The US COVID-19 Trends and Impact Survey: Continuous real-time measurement of COVID-19 symptoms, risks, protective behaviors, testing, and vaccination

Joshua A. Salomon[a,b,1], Alex Reinhart[c,d], Alyssa Bilinski[e], Eu Jing Chua[d,f], Wichada La Motte-Kerr[d,f], Minttu M. Rönn[g], Marissa B. Reitsma[a], Katherine A. Morris[h], Sarah LaRocca[h], Tamer H. Farag[i], Frauke Kreuter[j,k], Roni Rosenfeld[d,f], and Ryan J. Tibshirani[c,d,f]

[a]Department of Health Policy, Stanford University, Stanford, CA 94305; [b]Freeman Spogli Institute for International Studies, Stanford University, Stanford, CA 94305; [c]Department of Statistics and Data Science, Carnegie Mellon University, Pittsburgh, PA 15213; [d]Delphi Group, Carnegie Mellon University, Pittsburgh, PA 15213; [e]Department of Health Services, Policy and Practice, Brown School of Public Health, Providence, RI 02903; [f]Machine Learning Department, Carnegie Mellon University, Pittsburgh, PA 15213; [g]Department of Global Health and Population, Harvard T.H. Chan School of Public Health, Boston, MA 02115; [h]Demography and Survey Science, Meta, Menlo Park, CA 94025; [i]Health Partnerships, Meta, Menlo Park, CA 94025; [j]Joint Program in Survey Methodology, University of Maryland, College Park, MD 20742; and [k]Department of Statistics, Ludwig-Maximilians-Universität Munich, Munich 80539, Germany

The US COVID-19 Trends and Impact Survey (CTIS) is a large, cross-sectional, internet-based survey that has operated continuously since April 6, 2020. By inviting a random sample of Facebook active users each day, CTIS collects information about COVID-19 symptoms, risks, mitigating behaviors, mental health, testing, vaccination, and other key priorities. The large scale of the survey—over 20 million responses in its first year of operation—allows tracking of trends over short timescales and allows comparisons at fine demographic and geographic detail. The survey has been repeatedly revised to respond to emerging public health priorities. In this paper, we describe the survey methods and content and give examples of CTIS results that illuminate key patterns and trends and help answer high-priority policy questions relevant to the COVID-19 epidemic and response. These results demonstrate how large online surveys can provide continuous, real-time indicators of important outcomes that are not subject to public health reporting delays and backlogs. The CTIS offers high value as a supplement to official reporting data by supplying essential information about behaviors, attitudes toward policy and preventive measures, economic impacts, and other topics not reported in public health surveillance systems.

COVID-19 | SARS-CoV2 | survey

During 2020, the coronavirus disease 2019 (COVID-19) pandemic precipitated the need for new public health surveillance to inform urgent policy decisions. Effective pandemic policy-making requires information on a broad array of indicators, including local morbidity and mortality, preventive behaviors, healthcare capacity, and economic impacts. Given the critical importance of COVID-19 trends for policy, health departments set up routine public reporting systems for tracking cases, deaths, testing, and hospitalizations (1). However, supplemental data can both augment official reporting, for example by providing additional indicators of COVID-19 prevalence not subject to reporting delays and backlogs, and supply complementary information about public behavior, attitudes toward policy and preventive measures, mental health, economic impacts, and other items not observed in public health surveillance systems.

A number of efforts have used surveys to provide supplemental surveillance data. For example, symptom-tracking smartphone apps invite users to self-report symptoms, in some cases encouraging repeated participation to enable longitudinal tracking (2–4). Other surveys have addressed broader impacts of the pandemic, such as economic consequences (5). In this paper, we present findings from the Delphi Group at Carnegie Mellon University (CMU) US COVID-19 Trends and Impact Survey (CTIS), in partnership with Facebook, which has operated continuously since April 6, 2020 and collected over 20 million responses. [An international version of the survey is described in a companion paper in this theme issue (6).]

As the largest public health survey conducted in the United States to date, CTIS is designed to facilitate detailed demographic and geographic analyses, to track trends over time, and to accommodate rapid response to emerging priorities (7). A random sample of Facebook active users are invited each day to complete a questionnaire comprising survey items on symptoms, COVID testing, social distancing, vaccination, schooling, mental health, and economic security. The survey instrument has been updated frequently to incorporate new policy-relevant topics. Results are aggregated and made publicly available, and

Significance

The US COVID-19 Trends and Impact Survey (CTIS) has operated continuously since April 6, 2020, collecting over 20 million responses. As the largest public health survey conducted in the United States to date, CTIS was designed to facilitate detailed demographic and geographic analyses, track trends over time, and accommodate rapid revision to address emerging priorities. Using examples of CTIS results illuminating trends in symptoms, risks, mitigating behaviors, testing, and vaccination in relation to evolving high-priority policy questions over 12 mo of the pandemic, we illustrate the value of online surveys for tracking patterns and trends in COVID outcomes as an adjunct to official reporting, and showcase unique insights that would not be visible through traditional public health reporting.

[1]To whom correspondence may be addressed. Email: salomon1@stanford.edu.

microdata are available under institutional data use agreement, in both cases with fewer than 3 d of lag. These data provide information at a level of geographic and temporal detail that can supply essential inputs into short-term decision-making and longer-term strategic planning. These data also facilitate retrospective analysis of patterns, trends, and associations, supporting longer-term research on health policy decisions and the impacts of the pandemic.

In this study, we first compare COVID-19 indicators from CTIS with publicly reported case, hospitalization, and mortality data between April 2020 and April 2021. Despite potential limitations of our internet-based sample and the voluntary nature of the survey, we demonstrate high correspondence between the two, with CTIS less affected by holiday-related reporting anomalies. Second, we examine patterns and trends in symptoms, risks, mitigating behaviors, testing, and vaccination in US states and localities, in relation to evolving high-priority policy questions over 12 mo of the pandemic. The findings illustrate the value of online surveys for tracking patterns and trends in COVID-related outcomes as an adjunct to official reporting, while also showcasing insights that are only possible through a large-scale survey effort.

Methods

Sampling and Recruitment. The US CTIS launched on April 6, 2020 and has run continuously since that time, with an average of more than 350,000 people participating each week over the first year of operation. The survey is implemented by the Delphi Group at CMU, with participants recruited via the Facebook platform. Every day, Facebook invites a new sample of active users ages 18 y or older to participate in the survey. Facebook uses stratified random sampling within US states to randomly select a sample of its users to see the survey invitation at the top of their News Feed. Users who click on the invitation are taken to the CMU-administered survey hosted on Qualtrics. To ensure privacy, Facebook does not see any individual survey response during or after the data collection. The survey is available in English, Spanish, Brazilian Portuguese, Vietnamese, French, and simplified Chinese.

Survey Design. The survey instrument was deployed in multiple waves from launch through April 5, 2021, with contents of each survey version summarized in Table 1. Revisions are ongoing as new public health needs arise. A number of core items have been included consistently across all survey

versions, including questions about symptoms, contacts, and demographics. Key additions include items on mask wearing and occupation, added in September 2020, seasonal flu vaccination and schooling, added in November 2020, and COVID-19 vaccination, added in December 2020. As of April 5, 2021, the range of survey items spanned the following broad categories: household and individual symptoms, common comorbidities, contact patterns and mitigating behaviors, testing and diagnosis, worry and financial impact, schooling, vaccination, and demographics.

Full versions of all survey instruments can be found at https://cmu-delphi.github.io/delphi-epidata/symptom-survey/coding.html.

Weighting. Analytic weights have been developed to adjust for differences between Facebook users and the United States population, and to adjust for biases related to coverage and nonresponse (8). When Facebook links users to the survey, it generates a random unique identifier that is passed to CMU. For users who complete the survey, CMU returns the corresponding identifiers to Facebook, which then calculates analytic weights in two steps:

1. To adjust for nonresponse bias, Facebook calculates the inverse probability that sampled users complete the survey using their age, gender, and geographical variables, as reported on their Facebook profiles, as well as other characteristics known to correlate with nonresponse. The inverse probabilities are then used to create weights for responses, after which the survey sample reflects the active adult user population on Facebook.
2. To adjust for coverage bias, Facebook poststratifies the weights created in the first step so that the distribution of age, gender, and state or territory of residence in the survey sample reflects that of the general population.

The analytic weight does not identify the survey respondent. The weight for an individual is scaled to approximate the number of people in the adult population represented by that individual based on age, gender, location, and date. Facebook passes these weights to CMU. To protect respondent privacy, CMU cannot use these weights to identify specific Facebook users, and Facebook never receives individual survey responses and cannot link them to specific users.

Analysis. In this study we examined a range of different outcomes measured in the CTIS over the period April 6, 2020 to April 5, 2021. We analyzed both aggregated macrolevel data and individual-level data to highlight some of the key questions that may be examined using CTIS. Across the examples presented in this paper, we have stratified analyses by individual characteristics, including age, race/ethnicity (using categories consistent with National Center for Health Statistics), and occupation, and by geographic divisions, including Census region, Census division, state, and county.

Table 1. Summary of survey waves deployed between April 6, 2020 and April 5, 2021

Wave*	Contents	Start date	n (in millions)
1	Household and individual symptoms Common comorbidities Contacts with others Anxiety, depression Financial impact Demographics (age, gender)	April 6, 2020	1.1
2	New question: Symptoms among "people in your local community that you know personally" Minor textual revisions	April 15, 2020	2.6
3	Translated into simplified Chinese, Spanish, French, Brazilian Portuguese, Vietnamese Minor textual revisions	May 21, 2020	7.4
4	New questions: Medical care sought, COVID testing and results, mask wearing, social isolation Additional demographics, including race, ethnicity, occupation, education Textual revisions Some unused items removed	September 8, 2020	3.0
5	New questions: Seasonal flu vaccination, schooling, and school precautions Textual revisions	November 24, 2020	1.3
6	New questions: Vaccine intent. Vaccine status item enabled on January 6, 2021	December 19, 2020	1.2
7	Textual revisions to vaccine intent items	January 12, 2021	1.3
8	New questions: Reasons for vaccine hesitancy, vaccine dosing Minor textual revisions	February 8, 2021	0.9
10	New questions: Appointments for COVID vaccines, information about getting vaccinated Textual revisions	March 2, 2021	1.4

*There was no Wave 9 survey. The numbering of waves skipped from 8 to 10 to synchronize numbering conventions with the international version of the survey.

