



Published in final edited form as:

*Cancer Epidemiol Biomarkers Prev.* 2020 July ; 29(7): 1283–1289. doi:10.1158/1055-9965.EPI-20-0606.

## The COronavirus Pandemic Epidemiology (COPE) Consortium: A Call to Action

*A full list of authors and affiliations appears at the end of the article.*

### Abstract

The rapid pace of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2; COVID-19) pandemic presents challenges to the real-time collection of population-scale data to inform near-term public health needs as well as future investigations. We established the COronavirus Pandemic Epidemiology (COPE) consortium to address this unprecedented crisis on behalf of the epidemiology research community. As a central component of this initiative, we have developed a COVID Symptom Study (previously known as the COVID Symptom Tracker) mobile application as a common data collection tool for epidemiologic cohort studies with active study participants. This mobile application collects information on risk factors, daily symptoms, and outcomes through a user-friendly interface that minimizes participant burden. Combined with our efforts within the general population, data collected from nearly 3 million participants in the United States and United Kingdom are being used to address critical needs in the emergency response, including identifying potential hot spots of disease and clinically actionable risk factors. The linkage of symptom data collected in the app with information and biospecimens already collected in epidemiology cohorts will position us to address key questions related to diet, lifestyle, environmental, and socioeconomic factors on susceptibility to COVID-19, clinical outcomes related to infection, and long-term physical, mental health, and financial sequelae. We call upon additional epidemiology cohorts to join this collective effort to strengthen our impact on the current health crisis and generate a new model for a collaborative and nimble research infrastructure that will lead to more rapid translation of our work for the betterment of public health.

---

In facing the unprecedented challenge posed by the severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) disease pandemic (COVID-19), there has perhaps never been a moment in history when the importance of epidemiologic research has been clearer. Within the medical community, leaders have described an “urgent need to expand public health activities to elucidate the epidemiology of the novel [SARS-CoV-2] virus and characterize its potential impact” (1). For the most part, infectious disease epidemiologists have been answering this call. The call to action for cancer and, more broadly, chronic disease

---

**Corresponding Author:** Andrew T. Chan, Massachusetts General Hospital, 100 Cambridge Street, 15th Floor, Boston, MA 02114. Phone: 617-726-3212; Fax: 617-724-6832; achan@mgh.harvard.edu.

#### Disclosure of Potential Conflicts of Interest

A.T. Chan reports receiving a commercial research grant from Zoe Global Ltd. P.W. Franks is a consultant for Zoe Global Ltd. C.R. Marinac reports receiving other commercial research support from GRAIL, Inc. S. Ourselin is a consultant for Johnson & Johnson. J. Wolf is CEO of and has ownership interest (including patents) in Zoe Global Ltd. T. Spector is a consultant for Zoe Global Ltd. No potential conflicts of interest were disclosed by the other authors.

epidemiologists has been less obvious. Work-as-usual has been sidelined with many standard activities of cohort studies or clinical trials unable to proceed. Study visits for participants have been suspended. Biospecimens are left uncollected and unanalyzed. The short-term consequence is that research into cancer, an important cause of mortality and chronic disease with a long-time horizon, has received less attention given the immediacy and lethality of the COVID-19 threat before us. The cancer epidemiology community is left with many questions. To what extent should we adjust our research priorities? What long-term impact will COVID-19 have on cancer-related or other outcomes due to the virus itself or delays in screening, diagnosis, and treatment? What new tools should we use and what infrastructure needs to be set in place to assess such outcomes? Can we leverage ongoing cohort studies to address the COVID-19 crisis in the present while laying a foundation for the future? What can the cancer epidemiology community do together in ways that we cannot do alone? We herein describe the development of the COronavirus Pandemic Epidemiology (COPE) consortium ([www.monganinstitute.org/cope-consortium](http://www.monganinstitute.org/cope-consortium)), a collaborative effort in the cancer epidemiology and clinical research community that has come together to address these and other questions posed by the COVID-19 crisis.

## Adjusting Our Research Priorities

While COVID has acutely become a leading cause of death in the United States (2), cardiovascular disease, cancer and other chronic diseases retain their position as leading causes of death and will remain there upon resolution of the pandemic. Nonetheless, COVID-19 will have an indelible impact on both ongoing and future studies. We must understand how this highly prevalent infection with substantial near-term risk will impact our study participants, particularly those with chronic underlying health conditions. Practically, this has implications on participant response to our usual assessments and loss-to-follow-up. Scientifically, this has implications on shorter-term clinical outcomes related to infection as well as downstream consequences. This may include chronic complications in patients with cancer related to COVID-19 disease, as well as indirect effects related to deferring medical care (including delays in screening, preventive interventions, or cancer treatment), the short- and long-term effects of stress on mental and physical health (and related health disparities), changes in important behaviors (e.g., diet and/or physical activity), and economic and social disruptions related to the COVID-19 public health response (3-5). Indeed, the long-term potential impact of such disruptions may be realized as spikes in cancer incidence or mortality in the years to come. There is also intriguing potential mechanistic connections between cancer oncogenes (e.g., *TMPRSS2*) and COVID-19 that may not yet be fully realized (6). Moreover, as we consider shifting research priorities, we must also remain cognizant of the primary goal of existing cohort studies. We must strike a balance where we address the unique effects of COVID-19 on cohort participants, without compromising the original mission of the cohorts. For example, we must carefully develop instruments for our cohorts that are not so burdensome that they lead to decreased future response rates or greater drop-out rates among critical subgroups.

