biases, and these programs can be an initial step toward health equity.

5. Promote professional accountability and commitment to equity and diversity. Professional associations, medical organizations and academic institutions must be held accountable. These entities have been very vocal about their anti-racism stance of late, but we need evidence of their commitment to equity and diversity. Organizations should ensure diverse leadership in their companies and executive boards and should create initiatives for addressing and dismantling systemic racism. We need action now, not just words.

6. Engage in self-reflection. You should acknowledge if reading this creates a feeling of defensiveness, and then think about your perception of Black people. What did your family tell you about Black people growing up? Has that changed? Are you afraid of Black men? How many Black patients do you care for? Do you give them the same respect and time that you give your white patients? Do you talk to them about what it means to be Black in America? How comfortable do you feel talking with Black patients about their experience?

These are just a few initial steps we should consider taking to start addressing racism and the social determinants of health in clinical settings. After the murder of so many innocent Black people, it is urgent that we address these issues. Beyond training and patient care, physicians have a voice in addressing racism, including but not limited to discussing with our patients during clinical visits the dangers that Black people face, contacting our legislators, organizing conferences to discuss how to better serve Black people, and compelling the judicial branches of government to develop fair and more-equitable justice for those who kill them. As healthcare providers, we must use our platform to speak up. Black people need us now. ❐

Frinny Polanco Walters[1 ✉],
Adjoa Anyane-Yeboa [ID][1] and
Alden M. Landry[2]

[1]The Commonwealth Fund Fellowship in Minority Health Policy at Harvard University, Harvard Medical School, Boston, MA, USA. [2]The Office for Diversity Inclusion and Community Partnership, Harvard Medical School, Boston, MA, USA.
✉e-mail: frinny_polancowalters@hms.harvard.edu

References
1. AAMC. www.aamc.org/data-reports/students-residents/interactive-data/2019-facts-applicants-and-matriculants-data (2020).
2. Laurencin, C. T. & Murray, M. J. Racial Ethn. Health Disparities 4, 317–321 (2017).
3. Saha, S., Guiton, G., Wimmers, P. F. & Wilkerson, L. J. Am. Med. Assoc. 300, 1135–1145 (2008).
4. Williams, D. R. J. Health Psychol. 2, 335–351 (1997).
5. Project Implicit. https://implicit.harvard.edu/implicit/ (2020).

Competing interests
The authors declare no competing interests.

Check for updates

Building an international consortium for tracking coronavirus health status

We call upon the research community to standardize efforts to use daily self-reported data about COVID-19 symptoms in the response to the pandemic and to form a collaborative consortium to maximize global gain while protecting participant privacy.

Eran Segal, Feng Zhang, Xihong Lin, Gary King, Ophir Shalem, Smadar Shilo, William E. Allen, Faisal Alquaddoomi, Han Altae-Tran, Simon Anders, Ran Balicer, Tal Bauman, Ximena Bonilla, Gisel Booman, Andrew T. Chan, Ori Cohen, Silvano Coletti, Natalie Davidson, Yuval Dor, David A. Drew, Olivier Elemento, Georgina Evans, Phil Ewels, Joshua Gale, Amir Gavrieli, Benjamin Geiger, Yonatan H. Grad, Casey S. Greene, Iman Hajirasouliha, Roman Jerala, Andre Kahles, Olli Kallioniemi, Ayya Keshet, Ljupco Kocarev, Gregory Landua, Tomer Meir, Aline Muller, Long H. Nguyen, Matej Oresic, Svetlana Ovchinnikova, Hedi Peterson, Jana Prodanova, Jay Rajagopal, Gunnar Rätsch, Hagai Rossman, Johan Rung, Andrea Sboner, Alexandros Sigaras, Tim Spector, Ron Steinherz, Irene Stevens, Jaak Vilo and Paul Wilmes

The rapid and global spread of COVID-19 led the World Health Organization to declare it a pandemic on 11 March 2020. One factor contributing to the spread of the pandemic is the lack of information about who is infected, in large part because of the lack of testing. This facilitated the silent spread of the causative coronavirus (SARS-CoV-2), which led to delays in public-health and government responses and an explosion in cases. In countries that have tested more aggressively and that had the capacity to transparently share this data, such as South Korea and Singapore, the spread of disease has been greatly slowed[1].

