



## OPEN ACCESS

## EDITED BY

Jie Huang,  
Southern University of Science and  
Technology, China

## REVIEWED BY

Bin Zhu,  
Southern University of Science and  
Technology, China  
Leshan Xiu,  
Shanghai Jiao Tong University, China  
Xinxin Han,  
Southern University of Science and  
Technology, China

## \*CORRESPONDENCE

Marc Lipsitch  
✉ mlipsitc@hsph.harvard.edu

RECEIVED 27 March 2024

ACCEPTED 18 June 2024

PUBLISHED 15 July 2024

## CITATION

Lipsitch M, Bassett MT, Brownstein JS,  
Elliott P, Eyre D, Grabowski MK, Hay JA,  
Johansson MA, Kissler SM, Larremore DB,  
Layden JE, Lessler J, Lynfield R,  
MacCannell D, Madoff LC, Metcalf CJE,  
Meyers LA, Ofori SK, Quinn C, Bento AI,  
Reich NG, Riley S, Rosenfeld R, Samore MH,  
Sampath R, Slayton RB, Swerdlow DL,  
Truelove S, Varma JK and Grad YH (2024)  
Infectious disease surveillance needs for the  
United States: lessons from Covid-19.  
*Front. Public Health* 12:1408193.  
doi: 10.3389/fpubh.2024.1408193

## COPYRIGHT

© 2024 Lipsitch, Bassett, Brownstein, Elliott,  
Eyre, Grabowski, Hay, Johansson, Kissler,  
Larremore, Layden, Lessler, Lynfield,  
MacCannell, Madoff, Metcalf, Meyers, Ofori,  
Quinn, Bento, Reich, Riley, Rosenfeld,  
Samore, Sampath, Slayton, Swerdlow,  
Truelove, Varma and Grad. This is an  
open-access article distributed under the  
terms of the [Creative Commons Attribution  
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or reproduction  
is permitted which does not comply with  
these terms.

# Infectious disease surveillance needs for the United States: lessons from Covid-19

Marc Lipsitch<sup>1,2\*</sup>, Mary T. Bassett<sup>3</sup>, John S. Brownstein<sup>4</sup>,  
Paul Elliott<sup>5</sup>, David Eyre<sup>6</sup>, M. Kate Grabowski<sup>7</sup>, James A. Hay<sup>6</sup>,  
Michael A. Johansson<sup>8</sup>, Stephen M. Kissler<sup>9</sup>,  
Daniel B. Larremore<sup>9,10</sup>, Jennifer E. Layden<sup>11</sup>, Justin Lessler<sup>12</sup>,  
Ruth Lynfield<sup>13</sup>, Duncan MacCannell<sup>14</sup>, Lawrence C. Madoff<sup>15</sup>,  
C. Jessica E. Metcalf<sup>16</sup>, Lauren A. Meyers<sup>17</sup>, Sylvia K. Ofori<sup>2</sup>,  
Celia Quinn<sup>18</sup>, Ana I. Bento<sup>19</sup>, Nicholas G. Reich<sup>20</sup>, Steven Riley<sup>21</sup>,  
Roni Rosenfeld<sup>22</sup>, Matthew H. Samore<sup>23</sup>, Rangarajan Sampath<sup>24</sup>,  
Rachel B. Slayton<sup>25</sup>, David L. Swerdlow<sup>2</sup>, Shaun Truelove<sup>12</sup>,  
Jay K. Varma<sup>26</sup> and Yonatan H. Grad<sup>27</sup>

<sup>1</sup>Center for Forecasting and Outbreak Analytics, US Centers for Disease Control and Prevention, Atlanta, GA, United States, <sup>2</sup>Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA, United States, <sup>3</sup>François-Xavier Bagnoud Center for Health and Human Rights, Department of Social and Behavioral Sciences, Harvard T. H. Chan School of Public Health, Boston, MA, United States, <sup>4</sup>Boston Children's Hospital, Harvard Medical School, Boston, MA, United States, <sup>5</sup>Department of Epidemiology and Public Health Medicine, Imperial College London, London, United Kingdom, <sup>6</sup>Big Data Institute, Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom, <sup>7</sup>Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States, <sup>8</sup>Division of Vector-Borne Diseases, US Centers for Disease Control and Prevention, Atlanta, GA, United States, <sup>9</sup>Department of Computer Science, University of Colorado Boulder, Boulder, CO, United States, <sup>10</sup>BioFrontiers Institute, University of Colorado Boulder, Boulder, CO, United States, <sup>11</sup>Office of Public Health Data, Surveillance, and Technology, US Centers for Disease Control and Prevention, Atlanta, GA, United States, <sup>12</sup>Department of Epidemiology, UNC Gillings School of Public Health, Chapel Hill, NC, United States, <sup>13</sup>Minnesota Department of Health, Minneapolis, MN, United States, <sup>14</sup>US Centers for Disease Control and Prevention, Office of Advanced Molecular Detection, Atlanta, GA, United States, <sup>15</sup>Massachusetts Department of Public Health, Boston, MA, United States, <sup>16</sup>Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, United States, <sup>17</sup>Department of Integrative Biology, University of Texas at Austin, Austin, TX, United States, <sup>18</sup>Division of Disease Control, New York City Department of Health and Mental Hygiene, New York City, NY, United States, <sup>19</sup>Department of Public and Ecosystem Health, College of Veterinary Medicine, Cornell University, Ithaca, NY, United States, <sup>20</sup>Departments of Biostatistics and Epidemiology, University of Massachusetts Amherst, Amherst, MA, United States, <sup>21</sup>United Kingdom Health Security Agency, London, United Kingdom, <sup>22</sup>Departments of Computer Science and Computational Biology, Carnegie Mellon University, Pittsburgh, PA, United States, <sup>23</sup>Division of Epidemiology, Department of Medicine, University of Utah, Salt Lake City, UT, United States, <sup>24</sup>Siemens Healthcare Diagnostics, Inc., San Diego, CA, United States, <sup>25</sup>Division of Healthcare Quality Promotion, US Centers for Disease Control and Prevention, Atlanta, GA, United States, <sup>26</sup>SIGA Technologies, New York City, NY, United States, <sup>27</sup>Department of Immunology and Infectious Diseases, Harvard T. H. Chan School of Public Health, Boston, MA, United States

The COVID-19 pandemic has highlighted the need to upgrade systems for infectious disease surveillance and forecasting and modeling of the spread of infection, both of which inform evidence-based public health guidance and policies. Here, we discuss requirements for an effective surveillance system to support decision making during a pandemic, drawing on the lessons of COVID-19 in the U.S., while looking to jurisdictions in the U.S. and beyond to learn lessons about the value of specific data types. In this report, we define the range of decisions for which surveillance data are required, the data elements needed to inform these decisions and to calibrate inputs and outputs of transmission-dynamic models, and the types of data needed to inform decisions

by state, territorial, local, and tribal health authorities. We define actions needed to ensure that such data will be available and consider the contribution of such efforts to improving health equity.

#### KEYWORDS

pandemic, COVID-19, surveillance and forecast system, public health, infectious diseases, mathematical model

## Introduction and purpose

To monitor pandemic pathogens effectively, modern surveillance systems should make use of the growing wealth of routine data from the health sector and from a range of other sources with new applications to disease surveillance such as mobility, internet searches, and wastewater. Public health experts must integrate these data in new ways that increase their value. We need purpose-built systems to detect new and evolving threats and to provide information as quickly as possible about those threats. What are the characteristics of the new pathogens or new variants of existing pathogens? What is their incidence and prevalence? What is the vulnerability of the population to infection and disease? What is the impact of our efforts to respond to these threats?

Systems to generate, integrate, and interpret these data should be designed and built with the explicit purpose of providing timely evidence to inform decisions about disease control and mitigation. First, they will provide direct input into decision making. For example, evidence of low vaccine effectiveness may prompt efforts to boost or change formulations or doses. As another example, real-time lab order data for diagnostic tests may prompt adjustments to resource allocation. Second, these data will parameterize scenario and forecasting models (1–3). For instance, estimates of per-case severity of a new variant, incorporated into forecasts or other models of case burden, may influence planning for hospital capacity and supply stockpiling and distribution.

This document reflects the framing ideas and the discussions held at a symposium organized by Harvard T.H. Chan School of Public Health entitled “Quantitative Tools and Data Opportunities for Pandemic Surveillance and Response,” held June 29–30, 2022, involving a range of public health and public officials, surveillance experts and other epidemiologists, and epidemic modelers. We first aim to identify the most important decisions for disease control and mitigation and the evidence that is needed to inform them. We then describe a set of surveillance activities designed to provide timely, reliable, and appropriately scaled data to inform these decisions. Our focus in this report is limited to domestic detection, characterization, and estimation of the burden of a pandemic pathogen in terms of direct health effects. Although we note the importance of monitoring economic, social, and indirect public health impacts of a disease control measures, we do not offer a comprehensive treatment of this element of pandemic monitoring and response activities. This paper builds on earlier efforts (4) while incorporating both the new possibilities that technology now provides, as well as the lessons of COVID-19.

We differentiate between two related but distinct goals of surveillance, as this document will focus on only one. The first goal is

to provide early warning about a potential pandemic, and so this type of surveillance includes global monitoring and rapid identification of domestic introductions. The second goal is to provide support for decision making during an ongoing pandemic, including tracking incidence, prevalence, and the pathogen’s properties. While we will briefly remark on the former—surveillance for early warning—we will primarily focus on the latter—surveillance for decision making.

## Detecting a jurisdiction’s first cases of a new disease

The first set of decisions faced by a domestic public health jurisdiction, following the appearance of a pandemic threat somewhere else in the world, concerns the questions of whether, how, and to what extent to scale up a response to reduce the risk of importation or, if importation has happened, to control its spread within the jurisdiction. Measures to reduce importation via restriction or testing of inbound travelers may buy limited time to prepare (5, 6), though such measures lose relevance once local transmission is established (7).