## Novel Approaches to Data Collection to Address Unprecedented Questions during a Health Crisis

The usual approaches to data collection in cancer epidemiology cohorts leverage instruments that are methodologically rigorous but typically slow to deploy. For example, traditional surveys or biospecimen collection protocols usually require a stepwise approach to pilot testing and validation to ensure scientific rigor and participant engagement. Even after this initial, often time-consuming process, population-wide implementation is purposefully gradual. For the study of a slowly evolving, noncommunicable disease process such as cancer, such a deliberative approach makes sense. However, the speed at which the COVID-19 pandemic is unfolding poses an unprecedented challenge to rapidly collect data to characterize the full breadth of COVID-19 disease and its impact on chronic disease-related exposures and outcomes.

Although the urgency of the situation may require some compromise in our approach, we must also ensure scientific rigor and minimize bias while maximizing relevance (7). Therefore, the true challenge to epidemiology in the time of COVID-19 is identifying the optimal balance of speed and precision in our data collection and analysis. We recently discussed the need for novel tools that deploy an adaptable real-time data-capture platform to rapidly and prospectively collect actionable data that encompass the spectrum of subclinical and acute presentations of COVID-19 (8). To that end, we at Massachusetts General Hospital and the Harvard T.H. Chan School of Public Health partnered with physicians and epidemiologists at King's College London and software engineers at Zoe Global Ltd., a health data science company to develop a free mobile app (COVID Symptom Study, previously known as COVID Symptom Tracker; <http://covid.joinzoe.com/us>) to collect real-time data on COVID-19 symptoms, baseline health factors, infection status, and clinical outcomes, especially where they may lead to insights quickly enough to change practices that can improve the prognosis of persons with COVID-19 (8).

The app enables quick (less than 3–5 minutes initially and less than 1 minute per day on follow-up) self-report of data related to COVID-19 exposure and infections (Fig. 1). At first use, the app queries location, age, and core health risk factors. Daily prompts query for updates on interim symptoms, health care visits, and COVID-19 testing results. In those who are self-isolating or seeking health care, the level of intervention and related outcomes are collected. Individuals without symptoms are also encouraged to use the app so as to allow for prospective collection and sensitive capture of onset of symptoms. This prospective approach can overcome some of the challenges with cross-sectional assessment during an ongoing health crisis where variation at in the individuals participating at multiple timepoints may contribute to bias. Through pushed software updates, we can add or modify questions in real-time to test emerging hypotheses about COVID-19 or fine-tune data capture. Importantly, participants enrolled in ongoing epidemiology studies, clinical trials, or patient registries can provide informed consent to link survey information collected through the app through a one-way flow of data back to cohorts in a Health Insurance Portability and Accountability Act (HIPAA)–compliant manner. Cohorts can then link these data to their preexisting study cohort data. A specific module is also provided for participants who

identify as health care workers to determine the intensity and type of their direct patient care experiences, the availability and use of personal protective equipment (PPE), and work-related stress and anxiety. In future modules, it may be possible to ask additional questions from individuals living with cancer (e.g., treatment and survivorship), for which current cohort studies may not be well positioned to collect. Through deployment of this tool, we can rapidly gain critical insights into population dynamics of the disease (Fig. 2).

We have considered individual privacy at every step of planning for our approach and this is central to the one-way data flow structure to protect existing cohort data. Furthermore, we have minimized the number of personal identifiers collected by the app. For example, we collect year of birth, not day or month. Emails are used to register an individual to allow for secure access to the app via a user-set password, password recovery, and downstream linkage with existing cohort data. Participants may provide their names and phone numbers to assist with linkage, but these are optional data collection fields. Precise device location data are not collected; however, participants are asked to report their five-digit zip code to help with geolocations and public health recommendations based on area. All personal identifiers will be stripped in aggregate data repositories.

