An interactive web-based dashboard to track COVID-19 in real time

In December, 2019, a local outbreak of pneumonia of initially unknown cause was detected in Wuhan (Hubei, China), and was quickly determined to be caused by a novel coronavirus,[1] namely severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The outbreak has since spread to every province of mainland China as well as 27 other countries and regions, with more than 70 000 confirmed cases as of Feb 17, 2020.[2] In response to this ongoing public health emergency, we developed an online interactive dashboard, hosted by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University, Baltimore, MD, USA, to visualise and track reported cases of coronavirus disease 2019 (COVID-19) in real time. The dashboard, first shared publicly on Jan 22, illustrates the location and number of confirmed COVID-19 cases, deaths, and recoveries for all affected countries. It was developed to provide researchers, public health authorities, and the general public with a user-friendly tool to track the outbreak as it unfolds. All data collected and displayed are made freely available, initially through Google Sheets and now through a GitHub repository, along with the feature layers of the dashboard, which are now included in the Esri Living Atlas.

The dashboard reports cases at the province level in China; at the city level in the USA, Australia, and Canada; and at the country level otherwise. During Jan 22–31, all data collection and processing were done manually, and updates were typically done twice a day, morning and night (US Eastern Time). As the outbreak evolved, the manual reporting process became unsustainable; therefore, on Feb 1, we adopted a semi-automated living data stream strategy. Our primary data source is DXY, an online platform run
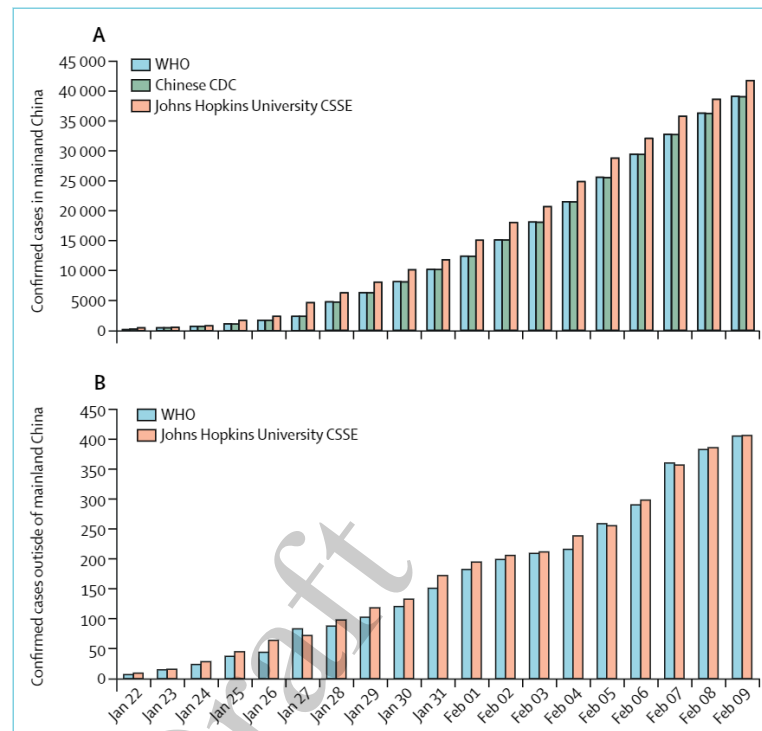

Figure: Comparison of COVID-19 case reporting from different sources
Daily cumulative case numbers (starting Jan 22, 2020) reported by the Johns Hopkins University Center for Systems Science and Engineering (CSSE), WHO situation reports, and the Chinese Center for Disease Control and Prevention (Chinese CDC) for within (A) and outside (B) mainland China.

by members of the Chinese medical community, which aggregates local media and government reports to provide cumulative totals of COVID-19 cases in near real time at the province level in China and at the country level otherwise. Every 15 min, the cumulative case counts are updated from DXY for all provinces in China and for other affected countries and regions. For countries and regions outside mainland China (including Hong Kong, Macau, and Taiwan), we found DXY cumulative case counts to frequently lag behind other sources; we therefore manually update these case numbers throughout the day when new cases are identified. To identify new cases, we monitor various Twitter feeds, online news services, and direct communication sent through the dashboard. Before manually updating the dashboard, we confirm the case numbers with regional and local health departments,

including the respective centres for disease control and prevention (CDC) of China, Taiwan, and Europe, the Hong Kong Department of Health, the Macau Government, and WHO, as well as city-level and state-level health authorities. For city-level case reports in the USA, Australia, and Canada, which we began reporting on Feb 1, we rely on the US CDC, the government of Canada, the Australian Government Department of Health, and various state or territory health authorities. All manual updates (for countries and regions outside mainland China) are coordinated by a team at Johns Hopkins University.