Salomon et al.
The US COVID-19 Trends and Impact Survey: Continuous real-time measurement of
COVID-19 symptoms, risks, protective behaviors, testing, and vaccination

To examine the representativeness of the study sample, we compared characteristics of the sample to data from the American Community Survey 2019 (ACS) supplemental estimates.

We evaluated reported symptoms and symptom patterns in comparison to surveillance data on confirmed COVID-19 hospitalizations from the Department of Health and Human Services (9) and reported COVID-19 cases and mortality aggregated by the Johns Hopkins University Center for Systems Science and Engineering (10). To summarize relevant symptom patterns, we defined "COVID-like illness" (CLI) as reporting a fever of at least 100 °F, along with cough, shortness of breath, or difficulty breathing, in line with a working definition of CLI used for syndromic surveillance purposes beginning in early 2020. A second indicator, which we call "CLI in Community," was based on responses to an item on the survey that asks whether respondents know someone personally in their community who is ill with COVID-like symptoms. We also compared reported diagnoses in CTIS to cumulative diagnoses in surveillance data.

To illustrate the utility of individual-level data to provide detailed information on characteristics that may be relevant to transmission risk, we compared reported diagnoses and reports of working outside of the home while symptomatic across occupational categories. To evaluate time trends in risk exposures and mitigating behaviors, we examined reported contacts, mask use, and use of public transit over time, across a range of stratifying variables, including counties grouped by levels on the Centers for Disease Control and Prevention (CDC) Social Vulnerability Index (SVI) (11), which is a composite measure constructed based on 15 social variables measured at the census tract level, age groups, and Census regions.

Finally, as an illustration of the information value of the survey at various levels of granularity, we assessed reported COVID-19 vaccination intent stratified by individual characteristics and across counties.

The study was approved by the CMU Institutional Review Board, under protocol STUDY2020_00000162. All respondents gave informed consent before participating in the survey.

Results

Characteristics of the Study Sample. As of April 5, 2021, a total of 20.2 million responses had been collected in the US CTIS. Table 2 summarizes characteristics of the survey respondents. Compared to the weighted sample, the unweighted sample had a higher proportion of women (66% vs. 52%) and a slightly higher proportion of respondents between ages 25 and 64 y (72% vs. 68%). Household size and prevalence of at least one comorbidity were similar in the unweighted and weighted samples.

Compared to 2019 ACS supplemental estimates, the weighted survey sample slightly overrepresented women, but had a broadly comparable age distribution and matched the ACS distribution across geographic regions. The weighted sample included a larger proportion of respondents with greater than a high school education, and a much smaller proportion

Table 2. Characteristics of the study sample, compared to 2019 ACS supplemental estimates

	Number	Unweighted proportion (%)	Weighted proportion (%)	Census proportion (%)
All responses	20,249,152			
Gender				
Female	11,409,227	66.3	52.3	50.8
Male	5,613,674	32.6	46.2	49.2
Nonbinary/self-described	174,124	1.0	1.5	—
Age groups				
18–24	1,001,345	5.8	10.7	11.9
25–34	2,856,685	16.5	16.4	17.9
35–44	3,212,187	18.5	16.5	16.4
45–54	3,129,334	18.1	17.7	16.0
55–64	3,337,427	19.3	17.5	16.6
65–74	2,752,379	15.9	15.3	21.2*
≥75	1,035,551	6.0	6.0	
Education[†]				
Less than high school	226,284	3.1	4.0	12.0
High school or equivalent	1,152,727	16.0	17.2	27.1
Some college, no degree	1,744,155	24.1	24.5	20.4
Associate's degree	829,618	11.5	11.2	8.5
Bachelor's degree	1,765,207	24.4	23.6	19.7
Graduate or professional degree	1,505,785	20.8	19.5	12.3
Region				
Northeast	3,426,497	17.4	17.6	17.4
Midwest	4,791,585	24.3	20.9	20.8
South	7,328,178	37.2	37.9	38.0
West	4,160,577	21.1	23.6	23.8
Household size				
1	2,698,400	13.7	13.3	
2	6,668,445	33.9	31.9	
3–5	8,489,946	43.1	43.7	
6–10	1,476,637	7.5	8.7	
>10	352,267	1.8	2.4	
Date of completion				
April–June 2020	6,814,488	33.7	23.6	
July–September 2020	5,280,596	26.1	25.2	
October–December 2020	3,832,698	18.9	25.2	
January–March 2021	4,131,800	20.4	24.7	
April 2021	189,570	0.9	1.4	
At least 1 comorbidity	9,949,181	53.7	52.0	

*Value reported for 65 y and older in ACS.
[†]Value reported for adults 25 and older in ACS; only collected in CTIS beginning in Wave 4.

Salomon et al.
The US COVID-19 Trends and Impact Survey: Continuous real-time measurement of COVID-19 symptoms, risks, protective behaviors, testing, and vaccination

PNAS | 3 of 9
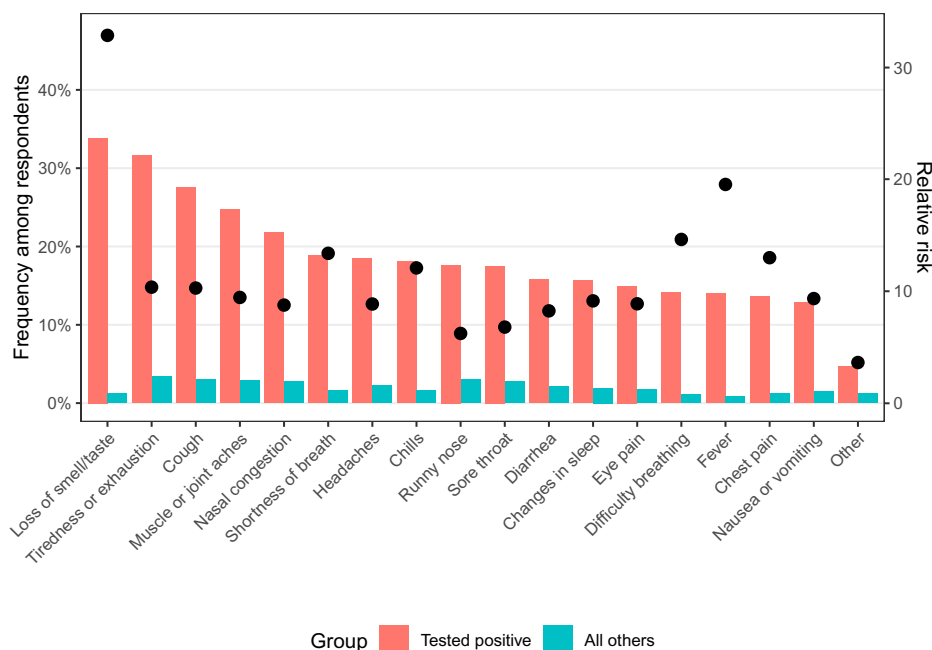https://doi.org/10.1073/pnas.2111454118

Fig. 1. Frequency of reported new or unusual symptoms, pooled over respondents to the CTIS, September 8, 2020 to April 5, 2021. Respondents are grouped by whether they indicated they tested positive in the past 14 d. Dots indicate the ratio of frequency among those who tested positive compared to all others.

with less than a high school education, suggesting the presence of a sampling or response bias correlated with education. This bias has remained consistent over time. As the weights provided by Facebook do not account for education, the weighting did not correct this bias.

COVID-19 Symptoms and Diagnoses. A large fraction of daily respondents reported new or unusual symptoms consistent with COVID-19 (Fig. 1). The most common single new or unusual

symptom among all respondents was "tiredness or exhaustion," with a prevalence of 3.9% over Waves 4 through 10. Patterns of symptoms were notably different among the subset of respondents who reported testing positive for COVID-19, compared to all other respondents, including a substantially higher probability of reporting loss of smell or taste (34% compared to 1.2%).

Fig. 2 compares time series for three indicators from the CTIS (reported anosmia, CLI, and CLI in community) against
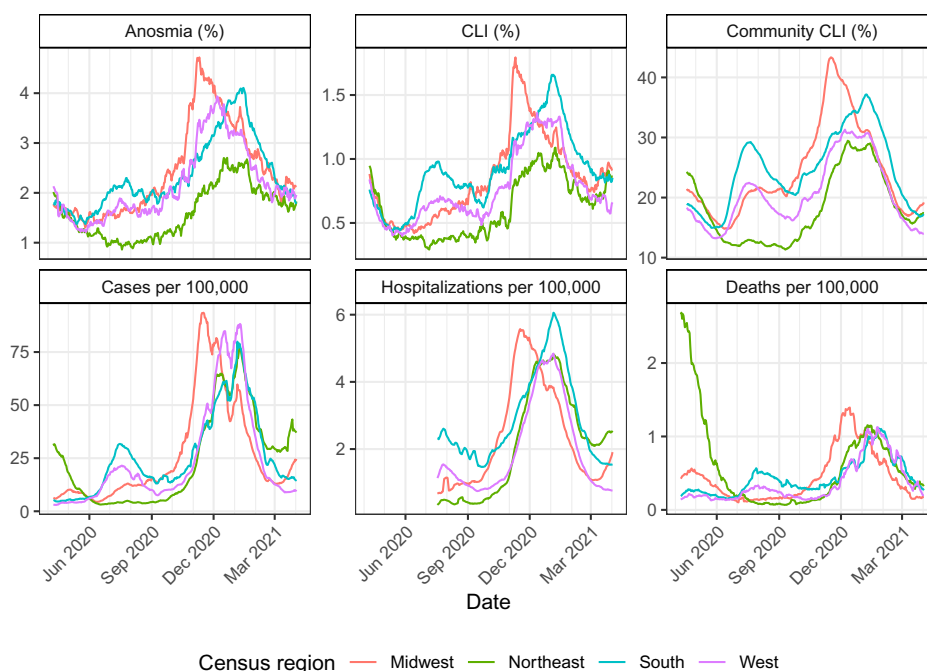