## Leveraging Cancer and Other Large-scale Epidemiology Cohorts

With our collective experience in designing and maintaining large prospective studies, investigators with epidemiologic cohorts are uniquely positioned to leverage the goodwill of a large number of highly engaged study participants across the United States. Thus, we moved rapidly to establish an international consortium of epidemiologic cohort studies, the COPE team, to bring together multidisciplinary expertise in “big data” research and state-of-the-art epidemiologic methods. ([www.monganinstitute.org/cope-consortium](http://www.monganinstitute.org/cope-consortium)). Within COPE, we have made the COVID Symptom Tracker app available for investigators to use in their cohorts. These cohorts range in size from small clinical populations, to large population cohort studies. Members of the general public can also download the app. To date, we have deployed the COVID Symptom Tracker app in long-standing U.S. prospective cohorts that have contributed significantly to cancer epidemiology research, including the Nurses' Health Study (NHS), NHSII, NHS3 (9), the Growing Up Today Study (GUTS; ref. 10), the Health Professionals Follow-up Study (HPFS; ref. 11), the Multiethnic Cohort Study (MEC; ref. 12), the American Cancer Society Cancer Prevention Study-3 (CPS-3; ref. 13), the California Teachers Study (CTS; ref. 14), the Black Women's Health Study (BWHS; ref. 15), the Sister Study (16), the CHASING COVID Cohort Study (17), Aspirin in Reducing Events in the Elderly (ASPREE; ref. 18), the PREDETERMINE study, and the Predicting Progression of Developing Myeloma in a High-Risk Screened Population (PROMISE; ref. 19), and Precursor Crowdsourcing (PCROWD) Studies (20). Several other cohorts with a focus on nutrition (Stanford Nutrition Studies) and environmental health [the Gulf Long-term Follow-up (GuLF) Study (21), the Agricultural Health Study (22), and the NIEHS Environmental Polymorphisms Registry (EPR; ref. 23)] have also joined our efforts. Our partners in the United Kingdom have enrolled participants within the TwinsUK Study, and U.K. Biobank, and collaborators in Canada, Australia, and Sweden are planning to launch versions of the COVID Symptom Study in their respective countries. Those already enrolled in participating cohort studies were sent an invitation to download the COVID Symptom

Tracker and asked to indicate their involvement in these cohorts. They were then invited to complete a study consent form that allows for the return of data collected through the app to the parent cohort study [under Institutional Review Board (IRB) for the specific institution and data sharing guidelines] for data linkage and future analysis.

## **COPE: Addressing the Near-term COVID-19 Crisis**

In the near term, COPE's goal is to provide data on the COVID-19 pandemic that can be used to contribute in real time to the emergency response and address critical needs raised by this outbreak. The use of the COVID Symptom Study app within COPE is embedded within a larger effort to encourage members of the general public to self-report symptoms through the app. This has also benefited from outreach led by organizations such as Stand Up to Cancer, the Susan Love Foundation for Breast Cancer Research, the Damon Runyon Cancer Research Foundation, and the American Cancer Society. Deidentified data collected through the app from the combined population cohort members with contributors from the general public will facilitate real-time assessment of the spread of COVID-19, including recognition of symptom presentation patterns and identification of new “hot spots” or regions at high risk for COVID-19 disease that may benefit from greater testing or expansion of hospital capacity. This information can also be used to inform guidelines around isolation measures. For example, within the United Kingdom, the app has grown to include over 2.7 million participants with the data collected being used to inform the U.K. government about areas predicted to experience a high burden of COVID approximately 5 days later (8). The health authorities, in turn, have acted upon the information and further promoted use of the app to its communities. In the United States, investigators at University of Texas School of Public Health aim to use data collected from the COVID Symptom Study app to conduct state-wide surveillance of COVID-19 symptoms in Texas to support public health decision making, especially as mitigation strategies are being deimplemented. Tools like these trackers are not only important for short-term tracking, but for long-term surveillance after areas attempt to reopen to normal activity.

Similarly, we are committed to working together with similar data participatory infectious disease surveillance efforts in the United States (24, 25), including collaborations with CovidNearYou and the HowWeFeel Project (26). Although several other digital collection tools for COVID-19 symptoms have been launched in the United States, these applications have largely been configured to offer a single assessment of symptoms to tailor recommendations for further evaluation. Others have been developed for researchers to report information on behalf of patients enrolled in clinical registries. While these approaches offer critical public health insights, they may not be tailored for scalable longitudinal data required for epidemiologists to perform comprehensive, well-powered investigations. Moreover, this tool is distinct from current efforts underway by large technology companies to embed digital contact tracing within cell phones, which are capable of alerting individuals who were in close contact with cell phone users who became infected with COVID-19, which may lead to additional surveillance or privacy concerns among potential app users. Nonetheless, a collective effort will be required to advance our understanding of COVID-19 and reach additional populations. For example, our app is currently available only in English and in the United States and the United Kingdom, and

while we are working to make this available in additional regions or languages, this takes significant development time and engineering resources. This creates a gap that complementary approaches may fill through continued collaboration, including translating the COVID Symptom Tracker questions and offering these via other platforms.