The case data reported on the dashboard aligns with the daily Chinese CDC[3] and WHO situation reports[2] for within and outside of mainland China, respectively (figure). Furthermore, the dashboard is particularly effective at capturing the timing of the first reported case of COVID-19 in new

For the **interactive dashboard of global COVID-19 cases** see https://arcg.is/0fHmTX

For the **COVID-19 data repository at GitHub** see https://github.com/CSSEGISandData/COVID-19

For **coronavirus data at DXY** see https://ncov.dxy.cn/ncovh5/view/pneumonia

countries or regions (appendix). With the exception of Australia, Hong Kong, and Italy, the CSSE at Johns Hopkins University has reported newly infected countries ahead of WHO, with Hong Kong and Italy reported within hours of the corresponding WHO situation report.

Given the popularity and impact of the dashboard to date, we plan to continue hosting and managing the tool throughout the entirety of the COVID-19 outbreak and to build out its capabilities to establish a standing tool to monitor and report on future outbreaks. We believe our efforts are crucial to help inform modelling efforts and control measures during the earliest stages of the outbreak.

We declare no competing interests.

We are grateful for the technical support from the Esri Living Atlas team and the Johns Hopkins University Applied Physics Lab.

*Ensheng Dong, Hongru Du, *Lauren Gardner*
l.gardner@jhu.edu

Department of Civil and Systems Engineering, Johns Hopkins University, Baltimore, MD 21218, USA (ED, HD, LG)

1 WHO. WHO statement regarding cluster of pneumonia cases in Wuhan, China. Jan 9, 2020. https://www.who.int/china/news/detail/09-01-2020-who-statement-regarding-cluster-of-pneumonia-cases-in-wuhan-china (accessed Feb 11, 2020).
2 WHO. Coronavirus disease 2019 (COVID-19) situation reports. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports (accessed Feb 17, 2020).
3 Chinese Center for Disease Control and Prevention. Tracking the epidemic. http://weekly.chinacdc.cn/news/TrackingtheEpidemic.htm (accessed Feb 11, 2020).

Open access epidemiological data from the COVID-19 outbreak

Coronavirus disease 2019 (COVID-19) is spreading rapidly across China, and as of Feb 16, 2020, had been reported in 26 countries globally.

The availability of accurate and robust epidemiological, clinical, and laboratory data early in an epidemic is important to guide public health decision-making.[1] Consistent recording of epidemiological information is important to understand transmissibility, risk of geographic spread, routes of transmission, and risk factors for infection, and to provide the baseline for epidemiological modelling that can inform planning of response and containment efforts to reduce the burden of disease. Furthermore, detailed information provided in real time is crucial for deciding where to prioritise surveillance.

Line list data are rarely available openly in real time during outbreaks. However, they enable a multiplicity of analyses to be undertaken by different groups, using various models and assumptions, which can help build consensus on robust inference. Parallels exist between this and the open sharing of genomic data.[2]

We have built a centralised repository of individual-level information on patients with laboratory-confirmed COVID-19 (in China, confirmed by detection of virus nucleic acid at the City and Provincial Centers for Disease Control and Prevention), including their travel history, location (highest resolution available and corresponding latitude and longitude), symptoms, and reported onset dates, as well as confirmation dates and basic demographics. Information is collated from a variety of sources, including official reports from WHO, Ministries of Health, and Chinese local, provincial, and national health authorities. If additional data are available from reliable online reports, they are included. Data are available openly and are updated on a regular basis (around twice a day).

We hope these data continue to be used to build evidence for planning, modelling, and epidemiological studies to better inform the public, policy makers, and international organisations and funders as to where

and how to improve surveillance, response efforts, and delivery of resources, which are crucial factors in containing the COVID-19 epidemic.

The epidemic is unfolding rapidly and reports are outdated quickly, so it will be necessary to build computational infrastructure that can handle the large expected increase in case reports. Data sharing will be vital to evaluate and maintain accurate reporting of cases during this outbreak.[3]

*Bo Xu, *Moritz U G Kraemer, on behalf of the Open COVID-19 Data Curation Group*
moritz.kraemer@zoo.ox.ac.uk

Department of Zoology, University of Oxford, Oxford OX1 3SZ, UK

1 Morgan O. How decision makers can use quantitative approaches to guide outbreak responses. *Philos Trans R Soc B Biol Sci* 2019; **374:** 20180365.
2 Yozwiak NL, Schaffner SF, Sabeti PC. Data sharing: make outbreak research open access. *Nature* 2015; **518:** 477–79.
3 Heymann DL. Data sharing and outbreaks: best practice exemplified. *Lancet* 2020; **395:** 469–70.

A family cluster of SARS-CoV-2 infection involving 11 patients in Nanjing, China

Human infection caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has become a global health concern.[1,2] Currently, human-to-human transmission of the virus accounts for most infections worldwide.[3] We report a family cluster of SARS-CoV-2 infection involving 11 patients in Nanjing, China.

The detailed timeline of exposure for the 11 confirmed patients is presented in the appendix 2. The index patient travelled to Nanjing on Jan 21, 2020, from Xiaogan (about 70 km from Wuhan), and switched trains in Wuhan. After arriving in Nanjing, she stayed with two of her

3