Fig. 2. Trends in anosmia, CLI, CLI in community, confirmed cases, hospitalizations, and deaths, by Census region, April 6, 2020 to April 5, 2021.
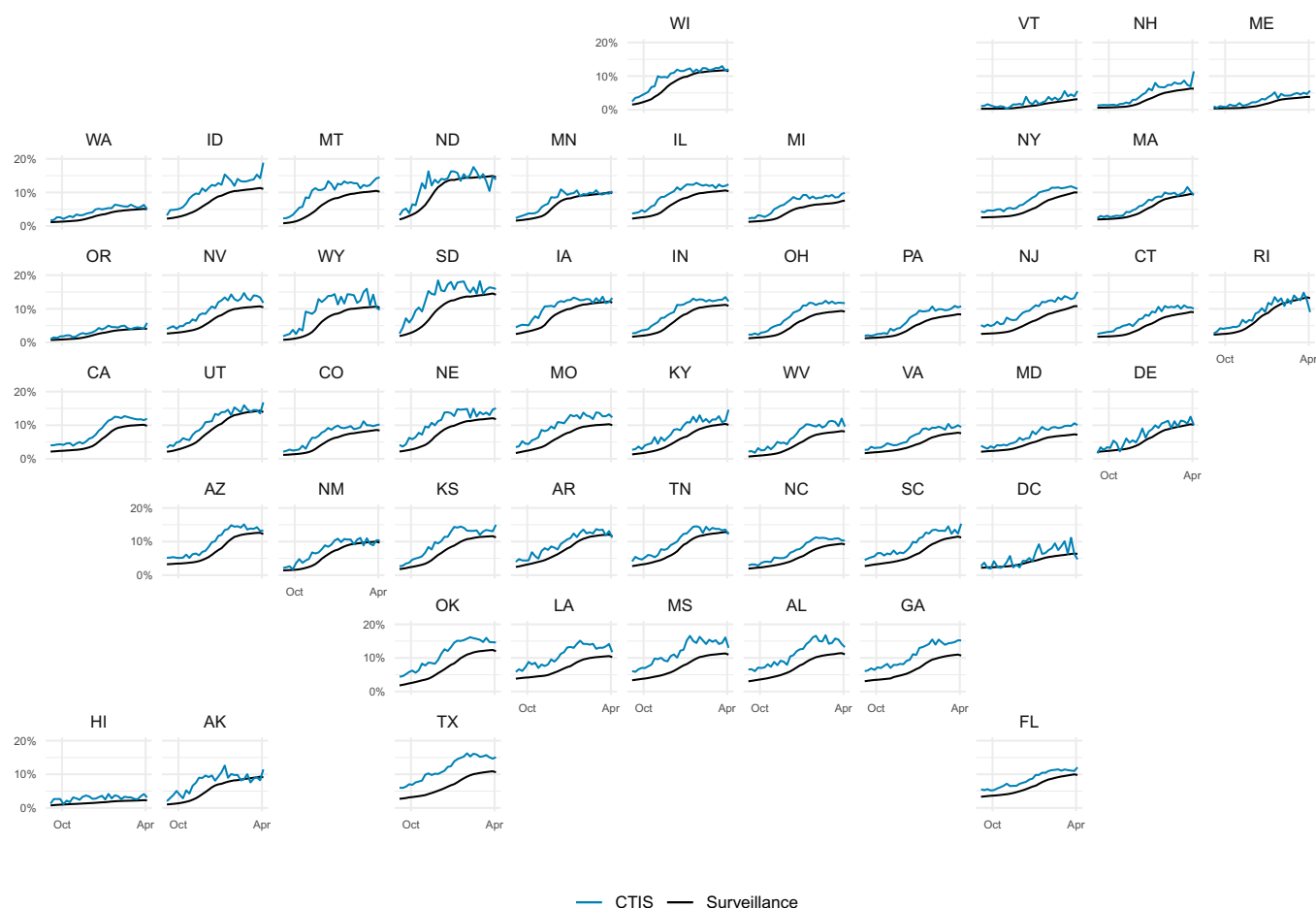
Fig. 3. Comparison of proportion of respondents in CTIS reporting ever having tested positive for COVID-19 and cumulative proportion of adult population with confirmed COVID diagnosis, by state, September 8, 2020 through April 5, 2021.

the three main surveillance indicators that have been used to monitor trends in the epidemic (confirmed cases, hospitalizations, and deaths) stratified by Census region. Over the period April 6, 2020 to April 5, 2021, the three survey indicators tracked both broad temporal trends and regional patterns in the surveillance indicators, and several notable features are evident in the comparison. First, the survey-based indicators were less susceptible to the daily fluctuations and reporting anomalies that appeared in cases and deaths, including abrupt discontinuities around certain holiday periods. Second, trends and patterns in anosmia were similar to patterns in CLI, and the anosmia series provided a closer match than the other two survey indicators to the trends and patterns observed in COVID-19 hospitalizations, mirroring temporal peaks and comparative levels across regions over different waves of the epidemic. Third, CLI in community provided the most temporally stable signals while also expressing broad differences over time and space that were generally similar to those in other indicators. In a companion paper, we performed extensive correlation analyses between reported cases and various auxiliary indicators, including the survey-based CLI and CLI in community signals (12), showing strong correlations between cases and these two survey signals during much of the pandemic.

The CTIS includes questions about testing and diagnosis, which since September 8, 2020 have been asked of all respondents. Fig. 3 compares weekly CTIS estimates of the proportion of adults reporting that they have ever had a positive test for COVID-19 against cumulative diagnoses from surveillance

reports by state in the same week. State surveillance reports were adjusted, using 2019 population estimates and CDC line-level demographic data on confirmed COVID-19 cases, to produce estimated diagnosis rates among the state's population over age 18 y. As of April 5, 2021, reported diagnoses in the survey ranged from 3.1% in Hawaii to 19% in Idaho, and the correlation between survey reported diagnoses and surveillance reports at the state level was 0.83, indicating strong convergent validity.

Transmission Risk by Individual Characteristics. Since Wave 4, the CTIS has included questions about occupation, which can offer valuable insights into exposures among essential workers and also supply signals of where transmission may be concentrated. Using responses from January 2021, we examined the probability of reporting a positive COVID-19 test across different occupation categories, as well as the probability of reporting working outside the home while having symptoms consistent with COVID-19 (Table 3). Substantial heterogeneity appeared across broad groups of occupation. The large proportion of people reporting never having been tested indicates the limitations of passive surveillance. Combining questions on symptoms, testing, and working into a single indicator, we examined the fraction of people who reported both working outside the home and currently having atypical symptoms; results ranged from more than 15% for respondents in food preparation and serving related occupations, to 4% of those in arts, design, entertainment, sports, and media.

Salomon et al.
The US COVID-19 Trends and Impact Survey: Continuous real-time measurement of COVID-19 symptoms, risks, protective behaviors, testing, and vaccination

PNAS | 5 of 9
https://doi.org/10.1073/pnas.2111454118

Table 3. Reporting testing, symptoms, and working outside the home, by reported occupation category, in January 2021

Occupation group	% Tested positive	% Working with symptoms	% Working outside and never tested	n
Arts, design, entertainment, sports, and media	7.2	4.0	17.1	20,585
Building and grounds cleaning and maintenance	11.9	10.8	45.8	12,528
Community and social service	13.1	9.0	23.3	28,223
Construction and extraction	11.8	11.3	47.7	10,528
Education, training, and library	10.5	6.8	24.0	72,098
Food preparation and serving related	13.6	15.5	39.4	30,817
Healthcare practitioners and technicians	15.1	9.6	25.4	70,793
Healthcare support	15.2	9.2	22.8	45,084
Installation, maintenance, and repair	10.5	10.9	49.2	15,511
Office and administrative support	11.0	6.2	24.4	84,285
Other	9.9	6.5	25.9	165,719
Personal care and service	12.1	9.2	32.6	15,115
Production	14.1	12.6	42.1	23,149
Protective service	14.6	11.8	33.5	8,314
Sales and related	11.7	10.8	36.1	63,066
Transportation and material moving	11.2	10.5	48.3	23,013

*There was no Wave 9 survey. The numbering of waves skipped from 8 to 10 to synchronize numbering conventions with the international version of the survey.

Mitigating Behaviors and Policy Analysis. A core set of questions included since the launch of the survey has addressed contacts and preventive behaviors. The survey has been amended over time to augment these items, with the addition of questions on mask use and specific high-risk behaviors in September 2020. In the context of recurrent surges in COVID-19 around the country over the course of 2020 and 2021, these items have illuminated how contacts and mitigating behaviors can shift in response to changes in local COVID-19 risk, sometimes preceding policy changes. For example, Fig. 4 shows selected variables relating to contacts and preventive behaviors over the period September 2020 to April 2021. Responses indicate sharp increases in risk-reducing behaviors during November and December as cases surged—including reduced contacts, increased use of masks, and reduced use of public transit—followed by relaxation of mitigating behaviors over the period January to April 2021 as cases fell.