## COPE: A Foundation for Future Research

A longer-term goal of COPE is to link real-time data collected on COVID-19 symptoms with previously collected data on lifestyle, diet, and mental and physical health conditions as well as germline genetic and other biomarker information among the individuals participating in cohort studies. Based largely on reports from other countries, COVID-19 incidence and outcomes appear to vary according to age, sex, race, and comorbidities (5, 27-31). Conflicting reports that angiotensin-converting enzyme (ACE) inhibitors, thiazolidinediones, and ibuprofen may contribute to disease severity (32-34) led the European Medicines Agency to call for “epidemiology studies to be conducted in a timely manner” (35). Moreover, the risk of COVID-19 associated mortality appears highest in individuals with extensive comorbidities. In Wuhan, China, the case fatality rate was elevated among patients with pre-existing comorbid conditions—10.5% for cardiovascular disease, 7.3% for diabetes, 6.3% for chronic respiratory disease, 6.0% for hypertension, and 5.6% for cancer (36) compared with 2.3% in all confirmed cases. Additional evidence suggests infection risk differs by blood group (37), implicating genetic susceptibility in pathogenesis. Other clinical biomarkers (e.g., IL6/IFN $\gamma$  ratio; ref. 38) may be useful to understand risk and/or severity of COVID-19. Although there is no evidence yet for COVID-19, data from pooled randomized trials suggests vitamin D supplementation may lower the risk of respiratory infections (39). The influence of air pollution and climate on SARS-CoV-2 outcomes also requires investigation. These critical questions are hypotheses that can be readily tested in the living laboratories offered by cancer epidemiology cohorts with populations that have been extensively phenotyped. In COPE, investigators can pursue such studies within their own cohorts. These cohorts are uniquely positioned to address other key questions such as the impact of socioeconomic factors and health disparities, and potential risk modifiers such as diet, exercise, and supplement use. Specific cohorts will have the ability to link COVID Symptom app data with existing data on environmental and neighborhood level data (e.g., socioeconomic status, air pollution, poverty, rurality, etc.). However, we hope that cohorts can leverage the power of larger sample sizes by working together on joint analyses. We will also have an opportunity to investigate the long-term sequelae of COVID-19 infection and the response to the pandemic. We will want to understand the long-term impact of COVID-induced damage to multiple organ systems that may contribute to mechanisms of tumorigenesis or how interruptions to screening and surveillance may impact future cancer incidence. Continued follow-up within these cohorts will allow us to study the long-term impact on mental health and chronic disease incidence and management, mortality, and social and financial strain and how they relate to cancer incidence and survival. This may require collaborations that incorporate data collected through other efforts, including supplementary questionnaires or the collection of biospecimens that might be relevant to assess disease status (e.g., serologies). Eventually, such work within COPE can be further connected with other large-scale collaborations,

including the COVID-19 Host Genetics Initiative <https://www.covid19hg.org/data-sharing/>. Taken together, COPE will provide critical insights into the symptoms of infection, epidemiologic, and molecular predictors of disease, its underlying biology, and impact on long-term health outcomes.

## Overcoming Collective Challenges

Typical challenges in developing consortia include identifying committed leaders within each cohort who can facilitate human subjects' approval and data use agreements. In COPE, we have encouraged participation through the development of a “toolkit” to help facilitate human subjects IRB approval and establishment of clear channels for data collection and return of information. We are encouraged that most IRBs and academic medical center contracting authorities have streamlined their approval of research protocols and agreements related to COVID-19 to minimize delays. Some IRBs have also determined that initial promotion of a data collection tool does not require approval; study protocols can be modified at a later time for data linkage. Concerns have also been raised about participant burden during a time of crisis, particularly for cohorts that have already planned their own data collection initiatives for COVID-19 related outreach. However, as cohort leaders we bear a responsibility to offer a few opportunities for our participants to report on their experiences during this unprecedented life-changing event. It is especially important to prioritize approaches that may lead to near-term data to address the public health emergency.

## A Call to Action

We encourage additional cohorts to join COPE, particularly as it has become clear that the COVID-19 crisis will likely be prolonged. We request member cohorts to promote the COVID Symptom Tracker to their study participants through local IRB approval. To enhance the potential impact on public health decision-making, we hope to recruit more participants across the U.S. We have made targeted efforts through outreach to individual investigators, the AACR Molecular Epidemiology Working Group, and Stand Up To Cancer. As of April 28, 2020 there were approximately 200,000 users in the United States and additional users are being added daily. We anticipate that these numbers will continue to increase as cohorts at the earliest stages of joining this effort reach out to additional participants. We hope that more investigators will break out of traditional research silos to join us in deploying new technologies to overcome the usual barriers of establishing consortia. Collectively, we will have a stronger voice in advocating for the needs of the cancer epidemiology community, particularly if we can demonstrate that our work can have near-term impact. Our colleagues leading clinical trials of experimental agents for the treatment of COVID-19 are experiencing unprecedented fast-tracking of their protocols for approval, funding, and dissemination of results. Cohort studies should also demand that the usual rules of engagement for conducting our research be adapted to accommodate a quick response and be adequately resourced to be successful. This call to action for the cancer epidemiology community to mount a collective response will not only have an impact on the current health crisis but also generate a new model for a collaborative and nimble research infrastructure that will lead to more rapid translation of our work for the betterment of public health.