Although efforts are underway around the world to substantially ramp up testing capacity, technology-driven approaches to collecting self-reported information can fill an immediate need and complement official diagnostic results. This type of approach has been used for tracking other diseases, notably influenza[2]. The information collected may include health status that is self-reported through surveys, including those from mobile apps; results of diagnostic laboratory tests; and other static and real-time geospatial data. The collection of privacy-protected information from volunteers about health status over time may enable researchers to leverage these data to predict, respond to and learn about

Fig. 1 | Examples of questionnaires for symptom tracking. a, The English-language version of the daily questionnaire used in Israel (at https://coronaisrael. org/en/), showing a map of Israel with colors corresponding to a weighted average of COVID-19-associated symptoms, which is updated daily. The color of each region indicates a category defined by the average symptoms ratio, calculated by averaging of the reported symptoms rate by the responders in that city or neighborhood (color scale ranges from bright yellow (low symptoms ratio) to pink (high symptoms ratio)). Publ. note: Springer Nature is neutral about jurisdictional claims in maps. **b**, HowWeFeel: a 30-second survey of the person's well-being to collect epidemiological data.

the spread of COVID-19. Given the global nature of the disease, we aim to form an international consortium, tentatively named the 'Coronavirus Census Collective', to serve as a hub for amassing this type of data and to create a unified platform for global epidemiological data collection and analysis.

The mission of the Coronavirus Census Collective

The Coronavirus Census Collective (CCC) will be committed to a mission of saving lives through the open sharing of information covering all aspects of the COVID-19 pandemic, to the greatest extent possible, while simultaneously ensuring privacy. This infrastructure could immediately help in the current COVID-19 pandemic, and its consolidation will also facilitate the global response to other communicable diseases that may emerge in the future, as well as those that are currently present. Although it is our hope that the capacity for diagnostic testing will rapidly increase, testing will probably never provide global population-wide coverage, and there is thus a critical and immediate need for

collecting additional data on self-reported symptoms and health status at a population level. Moreover, our plan is to integrate the growing official diagnostic testing data from reverse-transcriptase PCR results and serology results with real-time informative data, such as information on self-reported symptoms, to better estimate the incidence of symptoms experienced by patients diagnosed with COVID-19 and to improve the epidemiological and predictive models that we will develop.

The CCC will integrate these data streams and thereby increase the statistical power of analyses carried out on these data, and provide a framework for data collection and a single, central data bank that researchers from around the world can query securely. In addition to ensuring comparability in the data, the CCC will serve as a resource for entities in other countries that are developing surveys for self-reporting, with the goal of facilitating their rapid deployment around the world. The CCC will seek individual grants in the respective countries as well as international funding for its central activities. As many

countries are now struggling to find the best strategy to handle the pandemic, we believe that a global collaborative effort to obtain data that can be used to predict outbreaks of COVID-19 is urgently needed. In the long term, this will serve as a rich source of information for understanding disease outbreaks that could facilitate policy decision-making and ensure that society is better positioned to respond to and prepare for future pandemics.

In carrying out this global endeavor, it is very important to engage low- and middle-income countries and to make efforts to include populations that are under-represented or of lower socioeconomic status, such as by distributing the survey in several languages, engaging leaders in local communities and promoting the survey through several diverse channels to increase compliance across all sectors of the population.

Current status of data collection

Early epidemiological studies of the COVID-19 outbreak in China, where COVID-19 was first documented,
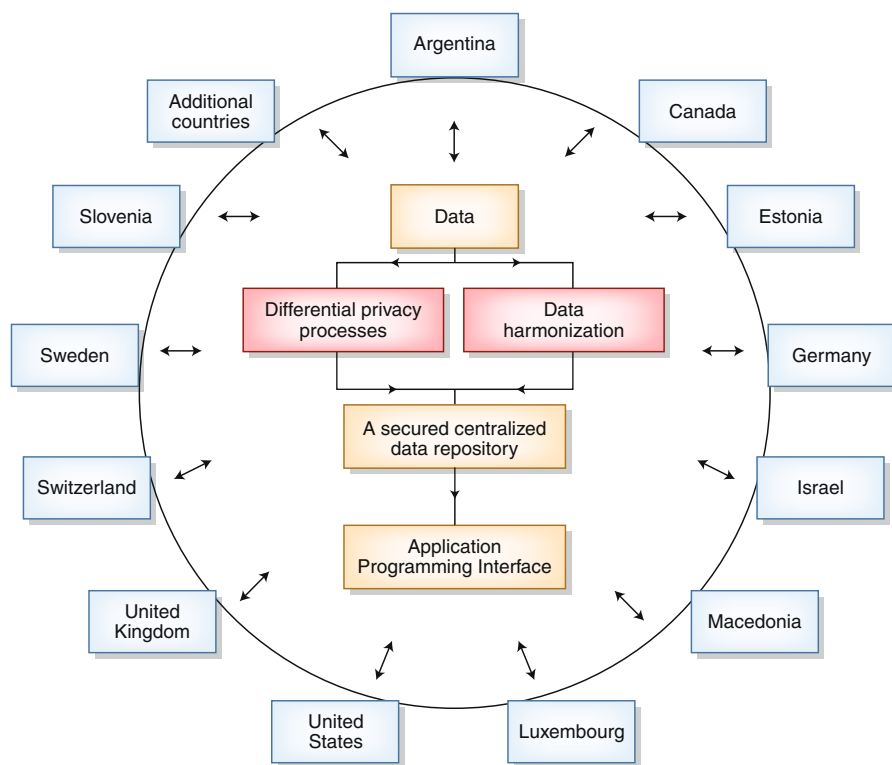
Fig. 2 | A flowchart of the data coordination process of the Coronavirus Census Collective (CCC).