Individual-level data allow for geographically detailed analysis that can also be disaggregated by demographic features. In Fig. 4,
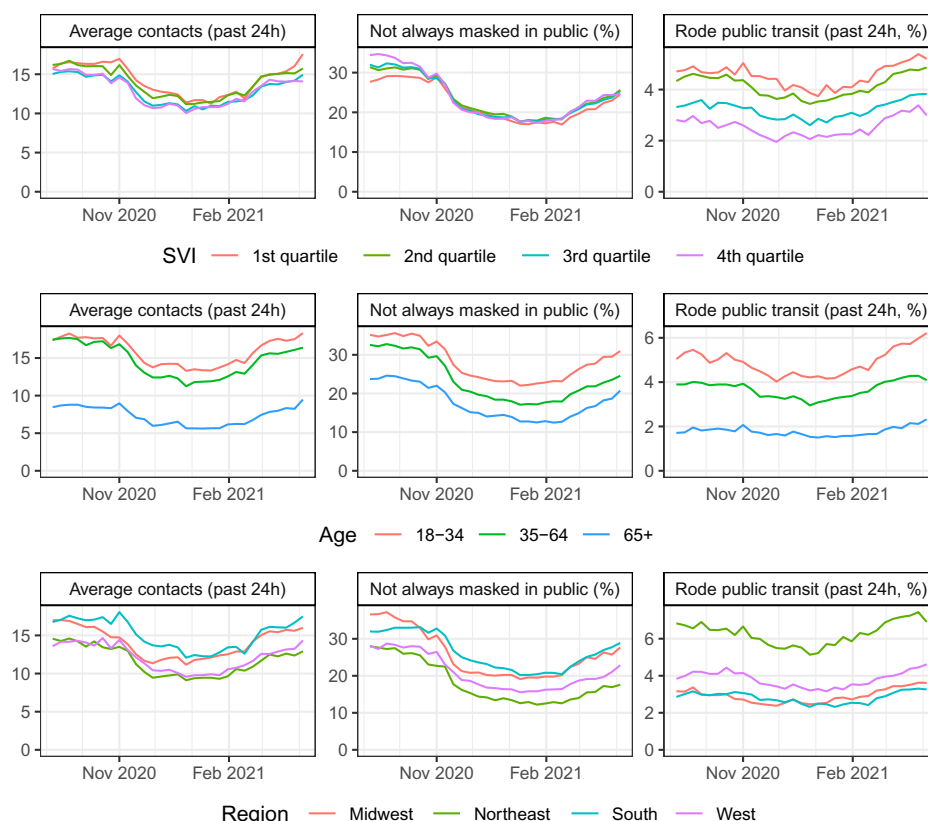

Fig. 4. Contacts, mask use, and use of public transport, by quartile of counties grouped by the CDC SVI (*Top*), age group (*Middle*), and Census region (*Bottom*), September 8, 2020 to April 5, 2021.

Salomon et al.
The US COVID-19 Trends and Impact Survey: Continuous real-time measurement of
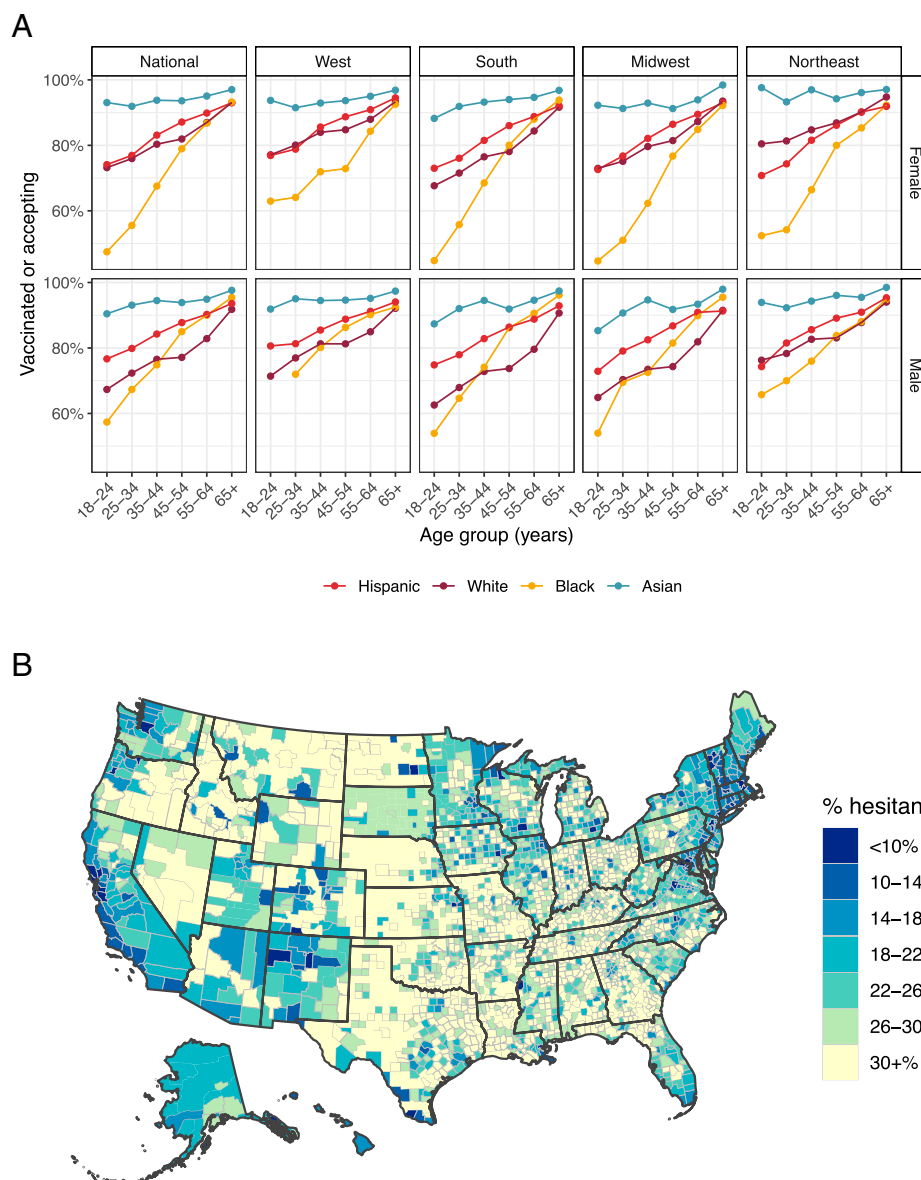COVID-19 symptoms, risks, protective behaviors, testing, and vaccination

A

B

Fig. 5. Reported vaccine acceptance and hesitancy by age group, race/ethnicity, gender, and Census region (*A*) and by county (*B*) during March 2021. Note: Results are pooled over the period March 1, 2021 to April 5, 2021. The map displays results computed for specific counties having at least 50 responses recorded over that period, with all other counties in a state combined into a residual group.

results are stratified in three different ways to illustrate this: by quartiles of the CDC SVI, by age, and by US Census region. There were minimal differences across counties grouped by the SVI in reported mask use, moderately higher contacts among those living in more vulnerable communities, and substantially higher use of public transit in more vulnerable counties. The second row shows age differences, which indicate a pronounced gradient of higher risk mitigation among older respondents, especially with respect to reduced contacts. The third row describes regional patterns that vary across indicators, with higher contacts and lower mask use in the South and Midwest regions compared to the Northeast and West, but greater use of public transit in the Northeast and West compared to South and Midwest.

Vaccination and Vaccine Acceptance. Since December 19, 2020, the CTIS has included questions on vaccination intent, and since January 6, 2021, the survey has asked about vaccination

status. The combination of geographic and demographic resolution in the survey allows a uniquely detailed view on vaccination acceptance and hesitancy across different United States population groups. Fig. 5*A* displays results by age group, race/ethnicity, gender, and Census region, pointing to high levels of acceptance among older respondents in all categories, but lower and more variable results at younger ages. (Respondents may identify as nonbinary or self-describe their gender, but this group was typically too small to break out and report reliable hesitancy estimates by region.) Fig. 5*B* shows the percentage of respondents indicating that they would probably not or definitely not get vaccinated across United States counties, indicating regional patterns but also high variability across counties within a given state. As the vaccination campaign slows across the country, high-resolution information on vaccine acceptance can inform policies that aim to increase uptake toward the goal of high levels of population immunity against COVID-19.

Salomon et al.
The US COVID-19 Trends and Impact Survey: Continuous real-time measurement of COVID-19 symptoms, risks, protective behaviors, testing, and vaccination

PNAS | 7 of 9
https://doi.org/10.1073/pnas.2111454118

Discussion

As SARS-CoV-2 spread throughout the United States during 2020 and into 2021, and policy makers faced decisions that would profoundly impact all sectors of society, the breadth and depth of information needed to support these decisions vastly exceeded the availability of data collected through existing surveillance systems designed to capture reported COVID-19 cases, hospitalizations, and deaths. A number of efforts to fill the urgent need for additional information relied on novel data collection and dissemination platforms that leveraged mobile phone technology and new media. In this study, we describe one of these efforts, the CTIS, which is the largest continuous health survey ever conducted in the United States, in operation since April 6, 2020, with more than 20 million responses collected over the first year of operation.

Comparisons to routine sources of surveillance information point to both the face validity and incremental value of the CTIS. Time trends and geographic patterns in COVID-19 outcomes measured in the CTIS—including specific symptoms strongly associated with SARS-CoV-2 infection, such as anosmia, syndromic patterns, such as COVID-like-illness, and the novel CTIS measure of CLI in community—mirror broad temporal and spatial features in standard surveillance measures on confirmed COVID-19 cases, hospitalizations, and deaths, while in many cases avoiding data artifacts and reporting anomalies that affect the official measures.

In this study we have highlighted several examples of how attributes of CTIS give it particular value and salience as an information platform for public health policy. The scope, scale, and recruitment strategy used in the survey support analysis at high geographic and temporal resolution, allowing detection of local trends on short timescales not available through other surveys, and accommodate a high level of stratification on relevant individual characteristics. Several examples illustrate the benefits of this granularity, including the ability to compare risks and preventive behaviors by occupational category, with further stratification possible by demography and geographic location; ability to describe variation in intentions; and use of key mitigating measures, including physical distancing, masking, and vaccination. Regular updating of the survey has enabled the survey content to adapt alongside the evolving policy response, for example through addition of survey items on mask use in September 2020, school mitigation strategies in November 2020, and vaccination in December 2020.