## Authors

Andrew T. Chan<sup>1,2,3</sup>, David A. Drew<sup>1,2</sup>, Long H. Nguyen<sup>1,2</sup>, Amit D. Joshi<sup>1,2</sup>, Wenjie Ma<sup>1,2</sup>, Chuan-Guo Guo<sup>1</sup>, Chun-Han Lo<sup>1</sup>, Raaj S. Mehta<sup>1,2</sup>, Sohee Kwon<sup>1</sup>, Daniel R. Sikavi<sup>1</sup>, Marina V. Magicheva-Gupta<sup>1,2</sup>, Zahra S. Fatehi<sup>1,2</sup>, Jacqueline J. Flynn<sup>1,2</sup>, Brianna M. Leonardo<sup>1,2</sup>, Christine M. Albert<sup>4</sup>, Gabriella Andreotti<sup>5</sup>, Laura E. Beane-Freeman<sup>5</sup>, Bijal A. Balasubramanian<sup>6</sup>, John S. Brownstein<sup>7</sup>, Fiona Bruinsma<sup>8</sup>, Annie N. Cowan<sup>9</sup>, Anusila Deka<sup>10</sup>, Michael E. Ernst<sup>11</sup>, Jane C. Figueiredo<sup>12</sup>, Paul W. Franks<sup>13,14</sup>, Christopher D. Gardner<sup>15</sup>, Irene M. Ghobrial<sup>9,16</sup>, Christopher A. Haiman<sup>17</sup>, Janet E. Hall<sup>18</sup>, Sandra L. Deming-Halverson<sup>19</sup>, Brenda Kirpach<sup>20</sup>, James V. Lacey Jr<sup>21</sup>, Loic Le Marchand<sup>22</sup>, Catherine R. Marinac<sup>9,16</sup>, Maria Elena Martinez<sup>23,24</sup>, Roger L. Milne<sup>8</sup>, Anne M. Murray<sup>20</sup>, Denis Nash<sup>25,26</sup>, Julie R. Palmer<sup>27</sup>, Alpa V. Patel<sup>10</sup>, Lynn Rosenberg<sup>27</sup>, Dale P. Sandler<sup>18</sup>, Shreela V. Sharma<sup>6</sup>, Shepherd H. Schurman<sup>18</sup>, Lynne R. Wilkens<sup>22</sup>, Jorge E. Chavarro<sup>3,13,28</sup>, A. Heather Eliassen<sup>28</sup>, Jaime E. Hart<sup>3,29</sup>, Jae Hee Kang<sup>3</sup>, Karestan C. Koenen<sup>28</sup>, Laura D. Kubzansky<sup>30</sup>, Lorelei A. Mucci<sup>28</sup>, Sebastien Ourselin<sup>31</sup>, Janet W. Rich-Edwards<sup>3</sup>, Mingyang Song<sup>1,2</sup>, Meir J. Stampfer<sup>3,28</sup>, Claire J. Steves<sup>31</sup>, Walter C. Willett<sup>13,28</sup>, Jonathan Wolf<sup>32</sup>, Tim Spector<sup>31,32</sup> **COPE Consortium**

## Affiliations

<sup>1</sup>Clinical & Translational Epidemiology Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts. <sup>2</sup>Division of Gastroenterology, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts. <sup>3</sup>Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts. <sup>4</sup>Department of Cardiology, Cedars-Sinai Hospital, Los Angeles, California. <sup>5</sup>Division of Cancer Epidemiology & Genetics, Occupational and Environmental Epidemiology Branch, National Institutes of Health, National Cancer Institute, Bethesda, Maryland. <sup>6</sup>Department of Epidemiology, Human Genetics, and Environmental Science, UTHealth School of Public Health, Houston, Texas. <sup>7</sup>Computational Epidemiology Group, Boston Children's Hospital, Boston, Massachusetts. <sup>8</sup>Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, Victoria, Australia. <sup>9</sup>Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts. <sup>10</sup>American Cancer Society, Atlanta, Georgia. <sup>11</sup>College of Pharmacy, University of Iowa, Iowa City, Iowa. <sup>12</sup>Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Hospital, Los Angeles, California. <sup>13</sup>Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, Massachusetts. <sup>14</sup>Department of Genetic and Molecular Epidemiology, Lund University, Malmö, Sweden. <sup>15</sup>Stanford Prevention Research Center, Department of Medicine, Stanford University, Stanford, California. <sup>16</sup>Harvard Medical School, Boston, Massachusetts. <sup>17</sup>Department of Preventive Medicine, Keck School of Medicine, Norris Comprehensive Cancer Center and the Epidemiology and Genetics Division, Department of Preventive Medicine, University of Southern California, Los Angeles, California. <sup>18</sup>National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North