To inform decisions about how to balance scarce public health resources between preventing importation vs. controlling local spread, it is critical to assess the risk that the infection has already arrived and started spreading within the jurisdiction. Testing and sequencing of specimens from international travelers at airports and analysis of wastewater from international flights may provide evidence of pathogen importation (8). Early evidence of local spread may come from informal communications among health care providers, reporting systems such as ProMED-mail (9), and “pre-health care” data (e.g., absenteeism, internet search queries). Signals may arise from monitoring of syndromes compatible with infections, or the volume, distribution, and results from clinical laboratory tests. Increasingly, wastewater monitoring can be the site of early detection of new threats or variants (10) and from routine programs or enhanced efforts at sequencing of clinical samples (11). They may also come from anomalous findings in sentinel and research efforts [e.g., the Seattle Flu Study at the start of COVID-19 (12)].

With each type of monitoring, there is first the question of what defines the signal we are looking for and then there is a tradeoff between having a highly sensitive and timely system capable of sounding an alarm early on one hand and producing too many false alarms on the other. In most situations, a high positive predictive value for such systems will be essential because the cost of responding to frequent false alarms is high. Much more work is needed to assess how to use and combine complementary monitoring signals to identify points at which an alarm should be escalated into a response.

## Surveillance for decision making during a pandemic

A comprehensive list of the decisions and guidance required in a pandemic would fill a much longer document than this. Still, based on the combined experience of the emergency response phase of COVID-19 (2020–2023) and H1N1 influenza (2009–10) pandemics, we propose a set of consistent themes that capture many of the major types of decisions arising both in the early days of each pandemic and throughout.

A brief, necessarily incomplete, list of such decisions follows, adapted from the list in (4):

- 1 Public health goals of a response (elimination, control, protection of high-risk groups, protection of health care functioning, or a combination) and overall scale of response needed to meet these goals.
- 2 Timing of scale-up and scale-down of response.
- 3 Choice of nonpharmaceutical countermeasures (individual-targeted such as quarantine, isolation, and personal protection; population-targeted such as closures). This includes decisions about the timing, magnitude, and geographic range of protective measures that may be socioeconomically costly. A related set of decisions concerns how such measures should be prioritized, i.e., who should receive protective equipment when it is scarce, and how closures should be targeted to reduce economic and social disruption.
- 4 Choice of medical countermeasures, including diagnostics, therapeutics, and vaccines. This includes decisions about development, stockpiling, procurement, expanding capacity (e.g., building alternative care sites), and more. Here too, questions of allocation and prioritization are central. This also includes planning for potential surges.
- 5 Specific policies for each of the issues above in special populations including vulnerable communities, and settings such as health care, schools, congregate settings, transport, etc. (See *Centering Equity* below)
- 6 Balance between community countermeasures to reduce severe disease or reduce transmission (e.g., allocation of resources to those at high risk of complications or high risk of transmission).
- 7 Design and implementation of staged alert systems to provide real-time risk awareness and trigger policy changes (13, 14).
- 8 Imposition and removal of international travel screening and restrictions.
- 9 Choice of public health communication strategies.

Each of these decisions requires specific data to decide how to improve health equitably, effectively, and efficiently while minimizing social and economic disruption. For example, decisions on testing, isolation, and quarantine policies require evidence on the natural history of infectiousness (or at least a proxy such as viral load), test sensitivity at different levels of viral shedding, the relationship between symptoms and infectiousness, and the potential economic and social consequences for various communities of the policies under consideration. In contrast, decisions about the timing of vaccine boosters require evidence on the effectiveness of existing vaccines against infection, transmission, and severe disease endpoints, stratified by such factors as pathogen variant, time since vaccination, and age,

as well as understanding of how vaccine protection is distributed across demographic groups.

## Decisions faced by state, territorial, local, and tribal authorities

In the federal system in the United States, public health is decentralized and typically not coordinated among states. State, territorial, local, and tribal (STLT) governments are responsible for nearly all binding policy decisions in public health, with governance health structures varying by state (15). The purview of these bodies includes (16) prescribing and enforcing isolation, quarantine, mask mandates, and restrictions on businesses and gatherings; vaccine prioritization and distribution; and (to a degree) diagnostic testing. They also hold responsibility for closely related areas, such as public education. STLT governments all have a desire for similar types of data, but vary in how much they need, how quickly they need it, and how they use it.

Many decisions involve procurement and distribution of countermeasures. Because STLT authorities are making allocation decisions within their jurisdictions (e.g., for counties, cities, hospitals, schools), jurisdiction-wide measures of disease activity are rarely sufficient; instead, more geographically granular numbers are required (Table 1).

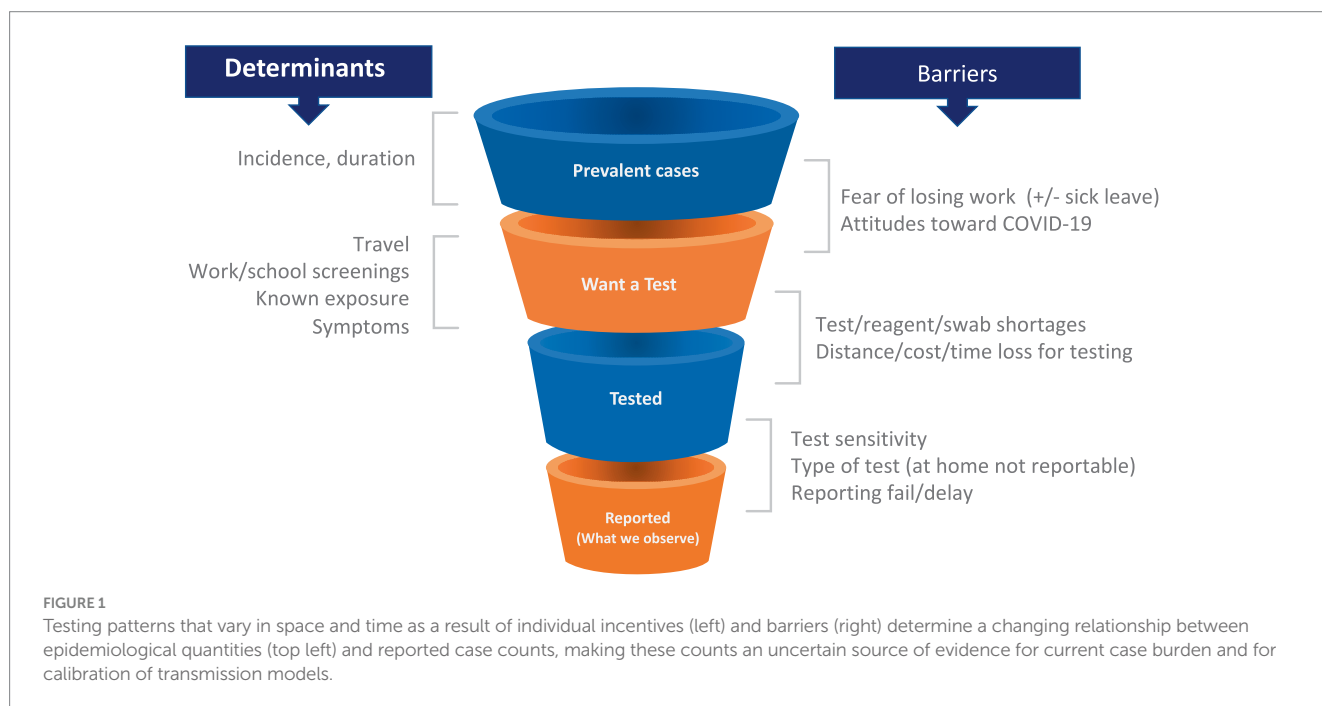
## Data needs for decision support: the COVID-19 experience

A range of data sources could and, during the COVID-19 pandemic, did provide evidence to support decisions by health authorities. Following initial social media reports of clusters of pneumonia, some of the earliest specific data to characterize the COVID-19 threat came from traditional sources, such as from case reports posted on Chinese public health department websites (17). A key challenge was the repeated change in the syndromic case definition in the early days (18). But other early data came from unexpected sources, such as cruise ships (19), restaurants (20), and fishing vessels (21), where conditions allowed inference of the path of transmission and thereby provided evidence about the degree and mechanisms of spread. Specifically, these provided some of the earliest evidence of asymptomatic and aerosol spread, which, when properly interpreted, aided in the design and prioritization of testing and other control measures. As had been true in the 2009 influenza pandemic (22, 23), sampling of travelers provided early estimates of the extent of global spread, growth rates, and likely under-detection (24, 25).

As the pandemic spread, the strengths and limitations of each data source became evident. Multiple data types were required to provide even an incomplete picture of trends in incidence and prevalence and behavior (26, 27). For example, case counts were used as an important indicator of disease burden. However, the relationship between new cases and true incidence varies as a function of numerous factors, including test availability, test reporting requirements by jurisdiction (which did not always include reports of negative tests), rates of testing through clinical facilities (which declined with the growth of rapid antigen testing), and incentives to get or avoid testing (Figure 1). Some of these limitations can be mitigated by breaking out case counts by

TABLE 1 summarizes key decisions and associated needs for jurisdiction-level data and analytics in COVID-19 cited by state and local leaders during the symposium.

Decision	Data/analytics need
Size of response needed	Rapid threat characterization
Choice of community countermeasures	County-level disease burden and transmission measures
How to ensure adequate supply of hospital beds, ventilators, personal protective equipment	Forecasts of demand for these items
School and congregate setting policies (closure of schools, infection control measures in jails, prisons, nursing homes, etc.)	Understanding of rates of transmission into, within, and from each of these settings and impact of testing and infection control on these rates as well as population specific health-risks
Countermeasure deployment within a jurisdiction	Age, racial, ethnic, and geographic patterns of transmission and disease burden. Note: these are often crude proxies for social determinants of infection and outcome risk, not adequate for scientific understanding of why particular groups are at risk, but nonetheless potentially useful for focusing prevention and treatment efforts on those with high vulnerability.
Efforts to distribute and promote vaccination	Variant prevalence, vaccine coverage, and vaccine effectiveness against dominant and emerging variants



the reason why an individual was tested (symptoms, travel, surveillance), but this was not consistently done in the US. As a result of these limitations, hospitalizations and even deaths were increasingly used as the more reliable indicators of case numbers, sacrificing some timeliness for a more consistent relationship to the underlying incidence of infection. Random sampling approaches (described below) can overcome these limitations and provide more consistent and reliable estimates of incidence and prevalence and how these change over time. Only the United Kingdom and Luxembourg used random sampling on a large scale, perhaps because of the cost and logistical challenges. Notwithstanding their limitations, case counts were the major early data source in the United States and provided critical evidence especially when linked to demographic information. Syndromic surveillance—done routinely as part of monitoring

influenza trends—from emergency room visits and hospital admissions were also valuable data sources, particularly when testing was limited. However, interpretation of syndromic surveillance was complicated by changes in healthcare seeking behavior and the increased use of telemedicine.