Other studies have used data from the CTIS to answer specific questions about key COVID-19 impacts and policies. A number of studies have analyzed relationships between reported risk-mitigating behaviors in the CTIS and other outcomes. For example, Rebeiro et al. (13) examined reported mask-wearing behavior as an outcome in relation to statewide mask-wearing requirements. Rader et al. (14) examined the relationship between mask wearing and physical distancing as measured in CTIS and measures of community transmission. Bilinski et al. (15) assessed trends across states in a number of indicators on risk perception and preventive behaviors in relation to COVID-19 case rates. Other studies have used CTIS measures to explore correlates of variation in risk. For example, Flaxman et al. (16) computed relative infection rates for healthcare workers vs. other respondents using information from the survey on occupation, testing, and test results. Lessler et al. (17) have assessed reported risks of COVID-19–related outcomes, including CLI, anosmia, or a positive COVID-19 test, in relation to whether a household includes a child who attends in-person schooling, and reported school-based mitigation measures.

Symptom measures from the survey have also been used to aid in forecasting of COVID cases and deaths. Through the COVID-19 Forecast Hub, the CDC collects standardized forecasts from dozens of teams. Rodríguez et al. (18) incorporated symptom surveillance data from CTIS into a deep-learning framework for real-time forecasting. In a companion paper in this issue (19), we demonstrate that symptom surveillance data and other auxiliary data streams (such as medical insurance claims) can improve forecasting and hotspot prediction accuracy over short (1 to 3 wk) time intervals.

Several limitations are important to note. First, because the survey uses Facebook active users as its sampling frame and because participation in the survey is strictly voluntary, respondents may not be fully representative of the United States population, despite incorporation of survey weights, which adjust for nonresponse and coverage biases based on a limited number of covariates. Comparison to the ACS indicates that our sample overrepresents respondents who are college-educated. Research users of the survey microdata can use additional demographic or other survey variables to construct improved poststratification adjustments to correct this for their purposes. However, any nonresponse biases not accounted for by Facebook's nonresponse weights would be much more difficult to correct.

Additionally, many of the outcome measures related to COVID-19 are based on self-reports, which may diverge from more objective measures due to recall bias, social desirability bias, and other sources of survey bias and measurement error. On the other hand, broad comparisons of indicators, such as cumulative COVID-19 diagnoses, suggest that measurement of key COVID-19 outcomes are relatively robust to response biases that may be present in the sample. Ultimately, the value of such a large-scale survey is not in accuracy afforded by its sample size, since survey biases persist no matter the size of the survey; smaller surveys, more carefully constructed to reduce sampling biases, would likely yield more accurate estimates (20). Instead, since these survey biases are unlikely to change rapidly in time or in space, CTIS can accurately track trends in key signals, even if the daily point estimates are systematically biased. This is demonstrated by the strong correlations between survey estimates of CLI in community and reported COVID case rates, for example; while CLI in community is not an unbiased population estimate of COVID case rates, it nonetheless provides useful information about trends in cases. The principal value of CTIS is hence in the detailed spatial and demographic comparisons it makes possible, and in its ability to track changes continuously over time and correlate them with key outcome measures.

Although CTIS was initially designed with a relatively limited scope, including a particular focus on syndromic surveillance, its value has ultimately derived in large part from its flexibility as a surveillance platform that can be rapidly adapted to changing information needs. Running a survey of this size has involved many challenges, particularly as it expanded to include key measures of public knowledge, attitudes, and behaviors, and as public health needs evolved continuously during the pandemic. Despite these challenges, however, CTIS has provided both a valuable public information resource during a global health emergency, as well as a potential model for ongoing health surveillance needs. Similar online surveys are likely to play important roles in future epidemics and pandemics by supplementing public reporting systems with information that is difficult to gather any other way.

8 of 9 | PNAS
https://doi.org/10.1073/pnas.2111454118

Salomon et al.
The US COVID-19 Trends and Impact Survey: Continuous real-time measurement of COVID-19 symptoms, risks, protective behaviors, testing, and vaccination

presented in an interactive online dashboard (https://delphi.cmu.edu/covidcast/survey-results/?date=20211103). Demographic breakdowns of key variables over time are available for public download at https://cmu-delphi.github.io/delphi-epidata/symptom-survey/contingency-tables.html. In order to safeguard the privacy of survey respondents, aggregate data are publicly presented only for cells containing at least 100 respondents. All code used to generate figures and tables in the paper is deposited with Zenodo (https://doi.org/10.5281/zenodo.5639567).

1. Centers for Disease Control and Prevention, COVID Data Tracker. https://covid.cdc.gov/covid-data-tracker/. Accessed 16 June 2021.
2. W. E. Allen *et al.*, Population-scale longitudinal mapping of COVID-19 symptoms, behaviour and testing. *Nat. Hum. Behav.* **4**, 972–982 (2020).
3. A. T. Chan, J. S. Brownstein, Putting the public back in public health—Surveying symptoms of Covid-19. *N. Engl. J. Med.* **383**, e45 (2020).
4. C. Menni *et al.*, Real-time tracking of self-reported symptoms to predict potential COVID-19. *Nat. Med.* **26**, 1037–1040 (2020).
5. United States Census Bureau, Household Pulse Survey (COVID-19). https://www.census.gov/programs-surveys/household-pulse-survey.html. Accessed 16 June 2021.
6. C. M. Astley *et al.*, Global monitoring of the impact of the COVID-19 pandemic through online surveys sampled from the Facebook user base. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2111455118 (2021).
7. F. Kreuter *et al.*, Partnering with a global platform to inform research and public policy making. *Surv. Res. Methods* **14**, 159–163 (2020).
8. N. Barkay *et al.*, Weights and methodology brief for the COVID-19 symptom survey by University of Maryland and Carnegie Mellon University, in partnership with Facebook. arXiv [Preprint] (2020). https://arxiv.org/abs/2009.14675 (Accessed 16 June 2021).
9. United States Department of Health and Human Services, COVID-19 reported patient impact and hospital capacity by state timeseries. https://healthdata.gov/dataset/COVID-19-Reported-Patient-Impact-and-Hospital-Capa/qqte-vkut. Accessed 16 June 2021.
10. E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **20**, 533–534 (2020).
11. Agency for Toxic Substances and Disease Registry, CDC/ATSDR Social Vulnerability Index 2021.https://www.atsdr.cdc.gov/placeandhealth/svi/index.html. Accessed 16 June 2021.
12. A. Reinhart *et al.*, An open repository of real-time COVID-19 indicators. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2111452118 (2021).
13. P. F. Rebeiro, D. M. Aronoff, M. K. Smith, The impact of state mask-wearing requirements on the growth of COVID-19 cases in the United States. *Clin. Infect. Dis.* **73**, 1703–1706 (2021).
14. B. Rader *et al.*, Mask-wearing and control of SARS-CoV-2 transmission in the USA: A cross-sectional study. *Lancet Digit. Health* **3**, e148–e157 (2021).
15. A. Bilinski, E. Emanuel, J. A. Salomon, A. Venkataramani, Better late than never: Trends in COVID-19 infection rates, risk perceptions, and behavioral responses in the USA. *J. Gen. Intern. Med.* **36**, 1825–1828 (2021).
16. A. Flaxman, D. Henning, H. Duber, The relative incidence of COVID-19 in healthcare workers versus non-healthcare workers: Evidence from a web-based survey of Facebook users in the United States. *Gates Open Res.* **4**, 174 (2020).
17. J. Lessler *et al.*, Household COVID-19 risk and in-person schooling. *Science* **372**, 1092–1097 (2021).
18. A. Rodríguez *et al.*, DeepCOVID: An operational deep learning-driven framework for explainable real-time COVID-19 forecasting. *Proc. Conf. AAAI Artif. Intell.* **35**, 15393–15400 (2021).
19. D. J. McDonald *et al.*, Can auxiliary indicators improve COVID-19 forecasting and hotspot prediction? *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2111453118 (2021).
20. X. Meng, Statistical paradises and paradoxes in big data (I): Law of large populations, big data paradox, and the 2016 US presidential election. *Ann. Appl. Stat.* **12**, 685–726 (2018).

Salomon et al.
The US COVID-19 Trends and Impact Survey: Continuous real-time measurement of COVID-19 symptoms, risks, protective behaviors, testing, and vaccination

PNAS | 9 of 9
https://doi.org/10.1073/pnas.2111454118

Supplementary Information for

Global Monitoring of the Impact of the COVID-19 Pandemic through Online Surveys
Sampled from the Facebook User Base

Christina M. Astley[a,b,c,d,1], Gaurav Tuli[b,1], Kimberly A. Mc Cord[b,1], Emily L. Cohn[b], Benjamin Rader[b,e], Tanner J. Varrelman[b], Samantha L. Chiu[f], Xiaoyi Deng[f], Kathleen Stewart[g], Tamer H. Farag[h], Kris M. Barkume[h], Sarah LaRocca[h], Katherine A. Morris[h], Frauke Kreuter[f,i,3], John S. Brownstein[b,c,3]

[a] Division of Endocrinology, Boston Children's Hospital, Boston, MA 02115;
[b] Computational Epidemiology Lab, Boston Children's Hospital, Boston, MA 02115;
[c] Harvard Medical School, Boston, MA 02115;
[d] Broad Institute of Harvard and MIT, Cambridge, MA 02142;
[e] Department of Epidemiology, Boston University, Boston, MA 02118;
[f] Joint Program in Survey Methodology, University of Maryland, College Park, MD 20742;
[g] Center for Geospatial Information Science, University of Maryland, College Park, MD 20742;
[h] Meta, Menlo Park, CA 94025; and
[i] Department of Statistics, Ludwig-Maximilians-Universität, Munich 80539, Germany

[1] C.M.A., G.T., and K.A. Mc Cord contributed equally to this work.
[2] To whom correspondence may be addressed. Email: christina.astley@childrens.harvard.edu.
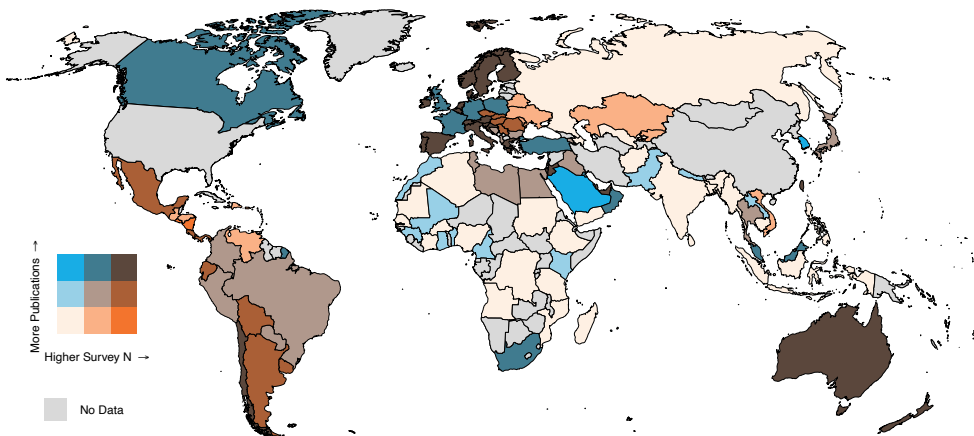[3] F.K. and J.S.B. contributed equally to this work.