Carolina. <sup>19</sup>Social & Scientific Systems, Durham, North Carolina. <sup>20</sup>Hennepin Health Care Research Institute, Berman Center for Outcomes and Clinical Research, Minneapolis, Minnesota. <sup>21</sup>Division of Health Analytics, Department of Computational and Quantitative Medicine, City of Hope, Duarte, California. <sup>22</sup>Cancer Center, University of Hawai'i, Honolulu, Hawaii. <sup>23</sup>Department of Family Medicine and Public Health, University of California, San Diego, La Jolla, California. <sup>24</sup>Moore's Cancer Center, University of California, San Diego, La Jolla California. <sup>25</sup>Institute for Implementation Science in Population Health, City University of New York (CUNY), New York, New York. <sup>26</sup>Department of Epidemiology and Biostatistics, School of Public Health, City University of New York (CUNY), New York, New York. <sup>27</sup>Slone Epidemiology Center, School of Medicine, Boston University, Boston, Massachusetts. <sup>28</sup>Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts. <sup>29</sup>Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts. <sup>30</sup>Department of Social and Behavioral Sciences and Lee Kum Sheung Center for Health and Happiness, Harvard T.H. Chan School of Public Health, Boston, Massachusetts. <sup>31</sup>Department of Twin Research & Genetic Epidemiology, Kings College, London, United Kingdom. <sup>32</sup>Zoe Global Limited, London, United Kingdom.

## Acknowledgments

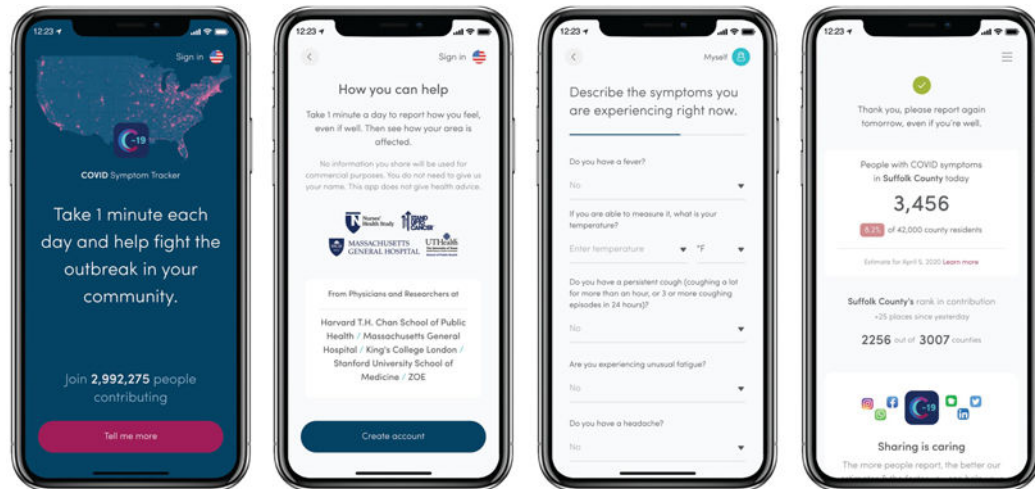
Zoe Global Ltd. provided in kind support for all aspects of building, running, and supporting the app and service to all users worldwide. S. Ourselin, C.J. Steves, and T. Spector were supported by the Wellcome Trust and Engineering and Physical Sciences Research Council (WT212904/Z/18/Z, WT203148/Z/16/Z, WT213038/Z/18/Z), the National Institute for Health Research (NIHR) Guy's and St Thomas'/King's College London Biomedical Research Centre, the Medical Research Council/British Heart Foundation (MR/M016560/1), the NIHR, and the Alzheimer's Society (AS-JF-17-011). A.T. Chan is the Stuart and Suzanne Steele Massachusetts General Hospital (MGH) Research Scholar. The Massachusetts Consortium on Pathogen Readiness (MassCPR) and Mark and Lisa Schwartz supported MGH investigators (A.T. Chan, D.A. Drew, L.H. Nguyen, A.D. Joshi, W. Ma, C.-G. Guo, C.-H. Lo, R.S. Mehta, S. Kwon, D.R. Sikavi, and M.V. Magicheva-Gupta).

## References

1. Lipsitch M, Swerdlow DL, Finelli L. Defining the epidemiology of Covid-19 - studies needed. *N Engl J Med* 2020;382:1194–6. [PubMed: 32074416]
2. Keating D, Esteban C. Covid-19 is rapidly becoming America's leading cause of death. 2020 4 16 The Washington Post Available from: <https://www.washingtonpost.com/outlook/2020/04/16/coronavirus-leading-cause-death/?arc404=true>.
3. Galea S, Merchant RM, Lurie N. The mental health consequences of COVID-19 and physical distancing: the need for prevention and early intervention. *JAMA Intern Med* 2020 4 10 [Epub ahead of print].
4. Pietrantonio F, Garassino MC. Caring for patients with cancer during the COVID-19 outbreak in Italy. *JAMA Oncol* 2020 4 10 [Epub ahead of print].
5. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA* 2020;323:1061.
6. Stopsack KH, Mucci LA, Antonarakis ES, Nelson PS, Kantoff PW. TMPRSS2 and COVID-19: serendipity or opportunity for intervention? *Cancer Discov* 2020 4 10 [Epub ahead of print].
7. Wynants L, Van Calster B, Bonten MMJ, Collins GS, Debray TPA, De Vos M, et al. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. *BMJ* 2020;369:m1328. [PubMed: 32265220]