Novel data streams provided confirmatory evidence as well as early warnings of trends that might not be evident in case counts. For example, wastewater surveillance for SARS-CoV-2 was adopted in numerous jurisdictions from 2020 to 2022 and provided evidence on local epidemic trends, although the precise relationship between wastewater abundance and the number of infected persons depends on the wastewater sampling scheme and on shedding patterns (among other issues), and thus difficult to quantify (28). Moreover, by its nature, wastewater cannot indicate who has been infected, thus

leaving the demographic profile of infected persons uncertain. Finally, wastewater surveillance as currently applied will miss infections in areas with high reliance on septic systems, which serve roughly one fifth of the households in the U.S. with heavy concentrations in certain geographies, inducing inequities in whose infections are tracked by this approach (29).

Another novel data stream is the use of population-wide distributions of viral load measured from PCR testing, which, in the aggregate, can provide information on the trajectory of the epidemic, even from a single cross-sectional analysis (30). This approach has reached proof of concept and has the advantage that it may be less sensitive to trends in testing behaviors than measures of incidence based on case counts, and unlike wastewater surveillance, it can provide some information on the demography and precise location of cases. However, further work is needed to see how a transition to non-PCR testing for many new cases, the halting of pre-procedural and asymptomatic testing, and the shifts in viral kinetics that come with immunity (from vaccination and infection) affects the nature of this signal (31–33). Moreover, the identifiability of time-since-infection from viral load, which is needed for the approach to work, depends on the asymmetry in viral load over time [fast rise, slower decline (30)], which may or may not be a feature of future infectious diseases.

Digital data can also be used for surveillance and to inform on epidemic trajectory. ProMED-mail (9) and HealthMap (34) are valuable for flagging and disseminating reporting and information on events known or suspected to be infections and outbreaks. Data from search engines, social media, and news reports data can also inform epidemic dynamics and for forecasting (35–37).

Finally, testing for antibodies in sera collected either for the purpose of serologic surveillance or in convenience samples (e.g., blood banks, discarded clinical samples) was used to characterize both the landscape of population immunity (i.e., who was and wasn't vulnerable to reinfection) and to distinguish between those who had acquired immunity via vaccination vs. infection (38). Secondary analyses of COVID-19 vaccine studies identified complexities in answering the latter question, finding that infection does not reliably induce antibodies to non-vaccine antigens in vaccinated individuals (39).

An important conclusion is that no one data source or surveillance tactic is sufficient. In a setting like the U.S., multiple surveillance approaches are needed at scale. Beyond the obvious need to combine data sources, several points stand out.

The first is the value of data completeness and of linking data types to produce evidence that is greater than the sum of the parts. For instance, while counts of cases and hospitalizations are valuable, missing race/ethnicity, geographic, and other patient characteristics have impeded efforts to improve services to groups that are underserved or experience high disease burden and to improve equity in health-related outcomes. Similarly, meticulously linking sequence data from patient isolates with demographic and clinical predictors of severe outcomes, including vaccination history, and clinical outcomes can help to evaluate the threat posed by novel variants (40). Unfortunately, despite prodigious amounts of SARS-CoV-2 sequencing in the U.S., this form of linkage has been relatively rare to date.

Second is the value of clear and accessible data dashboards with transparent data sources to make the state of the epidemic locally

evident to the public. The same data should also be available to analysts in public health departments, academia, and other sectors via application programming interfaces (APIs) to facilitate rapid data analysis. This can facilitate shared decision making and help to increase public support for control measures. For example, the city of Austin, Texas developed a COVID-19 staged alert system that guided local policy between May 2020 and March 2022 (13, 14). The public-facing dashboard featured a single graph that tracked COVID-19 hospital admissions and clearly indicated thresholds between the red, orange, yellow, green, and blue risk levels that were linked to specific actions. The county judge, city mayor, and public health authority cited the dashboard almost daily to communicate risks, explain changes in policy, and cultivate adherence via news outlets and social media. This system was only possible because local authorities required area hospitals to report daily admissions beginning in April 2020, long before such data were generally available.

Finally, discussion at the Symposium emphasized the value of metrics that could be compared across jurisdictions. Decision-makers expressed a desire for objective criteria by which their performance can be judged. Comparisons across states, for example, were hampered by differential testing rates that affected case counts in ways not reflecting actual prevalence. A CDC-supported academic effort called *covidestim* (41) used Bayesian evidence synthesis to harmonize estimates of current and cumulative infections across states and counties, providing an example of what could be done by health authorities. However, this effort was also hampered by unanticipated changes in reporting tempo, as well as 'data dumps' and data backfilling. Different definitions of COVID-19 hospitalization across states and over time impeded comparisons of outcomes that would have provided indications to elected leaders of the quality of their responses and informed improved responses.

## Surveillance inputs to forecasts, scenario projections, and analytic models

As noted above, many aspects of pandemic decision-making can directly incorporate evidence from surveillance, and will also make use of

- Nowcasts: estimates of current burden of cases, hospitalizations, deaths, and other quantities that account for delays in reporting (42);
- Forecasts: relatively short-term projections using time-series and other statistical modeling techniques, sometimes supplemented with transmission-dynamic approaches to estimate future case, morbidity, and mortality burden, typically on the scale of days to weeks (1, 3);
- Scenario models: longer-term estimates of pandemic dynamics using transmission-dynamic modeling approaches to anticipate multiple possible futures under stated assumptions about behavior, viral evolution, vaccine durability, etc., typically on the scale of months to years (43);
- Results from analytic models: estimates about different characteristics of the pathogen or a population of concern that are specifically designed to inform a decision or guidance, such

as school- or nursing-home-based testing policies (44), border restrictions or contact tracing procedures (45), or quarantine approaches (46).

These categories of decision-support tools require estimates of *input quantities* that represent the assumptions of the models—for example, in scenario models, estimates of per-case severity or vaccine effectiveness. These estimates need to be timely, representative, and specific to the pathogen or variants circulating or anticipated to circulate (for example, due to importation). In addition, scenario and forecast model *output* must be calibrated to existing measurements of disease burden: incidence of infection, diagnosed cases, hospitalizations, deaths, and other relevant metrics, as well as against cumulative measures such as seroprevalence. The categories of input and output are somewhat fluid, as a model with sufficient data to calibrate outputs may be able to estimate the values of some of the quantities described here as inputs. In a fully Bayesian framework, both external estimates (as priors) and calibration to output data may contribute to posterior parameter estimates.

For forecasts, evaluation can be performed quickly due to the short-term horizon of the predictions made, with results that can provide feedback to modelers about places where models are mis-specified. Evaluating scenario projections is more complicated, as multiple sets of counterfactual projections are made under different assumptions about how a pandemic situation will evolve over the course of months or years (47). Most (or perhaps all) of the scenarios will not be realized exactly as assumed, making evaluation less straightforward.

Together the quality and timeliness of these input parameters and output calibrations are important determinants of how useful a model is for decision making. While there are techniques to adjust for incomplete or lagged information, the absence of certain ingredients—especially model output calibration targets such as numbers of cases or hospitalizations—can critically compromise the ability to generate models that reflect reality to the point of hampering basic situational awareness. Data systems that support modeling and in turn decision-making during pandemics should be considered vital national security capabilities and prioritized accordingly.

A list of the key needs for model inputs is as follows, many of which may change as a pathogen evolves (referred to below by their letters):

- a Pathogen kinetics/epidemiological parameters (e.g., incubation period, latent period, infectious period, infection fatality ratio). Estimation of these inputs may itself require simple models, particularly at the early stages of a pandemic [ref: Gostic paper].
  - b Transmissibility and efficiency of various transmission mechanisms
  - c Risk factors for infection and severity
  - d Individual and population immunity (including effects of infection, vaccination, and waning)
  - e Diagnostic test characteristics, including specificity and sensitivity for active (acute) and past infection
  - f Vaccine effectiveness and waning of effectiveness, for infection, severe disease, and mortality endpoints.
  - g Treatment effectiveness
  - h Policies, uptake, and effectiveness of nonpharmaceutical interventions
  - i Population mobility and interactions: contact networks and patterns by setting
  - j Importation risk
  - k Other co-circulating pathogens of concern (e.g., if concurrent with significant influenza transmission)
  - l Capacity and utilization of healthcare resources (including hospital beds, therapeutics, and vaccines)
- Key additional data requirements for fitting models— as well as for general situational awareness – include:
- m Geographically and demographically stratified incidence, duration and prevalence of infection, hospitalization, ICU admission, death, and other relevant metrics associated with the pathogen, ideally by variant.
  - n Strain-specific incidence.

## Meeting these needs

This draft framework is a preliminary attempt to scope a system that could meet the needs listed above for situational awareness, decision support, and inputs and outputs for modeling and analytics for a new variant or a new pandemic. Capacity to achieve these would also be applicable to other pathogens, especially, but not only, respiratory ones.

### A. Estimating model inputs

#### System 1: high-frequency sampling for pathogen kinetics and diagnostic sensitivity (quantities a,e)

Possible Mechanism: Surveillance would be established to obtain repeated samples (for COVID-19, respiratory samples) from individuals exposed to a pathogen of interest (now, SARS-CoV-2) from the time of exposure through infection to the time of clearance. High-frequency sampling will provide detailed profiles of pathogen kinetics, which could be subgrouped by prior infection history, vaccination status, pathogen variant, demographics, and other predictors. Simultaneous use of nucleic acid amplification (NAAT), culture, and antigen-based testing on these specimens would provide detailed estimates of the sensitivity of each as a function of symptoms, pathogen load, pathogen infectious capacity, variant, and time since exposure/first-positive to inform choice of diagnostics and isolation/test/quarantine policy.