demonstrated the overwhelming importance of slowing the rate of transmission to reduce the spread of COVID-19[3–5]. Slowing the rate of transmission requires information about who is infected and where they are located. The Chinese government achieved this goal in part through testing large numbers of people thought to be infected and moving people who had positive test results into isolation[3–5]. In South Korea, government officials took an approach of combining large-scale testing with transparent data sharing on the whereabouts of affected people to contain the outbreak. Although the success of these approaches is clear, there are two major limitations that prevent their application in the vast majority of other countries: diagnostic testing capacity, and personal and health privacy laws. For example, in the USA, currently the world epicenter of the pandemic, large-scale testing for SARS-CoV-2 infection is still not available in the vast majority of states. For several weeks after community spreading had been documented in multiple US states, testing was still limited to people with severe symptoms and with international travel history to places with early outbreaks (such as China)[6]. With limited capacity for testing, the numbers reported as 'confirmed cases' in these countries (e.g., https://coronavirus.jhu.edu/) probably do not reflect the true

numbers or the actual rate of COVID-19 spread, as screening tests reveal a substantial number of asymptomatic people infected with SARS-CoV-2[7–9].

Technology-driven approaches, such as designated apps or online surveys for collecting voluntarily self-reported data on health status, can overcome these limitations. These data can be further integrated with other relevant real-time data resources, including meteorological data and population density at a given time and place, as well as other dynamic data sources. Together, these can provide crucial information that can be immediately leveraged for early identification of disease clusters, with the goal of slowing the spread of disease.

Several such efforts have been started independently. For example, in Israel, a daily one-minute online survey, called Predict-Corona (https://coronaisrael.org/) (Fig. 1a), has been developed, in which respondents are asked to report their daily body temperature and whether they are experiencing any of the symptoms found to be common in patients with COVID-19, according to the existing literature. Within approximately 2 months of its launch, there have been over 2.5 million responses, and initial analyses demonstrate the potential of this approach for detecting

future outbreak regions[10]. In the USA, the app HowWeFeel (http://www.howwefeel.org), which administers a 30-second survey of the person's well-being to collect epidemiological data, has been developed (Fig. 1b). Similarly, CovidNearYou (https://covidnearyou.org/) and covid19 Risk Survey, a web app with open-access visualization tools collecting the same information in a fully anonymous way (https://covid19.eipm-research.org), were launched. In the USA and the UK, the Coronovirus Pandemic Consortium was established (https://www.monganinstitute.org/cope-consortium) to use the Covid Symptom Tracker (https://covid.joinzoe.com) and join the efforts of international prospective observational cohorts (e.g., population based, clinical data) and clinical trials[11]. This has led to a model that predicted COVID-19 incidence in over 2.7 million users[12].

This work parallels efforts in other countries to develop similar surveys, such as the COVID-19:CH Survey Project, which was developed in Switzerland (https://covidtracker.ch), and COVID-19 self-reporting in Slovenia (https://covid-19-stats.si/) and Estonia (koroona.ut.ee). These apps are currently managed and distributed by researchers in their respective research institutions.