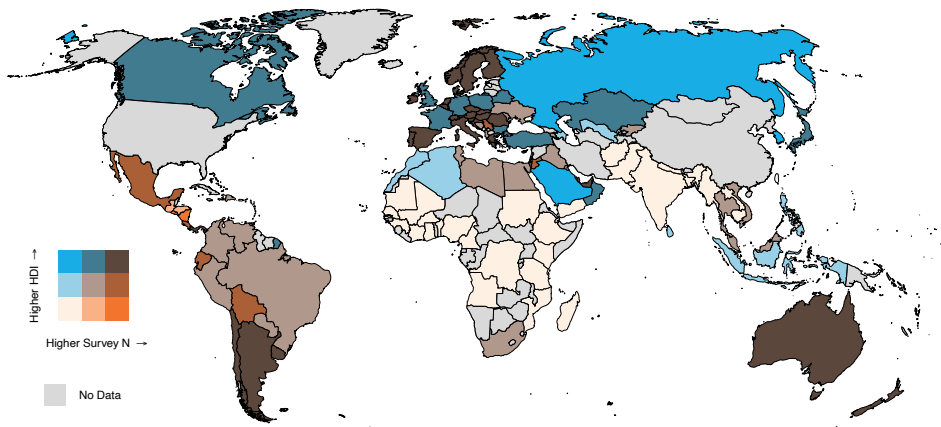**This PDF file includes:**

Supplementary Figures 1 to 4
Supplementary Tables 1 to 2

Supplementary Figure 1: UMD-CTIS coverage globally. Surveys per population (orange) during the study period (April 27, 2020 - December 20, 2020) in relation to COVID-19-related biomedical citations per population (blue) for the N=144 countries and territories with survey weights. Locales not sampled by UMD-CTIS or without survey weights are shown in grey. Surveys per population was moderate to high for much of Europe, and Central and South America, and low for much of Africa and Asia. Carnegie Mellon University in collaboration with Facebook oversees a similar, but not identical, study (CMU-CTIS) in the United States. The coverage of this syndromic surveillance platform spans a broad range of locales globally. (a) The survey serves many locales that have been relatively under-represented in COVID-19 biomedical literature to date. Citation counts pulled from Pubmed, Medxriv and Bioxriv on July 7, 2021 using the search logic ("COVID-19" OR "SARS-CoV-2" OR "2019-nCoV" OR "ncov2019") AND (Locale Name OR Alternative Locale Name). Biomedical citations are a very rough proxy for COVID-19-related research, and are not equivalent to availability of accurate and timely public health data in a country. (b) We also show the survey coverage in relation to Health Development Index (HDI, Our World In Data. https://ourworldindata.org/ Accessed October 7, 2021). Importantly, the platform requires no logistical or financial support on the part of local public health officials. For those locales for which additional syndromic surveillance data would be beneficial, UMD-CTIS should be considered and evaluated in the context of that specific locale's needs at that time.
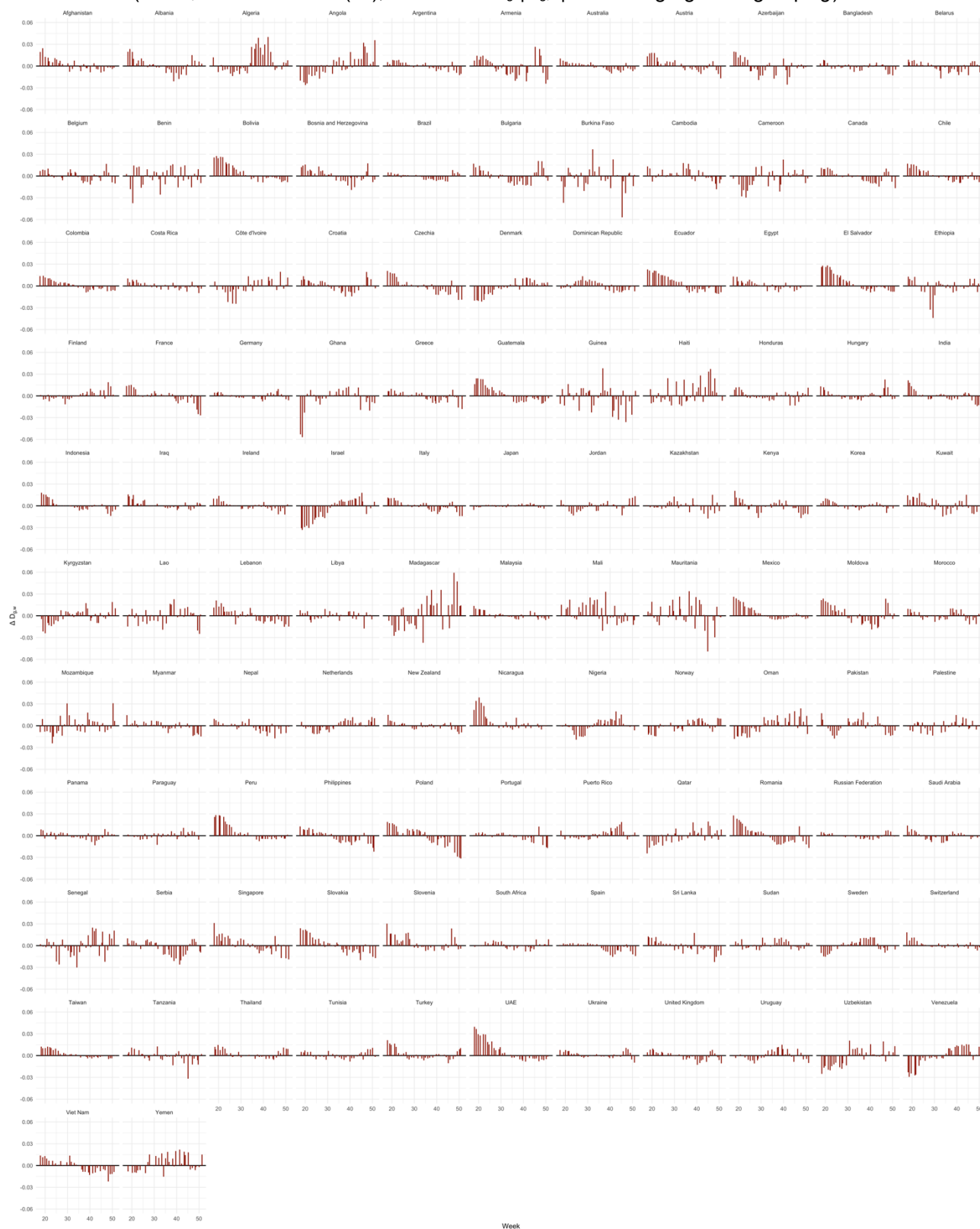
(a) UMD-CTIS coverage in relation to biomedical literature coverage