Performers might be research/surveillance networks or STLT health departments (recognizing that the health departments may have limited bandwidth in the context of an outbreak). The ability to scale up is critical. While pathogen kinetics are not likely to vary from place to place, geographic diversity in sites capable of performing these investigations will increase the timeliness of results in case one region is hit much earlier than others.

Settings may include households, universities, day cares and schools; intensely monitored cohorts such as sports leagues or health

care workers, congregate settings such as homeless shelters, correctional and detention facilities, or nursing homes.

Precedents: United Kingdom Household study (48) and United States National Basketball Association studies (32, 49).

### System 2: integrating routine sequencing with detailed clinical data (quantities b,c,d,f,g,n)

**Possible Mechanism:** A payer-provider network with diverse geographic and demographic representation (alternatively, a private sector entity or consortium of public health departments and laboratories capable of merging clinical data with sequence data) would track individuals as a cohort (not necessarily defined by long-term follow-up but perhaps with exposure or a positive test as an entry criterion) with known vaccine and prior infection history through diagnosis (outpatient or inpatient) and through the cascade of care to estimate the probability and severity of infection as a function of this history (vaccine effectiveness and infection-acquired immunity) and variant. Sequencing of positive clinical specimens would enable the variant-specific estimates. This system would provide a reliable infrastructure for assessing severity, vaccine effectiveness, and treatment effectiveness linked to infection and vaccination history for each new variant/virus. It would be crucial to link electronic health record (EHR) within the network to key external sources of data such as immunization registries. Improving completeness of such registries is also a high priority to improve the quality of these inferences. Strategies for linking pathogen genome sequencing with EHRs will depend on whether these data are from clinically validated systems and, if not, will require consideration to ensure use for research and not clinical purposes.

It would be valuable to explore to what extent such studies could be done in networks such as PCORNet (50) or the Vaccine Safety Datalink (51) that assemble EHRs from multiple health systems into a common data model; questions include how rapidly this could be done and whether sequence data could be linked to these records.

In addition to payer-providers, robust testing, reporting, and data collection capabilities should be considered for congregate settings at high-risk for transmission such as skilled nursing facilities, correctional facilities, detention facilities, and homeless shelters that can follow individuals from positive tests through outcomes.

**Precedents:** Cohort studies on variant-specific relative severity (52), relative vaccine effectiveness (53) and absolute vaccine effectiveness (54) have been performed during the COVID-19 pandemic. None of these included genomic sequencing or serological profiling integrated with clinical data collection, in part due to the issues of linking with EHRs as mentioned above. Integration of sequencing in particular is essential for the likely future scenario where one cannot rely on proxies for genetic variant that have been exceptionally convenient in COVID-19, notably the failure of the S-gene PCR target in certain polymerase chain reaction-based diagnostic tests.

In the US, this work could build upon or integrate with existing platforms such as VISION and Investigating Respiratory Viruses in the Acutely Ill (IVY) (55). Key additions would be sequencing and more comprehensive estimates of severity.

### System 3: behavioral surveillance and other routine data collection (quantities h,i,j,l)

Goals of behavioral surveillance are to provide real-time estimates of mobility, work-from-home frequency, proportion of schools open

or closed, use of other nonpharmaceutical interventions such as masks, and vaccine behavior/hesitancy. Data useful during COVID-19 included vaccine coverage from HHS Protect (56) and Census Pulse (57) and other surveys on vaccine intentions, mask use, work-from-home, and school opening/closure. Private sector [e.g., mobility (58)] and publicly available data [e.g. (59)], exist that measure many quantities of interest. These include self-reported mask use, absenteeism data from school and work, internet search queries, and much more. Further work needs to be done in several areas to enhance the value of these data streams:

- identify cost-effective sources of such data;
- quantify the degrees of representativeness in measurement from these different data sources by such factors geography, race/ethnicity, and social determinants of health;
- improve our mechanistic understanding of how these measures of mobility relate to transmission behavior, which will likely differ by social factors, pathogen transmission routes, and epidemic stage, among other factors (27)

A particular example of one such data stream is air travel and other travel data to estimate importation risk.

**Precedent:** Census Pulse and other surveys exist. Many local jurisdictions have used mobility data from private providers, often via academic intermediaries<sup>1</sup> to assess local trends. Vaccine coverage data exist with some limitations. The Center for Disease Control and Prevention's (CDC) Division of Global Migration and Quarantine maintains access to timely estimates of air travel volume.

## B. Fitting model outputs

### System 4: repeated testing for infection and immunity in a random sample of the population (m)

**Mechanism:** An academic, government (e.g., CDC or a coalition of state health departments), or private sector entity would identify a longitudinal sample and/or repeated cross-sections representative of the U.S. population for monthly testing for infection and immunity as evidence of prior infection. In the COVID-19 case, this would be PCR testing of respiratory samples and antibody measurement in blood; testing approaches might differ for future pathogens. Samples would be obtained by home visit or mail/courier. Specimens testing positive for one or more respiratory viruses would be sequenced. The initial sample would be powered to detect US-level trends; scale-up in a pandemic would enable regional/state-level and demographic-specific (e.g., age, race, sex-specific) estimates of virus prevalence and seroprevalence irrespective of symptoms and at the level of variant/subtype/species/type (depending on the pathogen).

In pathogens with antibody-based immunity, blood samples would be tested for multiple antibodies including vaccine and nonvaccine antigens of the novel pathogen. These would provide a population-based denominator for severity estimates, enable calibration of scenario and forecast models, track trends in viral

<sup>1</sup> e.g., <https://www.covid19mobility.org/>

species/variants in an unbiased way, and estimate the magnitudes of health inequities to better prioritize prevention measures (60).

Addition of serologic testing of a random, representative sample of the population would supplement existing passive serosurveillance approaches such as from blood donors (38), newborn heel sticks (61), or discarded specimens (62, 63). Longitudinal sampling would enable more precise estimates of rates of waning of antibody concentrations (64, 65) and the consequences for estimation of cumulative incidence using particular assays.

An alternative approach would be to use healthcare-based testing of individuals requiring admission for conditions not directly related to the pandemic, using weighting to standardize the population seeking health care to the background population (66, 67), though the quality of such data would need continuing validation.

Other alternatives would include the use of voluntary testing results, such as those gathered by test-proctoring telehealth services, retail pharmacies, or the like. CDC/FDA requirements to ask the reason for a test would facilitate interpretation (symptomatic vs. travel vs. exposure, for example).

**Precedents:** The main proposal could be roughly modeled on the United Kingdom COVID-19 Infection Survey and REACT-1 studies. One of the alternative approaches—universal testing of individuals requiring admission for non-pandemic reasons—was used in New York City early in the COVID-19 pandemic (68) and has been used in Indiana with reported high value (66, 67) for both prevalence and seroprevalence.

## System 5: maintain hospitalization surveillance data (l,m)

Hospitals have been required to report COVID-19 and influenza hospitalizations to HHS, and these formed the backbone of multiple forecasting and scenario modeling efforts in the US. It is critical to maintain the generation, interpretation, timeliness, and accuracy of these data to inform forecasts. In addition to the forecasting products, these data underlie hospital capacity and burden situational awareness, the ability to monitor outbreaks, and community burden indicators.

**Precedents:** Exists as of September 2023 but needs to be maintained at a base level outside of emergencies and be able to ramp up quickly at a time of new emergency (69).

## C. Actions needed

### Administrative and reporting preparedness

The response to COVID-19 required collaborations across sectors—public, private, and academic—but these collaborations were often forced to work through administrative frameworks that were not designed with speed and flexibility in mind. In turn, such mis-specified frameworks ultimately slowed or limited some critical public health projects and prevented others from being undertaken entirely. To address this class of problem, we propose six ideas below that would update, recast, or create key frameworks that establish links across sectors and that facilitate the urgent work of pandemics, while maintaining safeguards and oversight.

- 1 *Emergency data use agreements and formats.* Data use agreements (DUAs) are core elements to collaborative work across institutions, but they pose two types of challenges. First,

the process for negotiating an agreement acceptable to the institutions providing and receiving the data is often slow. The staff on each side tasked with reviewing and signing off on these agreements may have many competing priorities or be overwhelmed as an outbreak or pandemic may dramatically increase the volume of DUAs. Work on a sensitive or high-profile project, such as associated with an outbreak of infectious disease or a pandemic, generates additional scrutiny and often further lengthens the review process. Second, conflicting limitations can stall progress or even undermine a project before it starts. For example, in a partnership between academics and government public health institutions, academic institutions may deem the freedom to publish without interference to be non-negotiable. Public health institutions, however, may require veto power over what, if anything, is published, due to the sensitivity of the institution's data and ownership thereof. To address these problems, one solution is to establish Emergency Use Data Authorizations (EUDAs) for public health data with a standing framework vetted and updated regularly (e.g., annually), perhaps at the individual state level. Such EUDAs would catalyze collaborations and enable investigators at both institutions to shift the balance of effort up front from administrative to research tasks. As these are put in place, discussions about data formats can take place, ideally also in advance, to ensure that when data are delivered they are as ready-to-use as possible.

- 2 *Surveillance versus research: updating the Common Rule.* Projects designated as human subjects research require institutional review board (IRB) review, whereas those designated as public health surveillance are deemed not to be research, and thus do not require IRB review. This surveillance-research dichotomy has substantial implications for timeliness and speed of work, because writing, reviewing, and adjudicating IRB reviews—while vitally important for protecting the rights, welfare, and well-being of human subjects—may take days to weeks. The boundaries between surveillance and research are governed by the Common Rule, which states that public health surveillance activities “include those associated with providing timely situational awareness and priority setting during the course of an event or crisis that threatens public health (including natural or man-made disasters)” [45 CFR 46.102(I)(2)] (70). Unfortunately, these boundaries lacked clarity and standardization as questions arose during the COVID-19 pandemic. For example, while case monitoring is clearly surveillance and a routine public health activity, one could make a strong argument that “situational awareness and priority setting” includes assessing vaccine effectiveness and disease severity for new variants. However, analysis of variants requires pathogen genome sequencing, which is viewed by some as constituting research, as is evaluation of vaccine effectiveness, another critical public health function which is not exclusively a research objective. Modifying the text of the Common Rule to explicitly include examples such as these or providing an interpretation of the surveillance/situational awareness exemption that includes these activities would considerably improve the ability for public health agencies to maintain situational awareness and set priorities, quite in line with the spirit of the exemption.