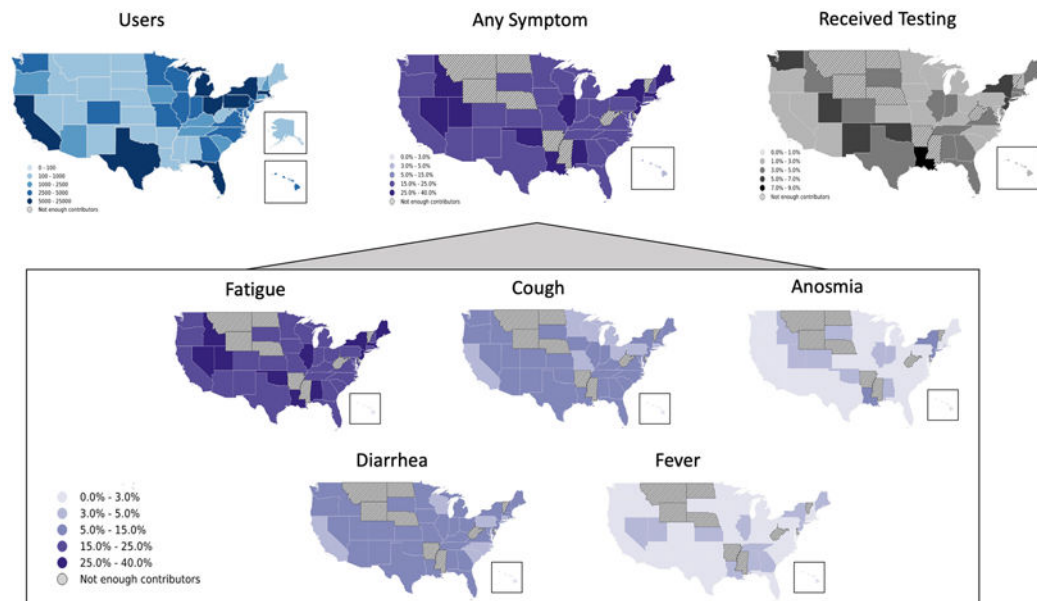
8. Drew DA, Nguyen LH, Steves CJ, Wolf J, Spector TD, Chan AT, et al. Rapid implementation of mobile technology for real-time epidemiology of COVID-19. *Science* 2020 5 5 [Epub ahead of print].
9. Bao Y, Bertoia ML, Lenart EB, Stampfer MJ, Willett WC, Speizer FE, et al. Origin, methods, and evolution of the three Nurses' Health Studies. *Am J Public Health* 2016;106:1573–81. [PubMed: 27459450]
10. Rockett HR, Berkey CS, Field AE, Colditz GA. Cross-sectional measurement of nutrient intake among adolescents in 1996. *Prev Med* 2001;33:27–37. [PubMed: 11482993]
11. Rimm EB, Giovannucci EL, Willett WC, Colditz GA, Ascherio A, Rosner B, et al. Prospective study of alcohol consumption and risk of coronary disease in men. *Lancet* 1991;338:464–8. [PubMed: 1678444]
12. Kolonel LN, Henderson BE, Hankin JH, Nomura AM, Wilkens LR, Pike MC, et al. A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. *Am J Epidemiol* 2000;151:346–57. [PubMed: 10695593]
13. Patel AV, Jacobs EJ, Dudas DM, Briggs PJ, Lichtman CJ, Bain EB, et al. The American Cancer Society's Cancer Prevention Study 3 (CPS-3): recruitment, study design, and baseline characteristics. *Cancer* 2017;123:2014–24. [PubMed: 28171707]
14. Bernstein L, Allen M, Anton-Culver H, Deapen D, Horn-Ross PL, Peel D, et al. High breast cancer incidence rates among California teachers: results from the California Teachers Study (United States). *Cancer Causes Control* 2002;13:625–35. [PubMed: 12296510]
15. Rosenberg L, Adams-Campbell L, Palmer JR. The Black Women's Health Study: a follow-up study for causes and preventions of illness. *J Am Med Womens Assoc* 1995;50:56–8.
16. Sandler DP, Hodgson ME, Deming-Halverson SL, Juras PS, D'Aloisio AA, Suarez LM, et al. The sister study cohort: baseline methods and participant characteristics. *Environ Health Perspect* 2017;125:127003. [PubMed: 29373861]
17. CHASING COVID Research Study; [about 4 screens]. Available from: <https://cunyisph.org/chasing-covid/>.
18. Group AI. Study design of ASPirin in Reducing Events in the Elderly (ASPREE): a randomized, controlled trial. *Contemp Clin Trials* 2013;36:555–64. [PubMed: 24113028]
19. The PROMISE Study [homepage on the Internet]. Available from: <https://www.enroll.promisestudy.org/>.
20. The PCROWD Study [homepage on the Internet]. Available from: <https://www.enroll.pcrowd.org/>.
21. Kwok RK, Engel LS, Miller AK, Blair A, Curry MD, Jackson WB, et al. The GuLF STUDY: a prospective study of persons involved in the deepwater horizon oil spill response and clean-up. *Environ Health Perspect* 2017;125:570–8. [PubMed: 28362265]
22. Alavanja MC, Akland G, Baird D, Blair A, Bond A, Dosemeci M, et al. Cancer and noncancer risk to women in agriculture and pest control: the Agricultural Health Study. *J Occup Med* 1994;36:1247–50. [PubMed: 7532217]
23. Chulada PC, Vainorius E, Garantziotis S, Burch LH, Blackshear PJ, Zeldin DC. The Environmental Polymorphism Registry: a unique resource that facilitates translational research of environmental disease. *Environ Health Perspect* 2011;119:1523–7. [PubMed: 21659040]
24. Wojcik OP, Brownstein JS, Chunara R, Johansson MA. Public health for the people: participatory infectious disease surveillance in the digital age. *Emerg Themes Epidemiol* 2014;11:7. [PubMed: 24991229]
25. Smolinski MS, Crawley AW, Baltrusaitis K, Chunara R, Olsen JM, Wojcik O, et al. Flu near you: crowdsourced symptom reporting spanning 2 influenza seasons. *Am J Public Health* 2015;105:2124–30. [PubMed: 26270299]
26. Segal E, Zhang F, Lin X, King G, Shalem O, Shilo S, et al. Building an international consortium for tracking coronavirus health status. *Nat Med*. In Press.
27. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 2020;382:1708–20. [PubMed: 32109013]
28. Wu C, Chen X, Cai Y, Xia Ja, Zhou X, Xu S, et al. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA Intern Med* 2020 3 13 [Epub ahead of print].