A framework for collecting and sharing data
Survey responses containing symptom information relevant to COVID-19, along with geospatial location, time, and demographic information on age and pre-existing and comorbid medical conditions, will be collected. Once the data have been collected, they will be de-identified and made further differentially private[13] before researchers obtain access. Indeed, where feasible, we are implementing differential privacy at the data-collection source — the cellphone, website or app — so that any information that leaves the control of the research subjects cannot be used to learn about any person or whether a person is part of our dataset. We thus plan to provide specific mathematical assurances for the privacy of research subjects while still enabling scholars to make statistically valid inferences from these data. Because the privacy of our participants is essential to our mission, we commit to remaining at the cutting edge of privacy-protective technologies, developing novel methods where needed and implementing improvements over time wherever possible. Given differences in privacy regulations from region to region, individual-level data from the surveys from some countries may

not be accessible, but we will aim to make the results of differentially private statistical analyses available to all researchers. The mechanism of data sharing has not been determined yet and may be provided by an independent third party.

One of the main goals of the CCC is to create a federated common data model to facilitate data sharing while ensuring data security and privacy across different countries. To achieve this goal, we will define guidelines to ensure that the underlying data model collected from all contributors can be easily amalgamated and harmonized, and a consortium data-sharing policy will be developed. We will also apply different methods, such as differential privacy[13], that will enable researchers to share and analyze the data while preserving the confidentiality of the participants. To maximize the impact of the data collected, we will provide an app programming interface to allow accredited researchers to access the data to perform additional statistical analyses (Fig. 2).

Our technologies and the code for our app will be open source so that others can adapt them for their particular situations, find bugs and help us improve it. We envision that all members will use surveys with a common set of 'core' questions, but additional region-specific questions may be added, in accordance with local regulations, researcher interest and community need. This set of core questions may grow over time as more is learned about COVID-19.

These data will allow researchers to perform several immediately useful analyses. These include the following: monitoring of the health status of the respondent population, using statistics on self-reported symptoms and/or official COVID-19 test results within particular geographic regions at any given time; analysis of the epidemiological factors associated with symptoms and testing results; and identification of the areas likely to have COVID-19 outbreaks on the basis of the co-occurrence of many people's reporting similar patterns of symptoms at the same time. This will be potentially useful for identifying areas to which additional testing or medical resources should be allocated.

Additional applications include using methods of quantification, distinct from those of classification, to produce accurate estimates of the population prevalence of COVID-19, based on participants who have positive or negative test results, even when people cannot be reliably classified from their symptoms alone[14], and evaluating the effectiveness of the various social-distancing measures taken and their

contribution to reducing the number of symptomatic people.

Data coordination by a consortium

We envision that the CCC will be coordinated by a board of directors that will be agreed upon by all consortium members and who will vet new members, maintain a secured centralized data repository or federated multisite data repositories and develop mechanisms to enable researchers from around the world to query the data. The selection of the board of directors will take into account issues of expertise in organizing consortia, data privacy, epidemiology and technology. Individual members will be responsible for ensuring adherence to local regulations.

Conclusions

In summary, we call for participation in an international consortium, the CCC, that will serve as a hub for the integration of COVID-19-related information. This collective effort to track and share information will be invaluable in predicting hotspots of disease outbreak; identifying which factors control the rate of spreading; informing immediate policy decisions; evaluating the effectiveness of measures taken by health organizations on pandemic control; and providing critical insights on the etiology of COVID-19. It will also help people stay informed on this rapidly evolving situation and contribute to other global efforts to slow the spread of disease. ❑

Eran Segal [1] ✉, Feng Zhang[2,3], Xihong Lin [4], Gary King [5], Ophir Shalem [6], Smadar Shilo[1,7], William E. Allen[8], Faisal Alquaddoomi[9], Han Altae-Tran[3,10], Simon Anders [11], Ran Balicer[12], Tal Bauman[13], Ximena Bonilla [14], Gisel Booman [15], Andrew T. Chan [16], Ori Cohen[1], Silvano Coletti[17], Natalie Davidson[14], Yuval Dor[18], David A. Drew [16], Olivier Elemento[19,20], Georgina Evans[21], Phil Ewels [22], Joshua Gale[23], Amir Gavrieli[1], Benjamin Geiger[24], Yonatan H. Grad [25], Casey S. Greene[26], Iman Hajirasouliha[19,20], Roman Jerala [27], Andre Kahles[14], Olli Kallioniemi[28], Ayya Keshet[1], Ljupco Kocarev[29], Gregory Landua[30], Tomer Meir[1], Aline Muller[31], Long H. Nguyen[16], Matej Oresic [32,33], Svetlana Ovchinnikova[11], Hedi Peterson [34], Jana Prodanova[29], Jay Rajagopal[35,36], Gunnar Rätsch[37], Hagai Rossman[1], Johan Rung [38], Andrea Sboner[19,39], Alexandros Sigaras [19,20], Tim Spector [40], Ron Steinherz[30], Irene Stevens[41], Jaak Vilo [34] and Paul Wilmes[42]