- 3 *Streamlined IRBs*. Where projects fall under human subjects research designation and require IRB review, generic, pathogen-agnostic study protocols for specific populations would accelerate research by decreasing the time to first data. Preapproval of a range of well-defined studies targeted at emergency response and using specific data sets would retain critical protections for human subjects, while allowing high-urgency protocols to be “on the shelf” and ready for fast rollout. As an additional feature, such preapproved protocols would also free up valuable researcher and IRB reviewer time, having converted per-submission efforts during a pandemic into fixed-cost efforts ahead of time. Moreover, designing consent processes for normal “peacetime” studies to allow use of data and specimens in public health emergencies could avoid some of the delays experienced during COVID-19 with, for example, use of the Seattle Flu Study’s specimens to understand early transmission of the virus in the United States (12).
- 4 *Case reporting standardization*. Tracking and understanding outbreaks, particularly at their beginnings, rely on case reporting. Ideally, public health efforts would follow case trends over time and across regions, compare and monitor clinical features including disease progression, resolution, and response to interventions, and track demographics of infected individuals. But lack of standardization of case report protocols, parallel or overlapping surveillance systems that result in duplication (often with varying completed fields for the same case), and inadequate systems for incorporating updates as further information about a case accumulates after the initial report, among other issues, result in case report data that require much time and effort to sort through. Worse, these issues may render some fraction of case reports unreliable. Improving national surveillance systems to be more uniform, timely, and flexible could serve both local and national surveillance needs would help address these issues (60).
- 5 *Dataset accessibility*. In the absence of a United States national healthcare system, research into the distribution and burden of clinical conditions depends on academic or private data streams, including surveys and surveillance systems constructed to address specific questions, and databases of insurance claims which represent utilization of the healthcare system. Insurance claims datasets include those from (i) employer-based insurance companies (e.g., MarketScan) (ii) all-payer claims databases available in some states (which, since a 2016 Supreme Court decision (71), are no longer necessarily ‘all payers’), (iii) Medicare for individuals over 65 years of age, (iv) Medicaid, which provides coverage to over 18% of the United States; and (v) data bases for other specific populations, such as those of the Veterans Affairs Health System, the Indian Health Service, and the Department of Defense. While these datasets can provide an important window into healthcare use across demographics and geography, access to these datasets can be expensive and time and labor intensive. Gaining access to Medicaid data, for example, presents a substantial burden, since this has to be acquired on a state-by-state basis. Establishing standing flexible DUAs for these datasets, with a single agreement across states for Medicaid and other state-controlled data, could enable both routine surveillance-type analysis to identify trends (such as disease outbreaks or

patterns of disease spread) and to evaluate the impact of clinical and public health interventions.

- 6 *Public health-health care partnerships*: While the United States does not have a national health system for all, it has a wealth of data in the health care sector that can inform public health decision making. Multiple studies at the Centers for Disease Control and Prevention (CDC) and other institutions harnessed such data to provide estimates of key quantities such as vaccine effectiveness (72, 73), variant severity (74), and antiviral effectiveness (75), as well as for surveillance of disease burden and its correlates (76). Building public health partnerships with the health care sector in advance to set in place the administrative, information technology, and financial arrangements to make possible high-quality analyses of this sort rapidly (and automated where possible) would greatly increase the timeliness and value of such efforts (77).

## Strengthening personnel and research ties, including globally

In response to the COVID-19 emergency and the need for expertise to gather, analyze, and interpret evidence around the pandemic and the clinical and public health responses, many academics put aside their usual research programs to engage directly in public health activities and research. The close interactions between academics and local, state, and national public health officials were often productive and important for guiding the pandemic response but raised issues that should be addressed before the next pandemic. These include the *ad hoc* way in which these academic-public health collaborations came into being, the lack of uniformity of access to academics with appropriate expertise across states, and the misalignment of incentives between public health and academic work.

Ideally, academia-public health collaborations can be rapidly scaled up in times of need through established pathways. One idea is to create a “rotator” program, in which academics (and potentially those in training, including doctoral students and postdoctoral fellows) are embedded within public health agencies—and similarly public health officials are embedded within academic groups—for intervals (such as 3 or 6 months) that build familiarity, collegiality, and accessibility. The LEAP fellowship through the Infectious Disease Society of America (78) and the joint Infectious Diseases/EIS fellowship (79) programs are efforts in this direction. Another approach is to establish an academic career path in which some fraction of time and effort are based in public health activities, analogous to academic medicine paths in which researchers spend some fraction of their time doing clinical work. Cooperative agreements established in 2023 between the CDC’s Center for Forecasting and Outbreak Analytics (CFA) and academic and other groups include a surge provision whereby the performers on these agreements would provide scientific assistance in times of crisis. Relatedly, an official “public health reserve corps” of analysts and modelers could provide a workforce available to be called up to prepare for and respond to emergencies. Formal recognition of these paths as prestigious and vital, and placing value on these activities within the academic systems of rewards and incentives, will be key to success (Box 1).

Often tools developed for one public health jurisdiction solve common problems and could be readily implemented in other jurisdictions, underscoring the importance of making code open access and ideally making tools generalizable. This would have the benefits of “not reinventing the wheel” and allowing those jurisdictions lacking local expertise access to useful tools. A curated clearinghouse of such tools, organized by research problem and perhaps hosted by CFA offers one strategy for providing access. Broad efforts to create and maintain state-of-the-art tools for epidemiologic modeling, such as Epiverse (82) and Recon are encouraging developments in this space (83).

We have focused on domestic systems in the United States, but that international cooperation is essential for multiple reasons. Maintaining systems to detect new pathogens at multiple locations in a globalized world will speed detection, facilitating timely responses including development of countermeasures (25) and providing early warning to other jurisdictions (84). For pathogen characterization, as we note in Box 2, each jurisdiction can benefit from findings in other jurisdictions on quantities that are relatively similar across populations, including for example the effectiveness of vaccines and treatments, as well as certain features of infection natural history (e.g., viral load kinetics). Strengthening these capacities globally means that locations that have exceptional data and study infrastructure and/or early experience with a pathogen or variant, can contribute to the global store of knowledge of pathogen characteristics (49, 54, 85–87).

## Educating the consumers

Tools are best deployed by those who understand how they work, how they are limited, and how they can be modified to improve their applicability to specific situations. Workforce development mandates to build subject matter expertise within public health departments, such as through CDC efforts via the Office of Science and the Office of Advanced Molecular Detection and through fellowships such as the CDC/Association of Public Health Laboratories (APHL) program, are critical efforts. While waiting for these workforce programs to get up and running, and since public health officials may not stay current with the frontiers of analytical and modeling methods, opportunities for regular formal trainings should be developed. For example, meetings such as the Council of State and Territorial Epidemiologists Annual Conference could provide a forum for workshops on advances in modeling, genomic epidemiology, and other fields. Relatedly, encouraging public health officials to attend field-specific meetings (e.g., Epidemics, Applied Bioinformatics in Public Health Microbiology) could provide opportunities for knowledge sharing, relationship building, and networking across sectors and disciplines.

## Improving knowledge flow

Successful communication of a health agency’s current understanding of a pandemic and outlook for its future requires a combination of approaches to communicate different kinds of data and outlooks, for different audiences. It has been suggested that principles for such communications include: thematic structure related to informing key decisions, synthesis of evidence from multiple sources, quantification of uncertainty, inclusion of visualizations as well as text and tables, and inclusion of forward-looking material

(outlooks for the future) (91); another important principle is open access to the data underlying figures in these reports. CDC’s Technical Reports on the Mpox epidemic in 2022 (92) sought to put these principles into practice, explicitly emulating aspects of the United Kingdom Health Security Agency (UKHSA) Technical Briefings from COVID-19 (93). Creating a regular cadence for such reports during an emergency, as was the case in the United Kingdom during the height of the COVID-19 pandemic, may help develop an audience and facilitate knowledge flow.

## Centering equity

The World Health Organization has stated that “Countries have an obligation to develop appropriate, feasible, sustainable public health surveillance systems” to ensure that the health needs of populations are quantified so that they can be addressed. While there has been a disproportionate impact of COVID-19 on racial and ethnic minorities and on socioeconomically disadvantaged populations in the United States (94) and elsewhere (95, 96), a persistent problem is that race/ethnicity data are too often missing from surveillance data. Under the plausible hypothesis that those with missing data on race/ethnicity are among the most disadvantaged, these missing data could lead to attenuated estimates of the degree of inequities; whether or not this is the case, it reduces the quality of the estimates by adding uncertainty. Improving the completeness of race-ethnicity reporting is an urgent priority to maximize the value of surveillance data to enhance health equity. Some symposium participants, while agreeing with the need for better reporting of such data, argued that in the presence of ongoing racial segregation, ZIP code or other geographic tags can be a useful proxy when such data are unavailable. Early maps of COVID-19 in New York City showed a higher prevalence of COVID-19 diagnoses in areas that were home to largely Black and Hispanic populations, as well as some areas where most residents were White and many believed to be first-responders (68). This was reflected in elevated COVID-19 mortality rates among Black, Hispanic, and Native American populations compared to White populations throughout the United States, particularly in the early waves of the pandemic (97).

The Presidential COVID-19 Health Equity Task Force final report from 2021 (98) recommends strategies for enhancing equity in data, analytics, and research. These recommendations include standardizing demographic and socioeconomic categories, supporting equity-centered data collection, tracking and reporting on health outcomes for people in congregate and high-risk settings, and research and analysis on behavioral health. In a similar spirit, for any clinical or public health intervention, one should ask in what ways the intervention exacerbates or alleviates inequities. To put this into practice, one goal is the development of real-time metrics that inform municipalities and states on the equity of interventions and health outcomes, enabling adjustments and responses to keep equity at the forefront of intervention decisions.