29. Yang X, Yu Y, Xu J, Shu H, Xia Ja, Liu H, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *Lancet Respir Med* 2020;8:475–81. [PubMed: 32105632]
30. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet North Am Ed* 2020;395:1054–62.
31. Gudbjartsson DF, Helgason A, Jonsson H, Magnusson OT, Melsted P, Norddahl GL, et al. Spread of SARS-CoV-2 in the Icelandic Population. *N Engl J Med* 2020.
32. Day M Covid-19: ibuprofen should not be used for managing symptoms, say doctors and scientists. *BMJ* 2020;368:m1086. [PubMed: 32184201]
33. Fang L, Karakiulakis G, Roth M. Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection? *Lancet Respir Med* 2020;8:e21. [PubMed: 32171062]
34. Zhang P, Zhu L, Cai J, Lei F, Qin JJ, Xie J, et al. Association of inpatient use of angiotensin converting enzyme inhibitors and angiotensin ii receptor blockers with mortality among patients with hypertension hospitalized with COVID-19. *Circ Res* 2020 4 17 [Epub ahead of print].
35. FitzGerald GA. Misguided drug advice for COVID-19. *Science* 2020;367:1434.
36. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72314 cases from the chinese center for disease control and prevention. *JAMA* 2020 2 24 [Epub ahead of print].
37. Zhao J, Yang Y, Huang H, Li D, Gu D, Lu X, et al. Relationship between the ABO blood group and the COVID-19 susceptibility. *medRxiv* 2020.03.11.20031096 [Preprint]. 2020 3 27 Available from: 10.1101/2020.03.11.20031096.
38. Lagunas-Rangel FA, Chavez-Valencia V. High IL-6/IFN-gamma ratio could be associated with severe disease in COVID-19 patients. *J Med Virol* 2020 4 16 [Epub ahead of print].
39. Martineau AR, Jolliffe DA, Hooper RL, Greenberg L, Aloia JF, Bergman P, et al. Vitamin D supplementation to prevent acute respiratory tract infections: systematic review and meta-analysis of individual participant data. *BMJ* 2017;356:i6583. [PubMed: 28202713]



**Figure 1.**

The COVID Symptom Study smartphone application. After providing informed consent and registering, participants are asked basic demographic information, baseline health questions, and whether they are experiencing any potential COVID-19 symptoms or have received testing or treatment for COVID-19. There is a specific module of questions for health care workers related to possible COVID-19 exposures and PPE use. Participants are requested to return to the app daily to log how they are feeling, even when they are well or symptoms have resolved. If they report feeling unwell, they are asked about specific symptoms. Finally, participants see an estimate of COVID-19 cases in their area and how many people in their county are reporting via the app. There is an option to share the app with others. The application is available for Android and Apple iOS operating systems.



**Figure 2.**

COVID Symptom Study app use, reported symptoms, and testing rates by state in the United States. More than 150,000 unique individuals have downloaded the application and shared clinical and demographic information, as well as daily symptoms and high-intensity occupational exposures across the United States (blue map). Population density of those presenting with any symptoms varied according to states with widespread reports of fatigue, cough, and diarrhea, followed by anosmia and relatively infrequently, fever (inlay). Reported testing rates vary between states, with all states reporting testing rates lower than 10%. This example of real-time visualization of data captured by the COVID Symptom Study may assist public health and government officials in reallocating resources, identifying areas with unmet testing needs, and detecting emerging hot spots earlier. Alaska is omitted from maps where there were not enough contributors.