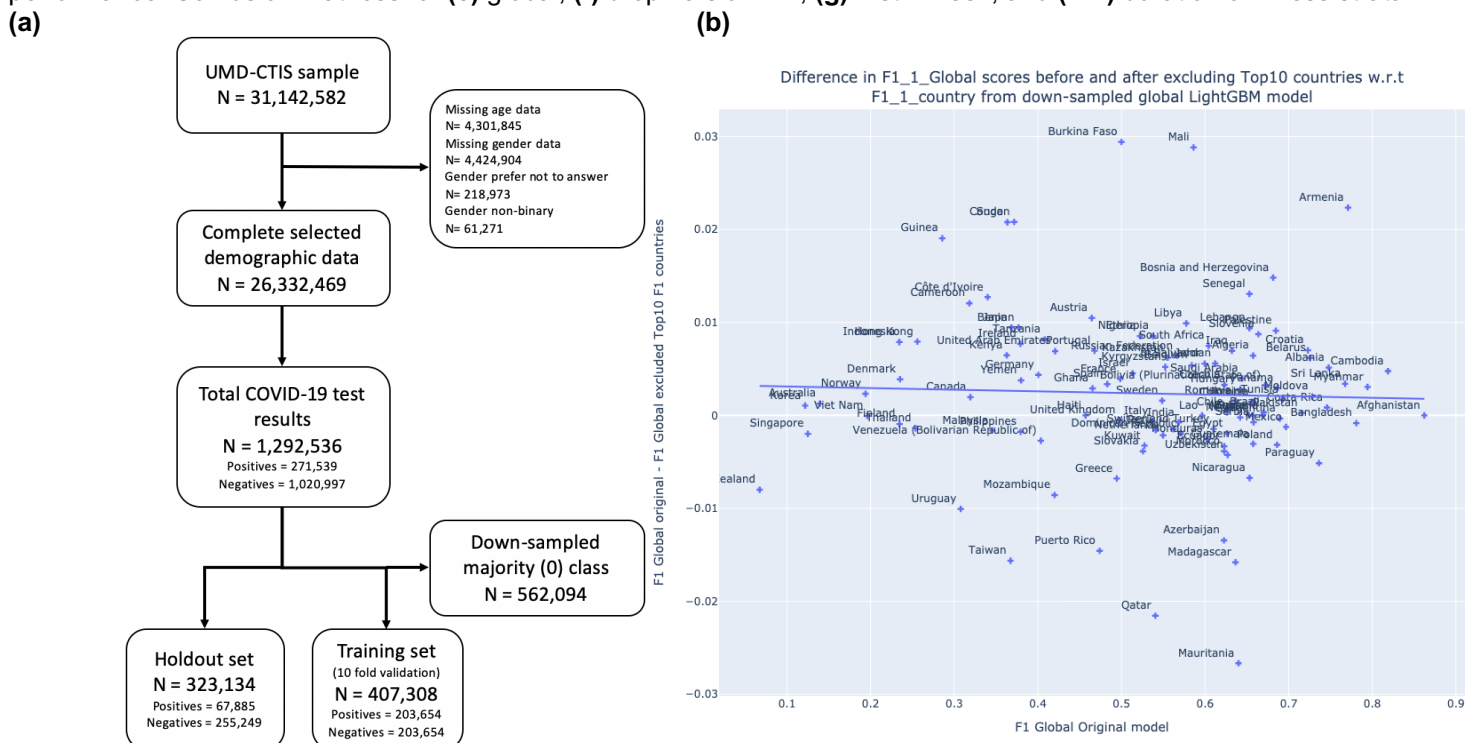(b) UMD-CTIS coverage in relation to HDI

Supplementary Figure 2: Within-locale variation in mean absolute differences from census over the study period. Within-locale variation in mean absolute differences across age-gender groups, by week, relative to the locale median across weeks (i.e. $\Delta_{\delta,w} = \delta_w - \text{median}(\delta_w)$, where $\delta_w = \Sigma_g |D_{g,w}|/6$ for all age-gender groups g).
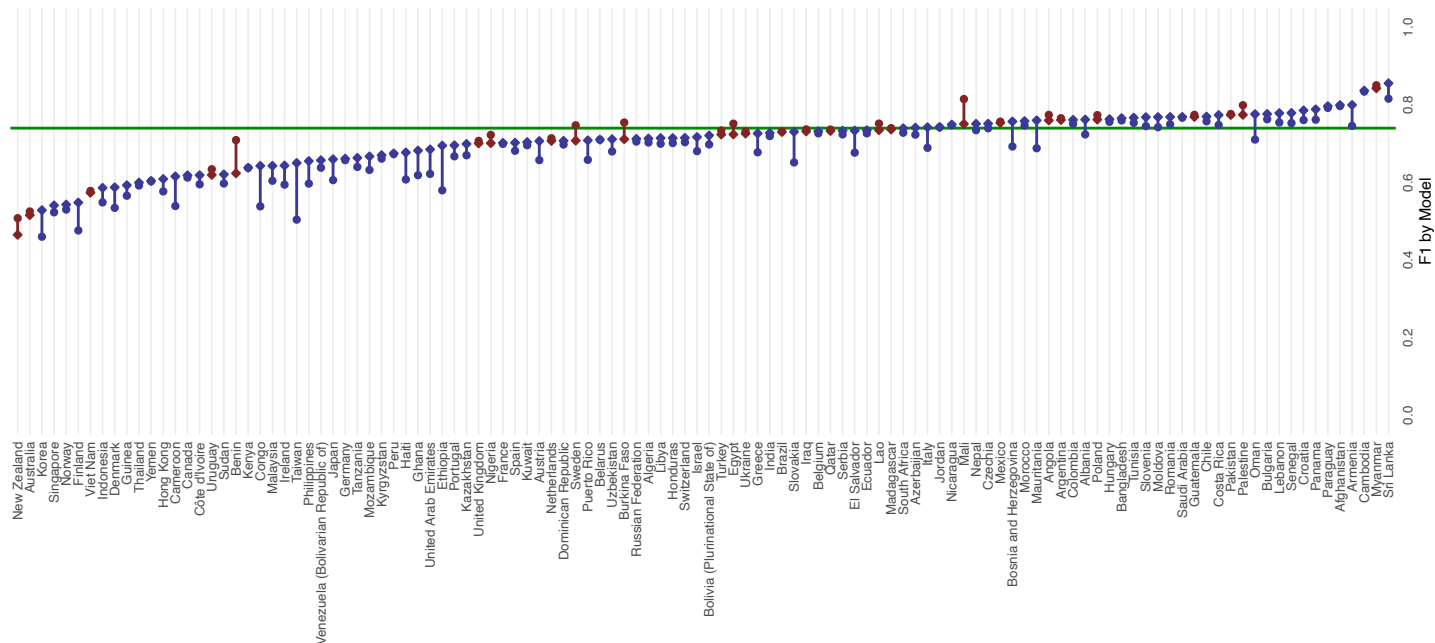
Supplementary Figure 3. Machine learning models to predict individual positive test result.

(a) Schematic diagram of the data flow for the global model. Diagram showing the full sample, and subsampling required to obtain a labelled data set for machine learning using symptoms and basic demographic factors. Down-sampling of the majority class to limit over-fitting was conducted after splitting the data into testing and validation sets, within each demographic group. After splitting the data, we used basic LightGBM to train and test country/territory- and globe-specific models. We hyperparameter tuned the global model with grid search. There was a 0.7% improvement in F1 with hyperparameter tuning. Subsequent models were not hyperparameter tuned. The global model performance overall F1 (F1 0.74), with performance for positive and negative tests in the cross-validation (0.71, 0.78) and holdout data (0.60, 0.89), respectively. Overall F1 for the holdout are presented subsequently. We took a similar approach for predicting receiving a test in the global data set. Not surprisingly, the prediction of recently tested performed worse (F1 0.59) than positive test result. This suggests that symptoms in the prior 24 hours and demographic features were less relevant, and there are other key predictors of receiving a test (e.g. spatiotemporal access, regulatory variability, contact history, high-density housing or neighborhood, and profession). These were not included the model, and further evaluation was beyond the scope of this study. **(b and c) Sensitivity analysis of model performance for individual country/territory models.** We compared the global model performance by applying the global model to holdout country/territory data, and country/territory models to their respective country/territory holdout data. We also conducted a sensitivity analysis restricted to the countries/territories that did not have the ten highest local model F1 scores. We aimed to test whether these countries were driving model features, perhaps due to unusual response patterns or data distributions. **(b)** There was no substantial change in the performance of the restricted global model (x-axis) on local data compared to the global model (difference on y-axis). Difference in performance is plotted versus performance, by country and territory, along with a regression line indicating limited correlation. **(c)** When comparing the performance of the global model (diamond) to the local (circle) models, for most countries and territories, the global model performed better on local data than the local model. Blue (higher) and red (lower) lines show change in performance when using the global model compared to the country/territory model. Graph is arranged by F1 of the global model applied to country/territory data. Median [IQR] F1 of 0.71 [0.64, 0.75], while the global model tested on local data F1 median [IQR] F1 of 0.72 [0.66, 0.76]. **(d-k) Comparison of model features and performance over time.** A model trained on early pandemic data from the first 2 weeks of the study period performed well (F1=0.74) when tested on holdout data from the full study period (the same holdout set used for the global model). Performance was also similar (F1=0.74) in sensitivity analyses omitting training and holdout data from survey versions 1-4 (prior to June 26, 2020), as these versions asked only symptomatic respondents about test results. We compared models trained and then tested on data limited to strata defined by duration of illness (brief 0-4 days, short 5-9 days, medium 10-14 days, long 15-21 days, after rounding duration down to the nearest integer), reasoning that access to testing throughout the course of illness likely varied. **(d)** Summary of model features and performance. Confusion matrices for **(e)** global, **(f)** drop version 1-4, **(g)** first 2 week, and **(h-k)** duration of illness strata.

(a)

(b)

Difference in F1_1_Global scores before and after excluding Top10 countries w.r.t F1_1_country from down-sampled global LightGBM model
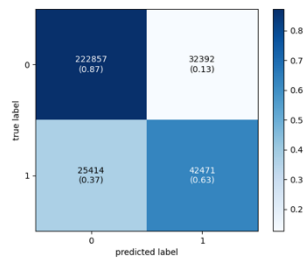
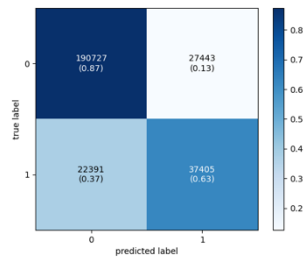(c)   Comparison of local versus global model performance.