Other important examples of the links between surveillance and health equity were discussed during the symposium. While documenting disparate impacts is a necessary starting point, identifying appropriate measures to rectify these inequities will often require understanding where, why, when, and how they arise (99). An early example in the U.S. was a documentation that higher SARS-CoV-2 prevalence among

**BOX 1 Modeling and analytic support for STLT jurisdictions: the role of academic groups.**

COVID-19 stimulated numerous collaborations between STLT health authorities and university and other research institutions to support decision making with modeling, analytics, and forecasts. These took multiple forms, ranging from the establishment of advisory councils to mayors and governors, to bilateral collaborations (80, 81) and formal consortia (<https://modelingconsortium.ucsf.edu/>). However, there are stark differences among jurisdictions in the number of such academic groups within the jurisdiction and/or with existing or prior links with the jurisdiction's health department, creating inequities in access to this kind of advice. The benefits of working with academic partners can include local knowledge and the capacity to surge efforts in an emergency. Potential barriers to such collaborations that should be addressed up front where possible include academics' need and incentives to publish, which may compete for time with their role in decision support, as well as the demands of academic schedules, whereby, for example, a key analyst on a project may have to devote effort to exams at times when they are needed for decision support. Administrative preparedness in the form of preexisting data use agreements can vastly accelerate these efforts.

Establishment of trust is essential to the success of academic – STLT collaborations. Elected and health officials at the symposium noted the repeated challenges of figuring out which models and modelers to trust, both locally and nationally. Participants observed that academic collaborations were most effective when there was a pre-existing relationship between the groups and the jurisdictions, and noted the benefits to both parties of cultivating these relationships in “peacetime” through collaboration on non-pandemic activities. Academic groups' ability to speak freely can lend credibility and objectivity to their analyses; however, trust can be undermined if academic groups with access to limited, publicly available data release analyses in publications or preprints that may be inconsistent with more complete data that are available to health departments but not publicly available. Frequent contact to share tentative conclusions and compare them against the evolving understanding of health officials can enhance the quality of analyses by incorporating more complete data, if these can be shared, and can enhance the trust between the parties, improving future interactions. When such interactions work well, they do not stifle the conclusions of academic groups but rather ensure that these conclusions are based on the best current understanding and to ensure that health officials are aware of what is being published about data from their jurisdictions. Academic incentives and structures are particularly not suited for routinely repeated analyses, such as reproductive number estimation, nowcasting, and forecasting, although academic centers have played key roles in these areas for over 2 years during COVID-19. Automation, as in the California consortium's dashboard, is one solution. CDC's Center for Forecasting and Outbreak Analytics is beginning to take on some of these tasks and will increasingly serve as a focal point for such repeated, real-time analyses.

**BOX 2 National insight from local evidence.**

Implementation of public health policies is a state/territorial/tribal and local responsibility in the U.S., as we noted above. Infectious disease surveillance is also decentralized, often with two levels of reporting (local/county and state/territorial) below the national level. From the perspective of national decision makers seeking a clear picture of an unfolding pandemic, decentralized surveillance has obvious limitations, particularly in a setting where data systems and data use agreements vary across jurisdictions. Efforts are underway, and should be expanded, to improve the speed, completeness, and accuracy of data flowing from states, localities, and health care systems to the CDC and other federal actors. Such efforts are essential for timely situational awareness and for calibrating the outputs of scenarios and forecasts to granular (state or county-level) data to form a national picture.

While incidence, prevalence, and health care burden are intrinsically local quantities that need to be estimated everywhere and over time, many aspects of surveillance and associated epidemiology are generalizable, such that findings in one local jurisdiction can inform control measures everywhere. These include characteristics of the pathogen, such as severity and natural history; and characteristics of countermeasures, such as test sensitivity and the effectiveness of drugs and vaccines. For these purposes, local conditions can facilitate detailed characterization that may not be possible, but also may not be necessary, on a larger geographic scale.

Some of the earliest evidence of low severity for the 2009 H1N1 influenza pandemic came from a study at the University of Delaware, where a comparatively self-contained population could be studied in detail (88). We noted in a postmortem of that pandemic that the findings from that study were not widely known until months later because of limited dissemination (4), arguably prolonging the state of alarm unnecessarily during that pandemic. In COVID-19, early findings of asymptomatic/presymptomatic infection and likely transmission from studies in a nursing home and a cruise ship, respectively (19, 89), were documented very early and widely disseminated, but still did not fully inform control measures.

In many other cases, detailed surveillance and epidemiology in local jurisdictions or health systems provided evidence of national and international importance. A few examples included:

- Evidence from the Yukon-Kuskokwim (Alaska) Health Corporation about the persistence of antigen test positivity 5 or more days after initial positive test or symptom onset during the early Omicron era (90)
- Evidence from the Kaiser Permanente Southern California health system about the relative clinical severity of Omicron BA.1 variant compared to Delta before it and BA.2 after (74)
- others

Each of these provided evidence that could be generalized beyond the location where it was generated, because it concerned generalizable features of the infection or countermeasures based on its biology. The degree to which these investigations informed policy and guidelines varied, indicating a need for a systematic approach to disseminating findings of wide importance and updating guidance in a way that reflects the totality of data.

mothers admitted for labor and delivery in New York City was associated with residence in boroughs with smaller reductions in mobility, suggesting inability to work from home as a potential driver of risk (68). Subsequent

modeling work explained racial disparities in infection rates in U.S. cities as a consequence of higher exposure by minority groups not only to infection generally, but particularly to more crowded venues with higher

infection risk (58). The age distribution of mortality by race/ethnicity (100), with its skew to younger ages among Black, Hispanic, and Indigenous individuals, also pointed to an increased exposure risk. The United Kingdom Government's Race Disparity Unit published a series of reports through the first 2 years of the pandemic enumerating hypotheses for mechanisms to explain disparate impacts, stating the current evidence related to these hypotheses, and recommending actions to address these drivers of higher incidence and severity in racial and ethnic minorities (60). In the United States, such studies may require linkage of disparate data bases to identify where disparities arise during the cascade of care (101), a strategy that has long been useful in HIV/AIDS surveillance to understand loss points in the continuum of care (102, 103). For COVID, a full cascade would require an estimate of the actual number of infected individuals, the number of people who have been identified by testing (reflecting under-diagnosis), the number treated when treatment became available (reflecting under-treatment), the number hospitalized (reflecting access to care and disease severity), and the number of fatalities, jointly stratified by race and ethnicity, age group and sex. Ascertaining these would require both modeling-based estimates and data from multiple sources (e.g., clinical laboratories and vital registries). For example, an analysis from a New York City hospital suggested no racial difference in case fatality among hospitalized patients, supporting the idea that racial differences in exposure (more infections) rather than racial differences in outcome contributed to racial differences in overall mortality (104).

Other sources of inequity can affect case ascertainment and thus identification of opportunities for intervention. Geographic and temporal variation in testing effort in the U.S. was very large, resulting in difficulties in comparing incidence across jurisdictions. Rural areas were often the least able to access testing, though there were important exceptions (90). Notably, the use of random sampling stratified by geography mitigated this problem significantly in the United Kingdom (85), though it did not solve it entirely because participation was of necessity voluntary. Equity considerations may change as public health authorities rely on new data sources; for example, mobility estimates may depend on smartphone ownership, while wastewater surveillance for pathogen abundance will be unavailable in areas using septic systems (29, 105).

## Expanding the range of data types

As described above, any health system, but particularly one as decentralized as that of the U.S., benefits from the ability to ingest and synthesize multiple types of data. Increased use of wastewater data (106, 107) has contributed to early warning of rising infection incidence and to surveillance for new variants. Further work to standardize collection and better define the quantitative relationships between true infection incidence and total and variant-specific concentrations of viral genomes in wastewater is needed to improve the value of such data, as well as a clear mapping of where it will not be informative, such as areas using septic systems. Likewise, mobility data from various sources (58) can be useful in informing strategies for disease monitoring and surveillance, modeling disease spread, and guiding interventions. Immune measures from serology provide a window onto past infection and a lens onto the landscape of risk (108). Here, further work is needed to ensure data standardization and accuracy as well as routine and frequent updating to capture important temporal variations. Such new forms of data may also raise

privacy considerations that have not entirely been solved (109, 110).

Crowdsourced and survey data (111, 112) can provide important insights into behaviors that affect the interpretations of other data; for example, the increasing prevalence of self-testing using antigen tests for COVID-19 reduces the utility of PCR-positive case counts.

A key to making use of this expanded range of data types is solving the problem of how to synthesize multiple data types into a single estimate of a quantity of interest, accounting for the different properties of each data type (76), including understanding the different biases that will affect each data stream. Significant further work is needed to advance the ability to do this in real time. A related but distinct problem is how to link data across data systems to understand the continuum of care and otherwise improve inference about the course of individual cases.

## Expanding the range of data sources

The use of claims data from health care payers (insurers) and electronic medical records from providers has exploded in many areas of health services research. There have been some notable examples of such data for surveillance to address the questions described in this report (53, 72–74), but in the United States there remains untapped potential to expand such efforts and improve their timeliness. This will require building relationships between public health entities and health care systems in their jurisdictions, including relationships between scientific investigators in each sector with regular discussions for bidirectional learning. In the spirit of administrative preparedness above, this will require up-front planning of master agreements to move resources in a timely fashion to address pressing questions. Health providers and public health have suffered from a “two cultures” challenge that results in the need to expand public health training of investigators and other personnel in health systems, acknowledge the contributions of health systems to community benefits, and find ways to produce incentives so that contributing to public health surveillance aligns with the business interests of health systems. Medical examiners and coroners are another group that has been disconnected from public health but with whom cooperation can enhance and help to calibrate surveillance for pathogen-specific deaths, as illustrated by some examples both domestically (113) and abroad (114).