[1]Department of Computer Science and Applied Mathematics, and Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel. [2]Howard Hughes Medical Institute, Chevy Chase, Maryland, USA. [3]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [4]Departments of Biostatistics and Statistics, Harvard T.H. Chan School of Public Health, Harvard University, Cambridge, MA, USA. [5]Institute for Quantitative Social Science, Harvard University, Cambridge, MA, USA. [6]Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [7]Pediatric Diabetes Unit, Ruth Rappaport Children's Hospital, Rambam Healthcare Campus, Haifa, Israel. [8]Harvard Society of Fellows, Harvard University, Cambridge, MA, USA. [9]ETH Zurich, NEXUS Personalized Health Technologies, Zurich, Switzerland. [10]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. [11]Center for Molecular Biology (ZMBH), University of Heidelberg, Heidelberg, Germany. [12]Clalit Research Institute, Clalit Health Services, Ramat Gan, Israel. [13]Mapping and Geo-Information Engineering, Civil and Environmental Engineering Faculty, The Technion, Haifa, Israel. [14]ETH Zurich, Department for Computer Science, Zurich, University Hospital Zurich, Medical Informatics, Zurich and SIB Swiss Institute of Bioinformatics, Zurich, Switzerland. [15]Regen Network, Mar del Plata, Argentina. [16]Massachusetts General Hospital (MGH), Boston, MA, USA. [17]Chelonia Applied Science, Allschwil, Switzerland. [18]IMRIC Developmental Biology and Cancer Research, School of Medicine, The Hebrew University, Jerusalem, Israel. [19]Englander Institute for Precision Medicine, Weill Cornell Medicine, New York, NY, USA. [20]Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, USA. [21]Institute for Quantitative Social Science, Harvard University, Cambridge, MA, USA. [22]Science for Life Laboratory (SciLifeLab), Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden. [23]http://symptometrics.org, Canada. [24]Department of Immunology, Weizmann Institute of Science, Rehovot, Israel. [25]Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Harvard University, Cambridge, MA, USA. [26]Department of Systems Pharmacology and Translational Therapeutics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [27]Department of Synthetic Biology and Immunology, National Institute of Chemistry, Ljubljana, Slovenia. [28]Science for Life Laboratory (SciLifeLab), Department of Oncology and Pathology, Karolinska Institutet, Stockholm, Sweden. [29]Macedonian Academy of Sciences and Arts, Skopje, Macedonia. [30]Regen Network, Great Barrington, MA, USA. [31]Luxembourg Institute of Socio-Economic Research and University of Luxembourg, Esch-sur-Alzette, Luxembourg. [32]School of Medical Sciences, Örebro University, Örebro, Sweden. [33]Turku Bioscience Centre, University of Turku and Åbo Akademi University, Turku, Finland. [34]Institute of Computer Science, University of Tartu, Tartu, Estonia. [35]Internal

Medicine, Harvard Medical School, Boston, MA, USA. ³⁶Department of Pulmonary Medicine and Critical Care, Massachusetts General Hospital (MGH), Boston, MA, USA. ³⁷ETH Zurich, Department for Computer Science, Zurich, University Hospital Zurich, Medical Informatics, Zurich and SIB Swiss Institute of Bioinformatics, Zurich and ELLIS Unit, ETH Zurich, Switzerland. ³⁸Science for Life Laboratory (SciLifeLab), Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden. ³⁹Department of Pathology and Laboratory Medicine, Weill Cornell Medicine, New York, NY, USA. ⁴⁰King's College, London, UK. ⁴¹Science for Life Laboratory (SciLifeLab), Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Stockholm, Sweden. ⁴²Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg.
✉e-mail: eran.segal@weizmann.ac.il

