidemiology in developing countries. *Ad Hoc Netw*. (2014) 13:83–93. doi: 10.1016/j.adhoc.2012.06.003

34. Yoneki E. Fluphone study: virtual disease spread using haggle. *Proceedings of the 6th ACM Workshop on Challenged Networks*. New York: Association for Computing Machinery (2011). p. 65–66. Available from: https://doi.org/10.1145/2030652.2030672.

35. Dandekar R, Henderson SG, Jansen HM, McDonald J, Moka S, Nazarathy Y, et al. Safe Blues: the case for virtual safe virus spread in the long–term fight against epidemics. *Patterns*. (2021) 2:100220. doi: 10.1016/j.patter.2021.100220

36. Asanjarani A, Shausan A, Chew K, Graham T, Henderson SG, Jansen HM, et al. Emulation of epidemics via Bluetooth-based virtual safe virus spread: experimental setup, software, and data. *PLOS Digit Health*. (2022) 1:e0000142. doi: 10.1371/journal.pdig.0000142

37. Kissler SM, Klepac P, Tang M, Conlan AJ, Gog JR. Sparking "The BBC Four Pandemic": leveraging citizen science and mobile phones to model the spread of disease. BioRxiv [Priprint] (2020). p. 479154. Available at: https://doi.org/10.1101/479154.

38. Klepac P, Kissler S, Gog J. Contagion! the BBC four pandemic–the model behind the documentary. *Epidemics*. (2018) 24:49–59. doi: 10.1016/j.epidem.2018.03.003

39. Firth JA, Hellewell J, Klepac P, Kissler SM, Kucharski AJ, Spurgin LG, et al. Combining fine-scale social contact data with epidemic modelling reveals interactions between contact tracing, quarantine, testing and physical distancing for controlling COVID-19. MedRxiv [Priprint] (2020). Available at: https://doi.org/10.1101/2020.05.26.20113720.

40. Moritz S, Gottschick C, Horn J, Popp M, Langer S, Klee B, et al. The risk of indoor sports and culture events for the transmission of COVID-19. *Nat Commun*. (2021) 12:1–9. doi: 10.1038/s41467-021-25317-9

41. Kim SK, Kim EO, Kim SH, Jung J. Universal Screening of Severe Acute Respiratory Syndrome Coronavirus 2 with Polymerase Chain Reaction Testing after Rally of Trainee Doctors. *J Korean Med Sci*. (2020) 35:e380. doi: 10.3346/jkms.2020.35.e380

42. Revollo B, Blanco I, Soler P, Toro J, Izquierdo-Useros N, Puig J, et al. Same-day SARS-CoV-2 antigen test screening in an indoor mass-gathering live music event: a randomised controlled trial. *Lancet Infect Dis* (2021) 21:1365–72. doi: 10.1016/S1473-3099(21)00268-1

43. Ridenhour B, Kowalik JM, Shay DK. Unraveling R_0: Considerations for public health applications. *Am J Public Health*. (2014) 104:e32–e41. doi: 10.2105/AJPH.2013.301704

44. Dion M, AbdelMalik P, Mawudeku A. Big data: big data and the global public health intelligence network (GPHIN). *Can Commun Dis Rep*. (2015) 41:209. doi: 10.14745/ccdr.v41i09a02

45. Carter D, Stojanovic M, Hachey P, Fournier K, Rodier S, Wang Y. Redesigning GPHIN [Preprint] (2020). Available at: https://doi.org/10.48550/arXiv.2004.04596.

46. Freifeld CC, Mandl KD, Reis BY, Brownstein JS. HealthMap: global infectious disease monitoring through automated classification and visualization of Internet media reports. *J Am Med Inform Assoc*. (2008) 15:150–7. doi: 10.1197/jamia.M2544

47. [Dataset] BlueDot. BlueDot products: AI-driven infectious disease surveillance (2013). Available from: https://bluedot.global.

48. Bogoch II, Watts A, Thomas-Bachli A, Huber C, Kraemer MU, Khan K. Pneumonia of unknown aetiology in Wuhan, China: potential for international spread via commercial air travel. *J Travel Med*. (2020) 27:taaa008. doi: 10.1093/jtm/taaa008

49. Morgan O, Pebody R. The WHO hub for pandemic and epidemic intelligence: supporting better preparedness for future health emergencies. *Eurosurveillance*. (2022) 27:2200385. doi: 10.2807/1560-7917.ES.2022.27.20.2200385

50. Kasamatsu A, Ota M, Shimada T, Fukusumi M, Yamagishi T, Samuel A, et al. Enhanced event-based surveillance for imported diseases during the Tokyo 2020 Olympic and Paralympic Games. *Western Pacific Surveillance and Response Journal: WPSAR*. (2021) 12:1. doi: 10.5365/wpsar.2021.12.4.903

51. Meng Z, Okhmatovskaia A, Polleri M, Shen Y, Powell G, Fu Z, et al. BioCaster in 2021: automatic disease outbreaks detection from global news media. *Bioinformatics*. (2022) 38:4446–8. doi: 10.1093/bioinformatics/btac497

52. Linge JP, Steinberger R, Fuart F, Bucci S, Belyaeva J, Gemo M, et al. MedISys: medical information system. *Advanced ICTs for disaster management and threat detection: collaborative and distributed frameworks*. IGI Global (2010). p. 131–142. Available from: https://doi.org/10.4018/978-1-61520-987-3.ch009.

53. Trifan A, Oliveira M, Oliveira JL. Passive sensing of health outcomes through smartphones: systematic review of current solutions and possible limitations. *JMIR Mhealth Uhealth*. (2019) 7:e12649. doi: 10.2196/12649

54. Kleinman RA, Merkel C. Digital contact tracing for COVID-19. *CMAj*. (2020) 192:E653–6. doi: 10.1503/cmaj.200922

55. Vogt F, Haire B, Selvey L, Katelaris AL, Kaldor J. Effectiveness evaluation of digital contact tracing for COVID-19 in New South Wales, Australia. *Lancet Public Health*. (2022) 7:e250–8. doi: 10.1016/S2468-2667(22)00010-X

56. Thomas R, Michaleff ZA, Greenwood H, Abukmail E, Glasziou P. Concerns and misconceptions about the Australian government's COVIDsafe app: Cross-sectional survey study. *JMIR Public Health and Surveillance*. (2020) 6:e23081. doi: 10.2196/23081

57. Mittelstadt B, Benzler J, Engelmann L, Prainsack B, Vayena E. Is there a duty to participate in digital epidemiology? *Life Sci Soc Policy*. (2018) 14:1–24. doi: 10.1186/s40504-018-0074-1

58. Aiello AE, Renson A, Zivich P. Social media-and internet-based disease surveillance for public health. *Annu Rev Public Health*. (2020) 41:101. doi: 10.1146/annurev-publhealth-040119-094402

59. Naudé W. Artificial Intelligence against COVID-19: an early review. SSRN [Preprint] (2020). Available at: http://dx.doi.org/10.2139/ssrn.3568314.