As we described above, new data sources become useful in proportion to our understanding of their “normal” behavior. As we expand the range of data types, it will be essential to monitor new data streams and continue to monitor old ones outside of epidemic periods to establish a baseline that can be used to calibrate signals of new outbreaks and estimate the exceedance caused by the ongoing transmission of novel pathogens (115).

## Conclusion

Data and modeling needs change over the course of a pandemic and vary by the jurisdictional dimensions, requiring anticipatory, rapid, dynamic, and locally adapted and scaled activities to optimize pandemic management and population health. Here, we have sought to describe concepts, tools, and strategies to address those needs, building on those enacted during the COVID-19 pandemic and those that could have

facilitated this work. While not a comprehensive list, we hope that the ideas we propose and envision serve as a useful resource and guide in efforts to manage ongoing infectious diseases challenges and preparedness for the inevitable next pandemic.

## Author contributions

ML: Conceptualization, Methodology, Project administration, Writing – original draft, Writing – review & editing. MB: Investigation, Writing – review & editing. JB: Investigation, Writing – review & editing. PE: Investigation, Writing – review & editing. DE: Investigation, Writing – review & editing. MG: Investigation, Writing – review & editing. JH: Investigation, Writing – review & editing. MJ: Investigation, Writing – review & editing. SK: Investigation, Writing – review & editing. DL: Investigation, Writing – review & editing. JeL: Investigation, Writing – review & editing. JuL: Investigation, Writing – review & editing. RL: Investigation, Writing – review & editing. DM: Investigation, Writing – review & editing. LCM: Investigation, Writing – review & editing. CM: Investigation, Writing – review & editing. LAM: Investigation, Writing – review & editing. SO: Investigation, Writing – review & editing. CQ: Investigation, Writing – review & editing. AB: Investigation, Writing – review & editing. NR: Investigation, Writing – review & editing. SR: Investigation, Writing – review & editing. RR: Investigation, Writing – review & editing. MS: Investigation, Writing – review & editing. RS: Investigation, Writing – review & editing. RBS: Investigation, Writing – review & editing. DS: Investigation, Writing – review & editing. ST: Investigation, Writing – review & editing. JV: Investigation, Writing – review & editing. YG: Funding acquisition, Investigation, Project administration, Writing – original draft, Writing – review & editing, Conceptualization.

## References

- Sherratt K, Gruson H, Grah R, Johnson H, Niehus R, Prasse B, et al. Predictive performance of multi-model ensemble forecasts of COVID-19 across European nations. *eLife*. (2023) 12:1916. doi: 10.7554/eLife.81916
- Cramer EY, Huang Y, Wang Y, Ray EL, Cornell M, Bracher J, et al. The United States COVID-19 forecast hub dataset. *Sci Data*. (2022) 9:462. doi: 10.1038/s41597-022-01517-w
- Lopez VK, Cramer EY, Pagano R, Drake JM, O'Dea EB, Adey M, et al. Challenges of COVID-19 case forecasting in the US, 2020–2021. *PLoS Comput Biol*. (2024) 20:e1011200. doi: 10.1371/journal.pcbi.1011200
- Lipsitch M, Finelli L, Heffernan RT, Leung GM, Redd SC. Improving the evidence base for decision making during a pandemic: the example of 2009 influenza A/H1N1 [research support, N.I.H., extramural]. *Biosecure Bioterror*. (2011) 9:89–115. doi: 10.1089/bsp.2011.0007
- Scalia Tomba G, Wallinga J. A simple explanation for the low impact of border control as a countermeasure to the spread of an infectious disease. *Math Biosci*. (2008) 214:70–2. doi: 10.1016/j.mbs.2008.02.009
- Quilty BJ, Clifford S, Flasche S, Eggo RMCMMID nCoV working group2. Effectiveness of airport screening at detecting travellers infected with novel coronavirus (2019-nCoV). *Euro Surveill*. (2020) 25:80. doi: 10.2807/1560-7917.ES.2020.25.5.2000080
- Cowling BJ, Lau LL, Wu P, Wong HW, Fang VJ, Riley S, et al. Entry screening to delay local transmission of 2009 pandemic influenza A (H1N1) [research support, N.I.H., extramural research support, non-U.S. Gov't]. *BMC Infect Dis*. (2010) 10:82. doi: 10.1186/1471-2334-10-82
- Li J, Hosegood I, Powell D, Tschärke B, Lawler J, Thomas KV, et al. A global aircraft-based wastewater genomic surveillance network for early warning of future pandemics. *Lancet Glob Health*. (2023) 11:e791–5. doi: 10.1016/S2214-109X(23)00129-8
- Carrión M, Madoff LC. ProMED-mail: 22 years of digital surveillance of emerging infectious diseases. *Int Health*. (2017) 9:177–83. doi: 10.1093/inthealth/ihx014
- Tisza MJ, Hanson BM, Clark JR, Wang L, Payne K, Ross MC, et al. Virome sequencing identifies H5N1 avian influenza in wastewater from nine cities. *medRxiv*. (2024) 2024:24307179. doi: 10.1101/2024.05.10.24307179
- Ma KC, Shirk P, Lambrou AS, Hassell N, Zheng XY, Payne AB, et al. Genomic surveillance for SARS-CoV-2 variants: circulation of omicron lineages - United States, January 2022–may 2023. *MMWR Morb Mortal Wkly Rep*. (2023) 72:651–6. doi: 10.15585/mmwr.mm7224a2
- Boeckh M, Chu HY, Englund JA, Lockwood CM, Nickerson DA, Shendure J, et al. The Seattle flu study: when regulations hinder pandemic surveillance. *Nat Med*. (2022) 28:7–8. doi: 10.1038/s41591-021-01587-0
- Yang H, Surer O, Duque D, Morton DP, Singh B, Fox SJ, et al. Design of COVID-19 staged alert systems to ensure healthcare capacity with minimal closures. *Nat Commun*. (2021) 12:3767. doi: 10.1038/s41467-021-23989-x
- Fox SJ, Lachmann M, Tec M, Pasco R, Woody S, Du Z, et al. Real-time pandemic surveillance using hospital admissions and mobility data. *Proc Natl Acad Sci USA*. (2022) 119:119. doi: 10.1073/pnas.2111870119
- CDC. *Health Department Governance*. (2022). Available at: <https://www.cdc.gov/publichealthgateway/sites/governance/index.html>.
- Gupta S, Nguyen T, Raman S, Lee B, Lozano-Rojas F, Bento A, et al. Tracking public and private responses to the COVID-19 epidemic: evidence from state and local government actions. *Am J Health Econ*. (2021) 7:361–404. doi: 10.1086/716197
- Du Z, Xu X, Wu Y, Wang L, Cowling BJ, Meyers LA. Serial interval of COVID-19 among publicly reported confirmed cases. *Emerg Infect Dis*. (2020) 26:1341–3. doi: 10.3201/eid2606.200357
- Tsang TK, Wu P, Lin Y, Lau EHY, Leung GM, Cowling BJ. Effect of changing case definitions for COVID-19 on the epidemic curve and transmission parameters in mainland China: a modelling study. *Lancet Public Health*. (2020) 5:e289–96. doi: 10.1016/S2468-2667(20)30089-X

## Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This project has been funded (in part) by contract 200–2016–91779 with the Centers for Disease Control and Prevention (CDC).

## Conflict of interest

RS was employed by company Siemens Healthcare Diagnostics. JV was employed by company SIGA Technologies.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Author disclaimer

The findings, conclusions, and views expressed are those of the author(s) and do not necessarily represent the official position of the CDC.