(d)   Comparison of global versus illness duration strata, early pandemic and dropped versions 1-4 models

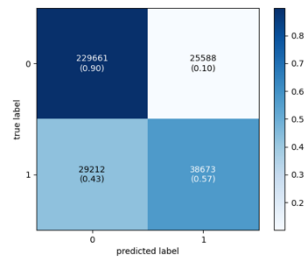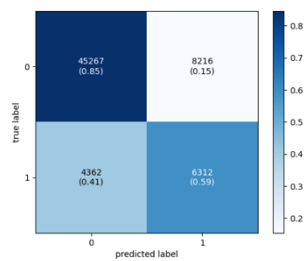| Feature	Global	Duration of Illness Strata				Study Period Strata	
		Brief 0-4 days	Short 5-9 days	Medium 10-14 days	Long 15-21 days	2 week (Early pandemic)	Versions 1-4 drop
Loss smell/taste	(1)	(1)	(1)	(1)	(1)	(1)	(1)
Cough	(2)	(3)	(7)	(6)	(6)	(4)	(2)
Fever	(3)	(2)	(2)	(4)	(2)	(2)	(3)
Chest pain	(4)	(4)	(3)	(2)	(3)	(5)	(4)
Fatigue	(5)	(10)	(6)	(3)	(5)	(11)	(5)
Ache/pain	(6)	(5)	(5)	(9)	(12)	(8)	(6)
Eye pain	(7)	(6)	(4)	(11)	(8)	(13)	(8)
Young male	(8)	(8)	(15)	(13)	(16)	(3)	(9)
Headache	(9)	(7)	(8)	(7)	(7)	(9)	(7)
Diff. breathing	(10)	(13)	(12)	(8)	(10)	(6)	(11)
Old female	(11)	(15)	(14)	(12)	(11)	(15)	(10)
Young female	(12)	(9)	(10)	(10)	(13)	(7)	(12)
Old male	(13)	(16)	(16)	(16)	(15)	(16)	(13)
Runny nose	(14)	(12)	(9)	(5)	(4)	(12)	(14)
Nausea	(15)	(14)	(13)	(15)	(14)	(14)	(15)
Sore throat	(16)	(11)	(11)	(14)	(9)	(10)	(16)
F1 score	**0.74**	**0.69**	**0.78**	**0.76**	**0.73**	**0.74**	**0.74**
Training Set (N)	**407,308**	**64,046**	**102,054**	**64,100**	**40,884**	**10,766**	**358,774**
Holdout Set (N)	**323,134**	**64,157**	**37,930**	**21,629**	**14,070**	**323,134**	**277,966**

5

(e) Global
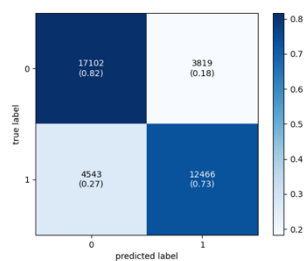

(f) Dropped survey version 1-4
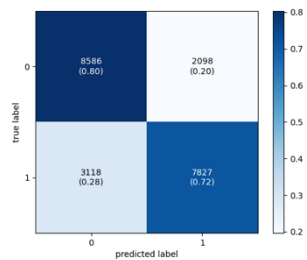

(g) First two-weeks
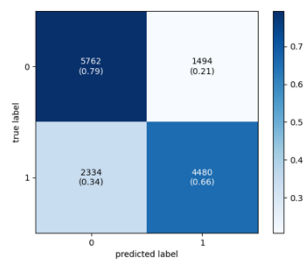

(h) Brief duration (0-4 days)

(i) Short duration (5-9 days)

(j) Medium duration (10-14 days)

(k) Long (15-21 days)

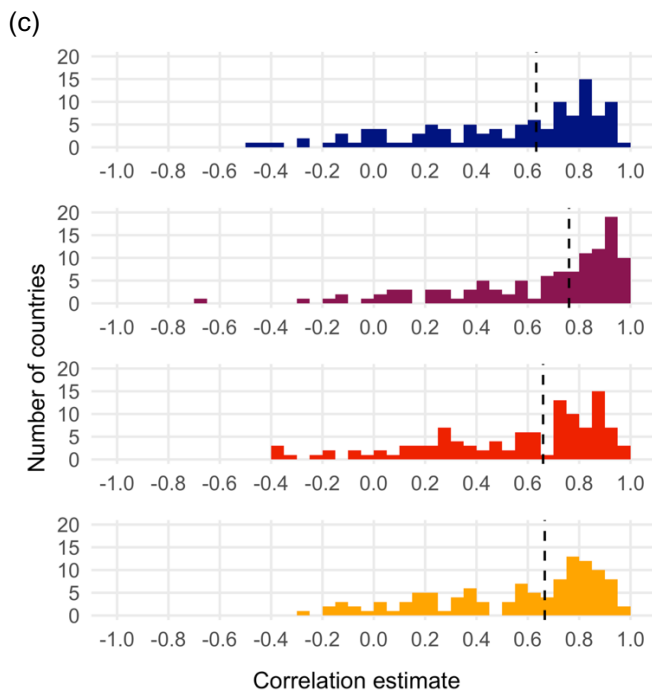**Supplementary Figure 4. Distribution of Spearman Correlations between Benchmark and UMD-CTIS Signals**
(a) Distribution summary of Spearman correlations (ρ) and difference in peak day (φ) measures comparing median within-country/territory signal rank, median value of measure by signal, and number of country/territories meeting surveillance thresholds by signal. (b) Distribution summary of correlations (ρ) and absolute difference in peak day (|φ|) within subgroups of the other similarity measure. Median |φ| <= 14 for the groups and signals with moderate to high ρ. Median ρ > 0.7 for groups and signals with |φ| <= 14. (c) Histograms of Spearman correlations for each of four primary signals, with color annotations matching those of Figure 4 in the main text. The median correlations are shown as dotted lines and are reported in the table below. N=112 countries were included in this analysis.

(a)

Variable	Median rank		Median value		
	Correlation (ρ)	Difference Peak Day (φ)	Correlation (ρ)	Difference Peak Day (φ)	Number countries/territories with 0.7 <= ρ & -14 <= φ <= 7
CCLI	1	2	0.76	1	37
CLI (broad)	3	3	0.66	4	27
CLI (narrow)	3	2	0.67	-4.5	25
Pred. positive test	3	3.5	0.66	7.5	26
Positive test	4	4	0.63	12	20

(b)

	Median |φ| by ρ Group and Signal				Median ρ by |φ| Group and Signal	
	-1 <= ρ < 0.4	0.4 <= ρ < 0.7	0.7 <= ρ < 0.9	0.9 <= ρ <= 1	|φ| <= 14	14 < |φ|
Positive test	92.5	17.5	13.0	6.0	0.78	0.47
Predicted positive test	104.0	19.0	10.0	6.5	0.77	0.42
CCLI	80.0	74.0	8.0	8.0	0.87	0.58
CLI (broad)	102.0	44.0	9.0	9.5	0.75	0.50
CLI (narrow)	72.0	15.5	14.0	7.0	0.76	0.42

	N by ρ Group and Sigal				N by |φ| Group and Signal	
	-1 <= ρ < 0.4	0.4 <= ρ < 0.7	0.7 <= ρ < 0.9	0.9 <= ρ <= 1	|φ| <= 14	14 < |φ|
Positive test	38	24	39	11	49	63
Predicted positive test	37	22	45	8	53	59
CCLI	24	22	37	29	52	60
CLI (broad)	36	21	45	10	50	62
CLI (narrow)	37	22	43	10	48	64

(c)

Supplementary Table 1. Survey questions, response options, and notes regarding survey changes and coding.

Covariate	Survey question/description	Response	Comments, versions and changes
Symptoms	B1: In the last 24 hours, have you had any of the following? Fever, cough, difficulty breathing, fatigue, stuffy or runny nose, aches or muscle pain, sore throat, chest pain, nausea, loss of smell or taste, eye pain, headache.	Respondents may select none or up to 14 symptoms 1 = Yes 2 = No	Chills (version 3) and changes in sleep (version 6) were not included in our analyses since they were added later and this presented with many missing values.
Duration of illness	B2: For how many days have you had at least one of these symptoms?	Open response: number validation	Only asked if B1 selected symptoms count greater than or equal to 1
Community COVID-19-Like Illness (CCLI)	B3: Do you personally know anyone in your local community who is sick with a fever and at least one other symptom?	1 = Yes 2 = No	In version 6 this was slightly changed to "Do you personally know anyone in your local community who is sick with a fever and either a cough or difficulty breathing?"
COVID-19 testing ever received	B6: Have you ever been tested for coronavirus (COVID-19)?	1 = Yes 2 = No	
COVID-19 testing recently received	B7: Have you been tested for coronavirus (COVID-19) in the last ${e://Field/sympdays} days? Note: sympdays is B2, but capped at 14 days.	1 = Yes 2 = No	Only asked if B6 is 'Yes' and B2 greater than or equal to 1 until version 4 (06/26/2020): then only asked if B6 is 'Yes' This was changed (version 7) to simply "Have you been tested for coronavirus (COVID-19) in the last 14 days?"
COVID-19 positive test	B8: Did your most recent test find that you had coronavirus (COVID-19)?	1 = Yes 2 = No 3 = I don't know	Only asked if B7 is 'YES'

Social Distancing	C1_m: In the last 24 hours, have you had direct contact with anyone who is not staying with you? Direct contact means spending longer than one minute within two meters of someone or touching, including shaking hands, hugging, or kissing.	1 = Yes 2 = No	Surveys answering NO coded as social distancing (per survey)
Either Social Distancing or Masking	C5: In the last 7 days, how often did you wear a mask when in public?	1=All of the time 2=Most of the time 3=Some of the time 4=A little of the time 5=None of the time 6=I have not been in public during the last 7 days	Surveys answering 1, 2 or 6, OR those social distancing (see above) coded as either social distancing or masking (per survey)
Gender	E3: What is your gender?	1 = Male 2 = Female 3 = Prefer to self-describe: [text entry] 4 = Prefer not to answer	This was changed in version 6 to 1=Male 2=Female 3=Other 4=Prefer not to answer
Age	E4: What is your age?	1 = 18-24 years 2 = 25-34 years 3 = 35-44 years 4 = 45-54 years 5 = 55-64 years 6 = 65-74 years 7 = 75 years or older	
Country	Country name as used in country level aggregates	Character	Ex: United Kingdom
Date recorded	Date that the response was recorded	Date/Time	Ex: 2020-05-01, 00:04:26
Survey duration (time)	Duration of response in seconds	Integer	Ex: 72

Weight	Survey weight to adjust to FB user population	Number (float)	Ex: 1526.36144 115 countries are weighted. Weighting documentation: https://arxiv.org/abs/2009.14675

Missing data was coded throughout the survey as follows: -99 = missing/valid skipped/invalid,
-77 = seen but unanswered

Supplementary Table 2. Parameters for global model.

Specific *lgbm.LGBMClassifier* function parameters for the global model.

Parameter	Specified Value
boosting_type	gbdt
objective	binary
n_jobs	16
silent	True
max_depth	-1
max_bin	255
subsample_for_bin	200
subsample	1
subsample_freq	1
min_child_weight	1
scale_pos_weight	1
early_stopping_rounds	30
colsample_bytree	0.6
learning_rate	0.07
min_child_samples	50
min_split_gain	0.75
n_estimators	500
num_leaves	32
random_state	501
reg_alpha	1.2
reg_lambda	1.2