References

1. Tariq, A. et al. *medRxiv* https://doi.org/10.1101/2020.02.21.20026435 (2020).
2. Smolinski, M. S. et al. *Am. J. Public Health* **105**, 2124–2130 (2015).
3. Tian, H. et al. *Science* **368**, 638–642 (2020).
4. Guan, W.-J. et al. *N. Engl. J. Med.* **382**, 1708–1720 (2020).
5. Wu, Z. & McGoogan, J. M. *JAMA* **323**, 1239–1242 (2020).
6. Adalja, A. A., Toner, E. & Inglesby, T. V. *JAMA* **323**, 1343–1344 (2020).
7. Gudbjartsson, D. F. et al. *N. Engl. J. Med.* https://doi.org/10.1056/NEJMoa2006100 (2020).
8. Mizumoto, K., Kagaya, K., Zarebski, A. & Chowell, G. *Eurosurveillance* https://doi.org/10.2807/1560-7917.ES.2020.25.10.2000180 (2020).
9. Sutton, D., Fuchs, K., D'Alton, M. & Goffman, D. *N. Engl. J. Med.* https://doi.org/10.1056/NEJMc2009316 (2020).
10. Rossman, H. et al. *Nat. Med.* **26**, 634–638 (2020).
11. Drew, D. A. et al. *Science* https://doi.org/10.1126/science.abc0473 (2020).
12. Menni, C. et al. *Nat. Med.* https://doi.org/10.1038/s41591-020-0916-2 (2020).
13. Dankar, F. K. & El Emam, K. *Trans. Data Priv.* **6**, 35–67 (2013).
14. King, G. & Lu, Y. *Stat. Sci.* **23**, 78–91 (2008).

Check for updates

Use of apps in the COVID-19 response and the loss of privacy protection

Mobile apps provide a convenient source of tracking and data collection to fight against the spread of COVID-19. We report our analysis of 50 COVID-19-related apps, including their use and their access to personally identifiable information, to ensure that the right to privacy and civil liberties are protected.

Tanusree Sharma and Masooda Bashir

Compared with prior infectious-disease outbreaks (e.g., the 'Spanish flu' pandemic of 1918 and the 'Asian flu' pandemic of 1957), the COVID-19 emergency is occurring in a vastly more connected and digital world. Governments in multiple countries are pushing for location surveillance to contain the spread of COVID-19[1]. Digital surveillance may be the most effective way to contain the spread of the outbreak, but how privacy rights may be impacted must be considered both now and as this crisis moves forward. Fear and uncertainty often win out over civil liberties; however, as has been learned from past crises, such as the terrorist attacks of 11 September 2001, it can be hard to regain lost liberties[2]. Thus, it is critical not only that virus-response opportunities provided by technology be embraced but also that technology be used to ensure that the right to privacy is secured (Fig. 1).

Some countries, such as China, Israel, Singapore and South Korea, have launched tracking apps to fight the pandemic, and many more commercial apps have been released since the beginning of the outbreak. Here we examine 50 apps

available in the Google Play Store that have been developed specifically for COVID-19 (Supplementary Table 1).

The most common functionalities of the apps are as follows: live maps and updates of confirmed cases; real-time location-based alerts; systems for monitoring and controlling home isolation and quarantine, direct reporting to government, and self-reporting of symptoms; and education about COVID-19. Some more-advanced services include self-assessment of daily physiological status; monitoring of vital parameters, such as temperature, heart rate, oxygen and blood pressure, through the use of Bluetooth-enabled medical devices; virtual medical consultations (ADiLife Covid-19 in Italy); social science–based interventions based on predictive analysis of diseases in specific locations (OpenWHO); and community-driven contact tracing (TraceTogether and mfineRadar).

We found that 30 of the 50 apps require permission for numerous types of access to users' mobile devices. For example, some demand access to contacts, photos, media, files, location data, the camera, the device ID, call information, the WiFi connection,

the microphone, full network access, the Google service configuration, and the ability to change network connectivity and audio settings, to name just a few types of access. In addition, some apps explicitly state that they will collect information about the person's age, email address, phone number and postal code; the device's location, unique device identifiers, mobile IP address and operating system; and the types of browsers used on the mobile device.

A troubling discovery is that only 16 of the 50 apps indicate that the user's data will be made anonymous, encrypted and secured and will be transmitted online and reported only in an aggregated format. Our data represent a number of government-issued COVID-19 tracing apps that are from both developing countries and developed countries. Somewhat worryingly, 20 apps from our sample were issued by governments, health ministries and other such official sources. While the US government is not currently requiring citizens to download any tracking apps, there are apps in the Play Store that were developed by US healthcare providers (Sentinel Healthcare, 98point6