Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Crowdsourced predictions for emerging infectious diseases

Alexander Spatari/Getty Images

Published **Online**
July 7, 2022
https://doi.org/10.1016/S2589-7500(22)00131-5

As of June 22, 2022, 3413 cases of monkeypox from 50 countries have been confirmed by WHO. Accurate and timely predictions on where and how fast monkeypox might spread are crucial to control the outbreak, and these could be provided by crowdsourced data. Such data were essential in characterising and predicting the spread of SARS-CoV-2 during the early stages of the pandemic. Given the promise they show for emerging infectious diseases, have crowdsourced predictions for such diseases come of age?

Traditional disease surveillance, using private data collected by health institutions, has informed public health decisions for decades, but there are shortcomings—particularly in the context of infectious disease outbreaks. For example, obtaining reliable data depends on timely diagnosis of the disease, which is less likely to be possible during the early stages of an outbreak. As such, disease can spread undetected for some time before being identified by surveillance. In unpredictable situations, such as the ongoing SARS-CoV-2 pandemic and the monkeypox outbreak, using faster alternatives might allow for more efficient monitoring.

Crowdsourcing is a resource that can be used to rapidly collect real-world infectious disease data, alongside technologies such as wearable devices to track influenza-like disease, and social media to gather data on COVID-19 symptoms. Other crowdsourcing projects collected data on the then-emerging COVID-19 pandemic, and notably the COVID Symptom Study mobile app allows the capture of self-reported COVID-19 symptoms, and provided predictions for COVID-19 infection rates and those at risk of long COVID. Now, two crowdsourced prediction platforms (centralised repositories that record real-world data provided by online contributors) are forecasting monkeypox infection and mortality rates: Metaculus and Map. As of July 5, Map reports 6178 confirmed cases of monkeypox, and ongoing research using data submitted to Metaculus suggests there is a 25% probability that WHO will declare monkeypox a Public Health Emergency of International Concern by the end of 2022. Unlike traditional disease surveillance, crowdsourced prediction platforms are usually easy and free to access, allow rapid sharing of real-world information, and are open to anyone with access to the internet to contribute, though this does mean contributions are limited to those with internet access.

Although there are benefits to using crowdsourced data for the monkeypox outbreak, there are limitations to consider, particularly while it is in its infancy and datasets are small and scarce. Identities of contributors to Metaculus cannot be verified, meaning predictions could be provided by non-experts and might not be accurate. As observations can be biased in the early stages due to availability of testing in different locations, Prof Moritz Kraemer, a computational and genomic epidemiologist at the University of Oxford, UK, using the Map tool to investigate monkeypox, notes it is important to "re-evaluate basic epidemiological parameters as the outbreak evolves". These parameters could include the basic reproduction number, mortality rates, and an updated incubation period. Computational scientist Assist Prof Thomas McAndrew of Lehigh University, USA, working on crowdsourcing for forecasting infectious disease transmission using Metaculus, notes that "an increase in timeliness of reporting often causes a decrease in completeness in datasets and vice versa" and, as such, a compromise must be made to maximise the accuracy of the predictions; timeliness of reporting is particularly important for monkeypox, due to its long incubation time (between 5 and 21 days). Traditional surveillance, while slower, is built on data collected by health professionals, and benefits from data that cannot be collected through an online platform—for example, antibody levels in an infected individual. Thus, traditional surveillance, alongside newer technologies, is still an essential resource.

Crowdsourced data have provided valuable information for several infectious diseases and represent a potentially powerful tool in understanding the monkeypox outbreak. Though traditional surveillance has its place, infectious disease outbreaks and pandemics are becoming increasingly common. As such, approaches need to adapt, and using timely and accurate crowdsourced predictions in tandem with traditional surveillance could be key in predicting and controlling the spread of monkeypox, and other infectious diseases that might arise in the future. ■ *The Lancet Digital Health*

For **WHO estimates on monkeypox infections** see https://www.who.int/emergencies/disease-outbreak-news/item/2022-DON396

For more on **wearable devices to track influenza-like disease** see **Articles** *Lancet Digit Health* 2020; **2:** e85–93

For more on the **use of social media to gather COVID-19 symptoms** see ECliniclaMedicine 2021; 38: 101019

For **examples of crowdsourcing initiatives used during the COVID-19 outbreak** see *R&D Management* 2020; published online Nov 15. https://doi.org/10.1111/radm.12443

For more on the **COVID smartphone app** see Nat Med 2020; 26: 1037–40

For more on the **Long COVID smartphone app** see Nat Med 2021; 27: 626–31

For more on **Metaculus** see https://www.metaculus.com/project/monkeypox/

For more on **Map** see https://map.monkeypox.global.health/country

For **Metaculus predictions for monkeypox** see **Comment** page e569

For more on the **use of Map in the context of monkeypox** see Correspondence Lancet Infect Dis 2022; 22: 941–42

For **facts about monkeypox** see https://www.who.int/news-room/fact-sheets/detail/monkeypox