19. Emery JC, Russell TW, Liu Y, Hellewell J, Pearson CA, Group CC-W et al. The contribution of asymptomatic SARS-CoV-2 infections to transmission on the Diamond princess cruise ship. *eLife*. (2020) 9:9. doi: 10.7554/eLife.58699
20. Lu J, Gu J, Li K, Xu C, Su W, Lai Z, et al. COVID-19 outbreak associated with air conditioning in restaurant, Guangzhou, China, 2020. *Emerg Infect Dis*. (2020) 26:1628–31. doi: 10.3201/eid2607.200764
21. Addetia A, Crawford KHD, Dingens A, Zhu H, Roychoudhury P, Huang ML, et al. Neutralizing antibodies correlate with protection from SARS-CoV-2 in humans during a fishery vessel outbreak with a high attack rate. *J Clin Microbiol*. (2020) 58:20. doi: 10.1128/JCM.02107-20
22. Lipsitch M, Lajous M, O'Hagan JJ, Cohen T, Miller JC, Goldstein E, et al. Use of cumulative incidence of novel influenza A/H1N1 in foreign travelers to estimate lower bounds on cumulative incidence in Mexico. *PLoS One*. (2009) 4:e6895. doi: 10.1371/journal.pone.0006895
23. Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, Hollingsworth TD, et al. Pandemic potential of a strain of influenza A (H1N1): early findings. *Science*. (2009) 324:1557–61. doi: 10.1126/science.1176062
24. Niehus R, De Salazar PM, Taylor AR, Lipsitch M. Using observational data to quantify bias of traveller-derived COVID-19 prevalence estimates in Wuhan, China. *Lancet Infect Dis*. (2020) 20:803–8. doi: 10.1016/S1473-3099(20)30229-2
25. De Salazar PM, Niehus R, Taylor A, Buckee CO, Lipsitch M. Identifying locations with possible undetected imported severe acute respiratory syndrome coronavirus 2 cases by using importation predictions. *Emerg Infect Dis*. (2020) 26:1465–9. doi: 10.3201/eid2607.200250
26. Bento AI, Nguyen T, Wing C, Lozano-Rojas F, Ahn YY, Simon K. Evidence from internet search data shows information-seeking responses to news of local COVID-19 cases. *Proc Natl Acad Sci USA*. (2020) 117:11220–2. doi: 10.1073/pnas.2005335117
27. Kishore N, Taylor AR, Jacob PE, Vembar N, Cohen T, Buckee CO, et al. Evaluating the reliability of mobility metrics from aggregated mobile phone data as proxies for SARS-CoV-2 transmission in the USA: a population-based study. *Lancet Digit Health*. (2022) 4:e27–36. doi: 10.1016/S2589-7500(21)00214-4
28. National Academies of Science Engineering and Medicine. Wastewater-based disease surveillance for Public Health action. Washington, DC: National Academies of Science Engineering and Medicine (2023).
29. Yu Q, Olesen SW, Duvall C, Grad YH. Assessment of sewer connectivity in the United States and its implications for equity in wastewater-based epidemiology. *PLoS Glob Public Health*. (2024) 4:e0003039. doi: 10.1371/journal.pgph.0003039
30. Hay JA, Kennedy-Shaffer L, Kanjilal S, Lennon NJ, Gabriel SB, Lipsitch M, et al. Estimating epidemiologic dynamics from cross-sectional viral load distributions. *Science*. (2021) 373:635. doi: 10.1126/science.abh0635
31. Kissler SM, Hay JA, Fauver JR, Mack C, Tai CG, Anderson DJ, et al. Viral kinetics of sequential SARS-CoV-2 infections. *Nat Commun*. (2023) 14:6206. doi: 10.1038/s41467-023-41941-z
32. Kissler SM, Fauver JR, Mack C, Tai CG, Breban MI, Watkins AE, et al. Viral dynamics of SARS-CoV-2 variants in vaccinated and unvaccinated persons. *N Engl J Med*. (2021) 385:2489–91. doi: 10.1056/NEJMc2102507
33. Hay JA, Kissler SM, Fauver JR, Mack C, Tai CG, Samant RM, et al. Quantifying the impact of immune history and variant on SARS-CoV-2 viral kinetics and infection rebound: a retrospective cohort study. *eLife*. (2022) 11:11. doi: 10.7554/eLife.81849
34. Healthmap Homepage. (n.d.). Available at: <http://healthmap.org/en/>.
35. McGough SF, Brownstein JS, Hawkins JB, Santillana M. Forecasting Zika incidence in the 2016 Latin America outbreak combining traditional disease surveillance with search, social media, and news report data. *PLoS Negl Trop Dis*. (2017) 11:e0005295. doi: 10.1371/journal.pntd.0005295
36. Santillana M, Nguyen AT, Dredze M, Paul MJ, Nsoesie EO, Brownstein JS. Combining search, social media, and traditional data sources to improve influenza surveillance. *PLoS Comput Biol*. (2015) 11:e1004513. doi: 10.1371/journal.pcbi.1004513
37. Kogan NE, Clemente L, Liautaud P, Kaashoek J, Link NB, Nguyen AT, et al. An early warning approach to monitor COVID-19 activity with multiple digital traces in near real time. *Sci Adv*. (2021) 7:989. doi: 10.1126/sciadv.abd6989
38. Jones JM, Stone M, Sulaeman H, Fink RV, Dave H, Levy ME, et al. Estimated US infection- and vaccine-induced SARS-CoV-2 Seroprevalence based on blood donations, July 2020–May 2021. *JAMA*. (2021) 326:1400–9. doi: 10.1001/jama.2021.15161
39. Follmann D, Janes HE, Chu E, Jayashankar L, Petropoulos CJ, Serebryanny L, et al. Kinetics of the antibody response to symptomatic SARS-CoV-2 infection in vaccinated and unvaccinated individuals in the blinded phase of the mRNA-1273 COVID-19 vaccine efficacy trial. *Open Forum Infect Dis*. (2023) 10:ofad069. doi: 10.1093/ofid/ofad069
40. Luoma E, Rohrer R, Parton H, Hughes S, Omeregbe E, Taki F, et al. Notes from the Field: epidemiologic characteristics of SARS-CoV-2 recombinant variant XBB.1.5 - new York City, November 1, 2022–January 4, 2023. *MMWR Morb Mortal Wkly Rep*. (2023) 72:212–4. doi: 10.15585/mmwr.mm7208a4
41. Chitwood MH, Russi M, Gunasekera K, Havumaki J, Klaassen F, Pitzer VE, et al. Reconstructing the course of the COVID-19 epidemic over 2020 for US states and counties: results of a Bayesian evidence synthesis model. *PLoS Comput Biol*. (2022) 18:e1010465. doi: 10.1371/journal.pcbi.1010465
42. Lison A, Abbott S, Huisman J, Stadler T. Generative Bayesian modeling to nowcast the effective reproduction number from line list data with missing symptom onset dates. *PLoS Comput Biol*. (2024) 20:e1012021. doi: 10.1371/journal.pcbi.1012021
43. Howerton E, Contamin L, Mullany LC, Qin M, Reich NG, Bents S, et al. Evaluation of the US COVID-19 scenario modeling hub for informing pandemic response under uncertainty. *Nat Commun*. (2023) 14:7260. doi: 10.1038/s41467-023-42680-x
44. Kahn R, Holmdahl I, Reddy S, Jernigan J, Mina MJ, Slayton RB. Mathematical modeling to inform vaccination strategies and testing approaches for COVID-19 in nursing homes. *Clin Infect Dis*. (2021) 74:597–603. doi: 10.1093/cid/ciab517
45. Pung R, Clapham HE, Russell TW, Group CC-WLee VJ, Kucharski AJ. Relative role of border restrictions, case finding and contact tracing in controlling SARS-CoV-2 in the presence of undetected transmission: a mathematical modelling study. *BMC Med*. (2023) 21:802. doi: 10.1186/s12916-023-02802-0
46. Quilty BJ, Clifford S, Hellewell J, Russell TW, Kucharski AJ, Flasche S, et al. Centre for the Mathematical Modelling of infectious diseases C-wg. Quarantine and testing strategies in contact tracing for SARS-CoV-2: a modelling study. *Lancet Public Health*. (2021) 6:e175–83. doi: 10.1016/S2468-2667(20)30308-X
47. Howerton E, Contamin L, Mullany LC, Qin M, Reich NG, Bents S, et al. Informing pandemic response in the face of uncertainty: an evaluation of the U.S. COVID-19 scenario modeling hub. *medRxiv*. (2023) 2023:23291998. doi: 10.1101/2023.06.28.23291998
48. Singanayagam A, Hakki S, Dunning J, Madon KJ, Crone MA, Koycheva A, et al. Community transmission and viral load kinetics of the SARS-CoV-2 delta (B.1.617.2) variant in vaccinated and unvaccinated individuals in the UK: a prospective, longitudinal, cohort study. *Lancet Infect Dis*. (2022) 22:183–95. doi: 10.1016/S1473-3099(21)00648-4
49. Kissler SM, Fauver JR, Mack C, Olesen SW, Tai C, Shiu KY, et al. Viral dynamics of acute SARS-CoV-2 infection and applications to diagnostic and public health strategies. *PLoS Biol*. (2021) 19:e3001333. doi: 10.1371/journal.pbio.3001333
50. Block JP, Boehmer TK, Forrest CB, Carton TW, Lee GM, Ajani UA, et al. Cardiac complications after SARS-CoV-2 infection and mRNA COVID-19 vaccination - PCORnet, United States, January 2021–January 2022. *MMWR Morb Mortal Wkly Rep*. (2022) 71:517–23. doi: 10.15585/mmwr.mm7114e1
51. McNeil MM, Gee J, Weintraub ES, Belongia EA, Lee GM, Glanz JM, et al. The vaccine safety datalink: successes and challenges monitoring vaccine safety. *Vaccine*. (2014) 32:5390–8. doi: 10.1016/j.vaccine.2014.07.073
52. Lewnard JA, Hong VX, Patel MM, Kahn R, Lipsitch M, Tartof SY. Clinical outcomes associated with omicron (B.1.1.529) variant and BA.1/BA.1.1 or BA.2 subvariant infection in southern California. *Nat Med*. (2022) 28:1933–1943.
53. Dickerman BA, Gerlovina H, Madenci AL, Kurgansky KE, Ferolito BR, Figueroa Muniz MJ, et al. Comparative effectiveness of BNT162b2 and mRNA-1273 vaccines in U.S. *Vet N Engl J Med*. (2022) 386:105–15. doi: 10.1056/NEJMoa2115463
54. Dagan N, Barda N, Kepten E, Miron O, Perchik S, Katz MA, et al. BNT162b2 mRNA COVID-19 vaccine in a Nationwide mass vaccination setting. *N Engl J Med*. (2021) 384:1412–23. doi: 10.1056/NEJMoa2101765
55. CDC. CDC's Influenza Vaccine Effectiveness Networks. (2021). Available at: <https://www.cdc.gov/flu/vaccines-work/vaccine-effectiveness-networks.htm>.
56. CDC. COVID-19 Vaccinations in the United States. (2024). Available at: <https://covid.cdc.gov/covid-data-tracker/#vaccination-states-jurisdictions>.
57. US Census Bureau. Household Pulse Survey COVID-19 Vaccination Tracker. (2021). Available at: <https://www.census.gov/library/visualizations/interactive/household-pulse-survey-covid-19-vaccination-tracker.html>.
58. Chang S, Pierson E, Koh PW, Gerardin J, Redbird B, Grusky D, et al. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature*. (2021) 589:82–7. doi: 10.1038/s41586-020-2923-3
59. Carnegie-Mellon University. (2021). Available at: <https://delphi.cmu.edu/covidcast/>.
60. UK Government Equality Hub and Race Disparity Unit. *Final report on progress to address COVID-19 health inequalities*. (2021). Available at: <https://www.gov.uk/government/publications/final-report-on-progress-to-address-covid-19-health-inequalities/final-report-on-progress-to-address-covid-19-health-inequalities>.
61. Ma KC, Hale JE, Grad YH, Alter G, Luzuriaga K, Eaton RB, et al. Trends in severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) Seroprevalence in Massachusetts estimated from newborn screening specimens. *Clin Infect Dis*. (2022) 75:e105–13. doi: 10.1093/cid/ciac158
62. Wiegand RE, Deng Y, Deng X, Lee A, Meyer WA3rd, Letovsky S, et al. Estimated SARS-CoV-2 antibody seroprevalence trends and relationship to reported case prevalence from a repeated, cross-sectional study in the 50 states and the District of Columbia, United States–October 25, 2020–February 26, 2022. *Lancet Reg Health Am*. (2023) 18:100403. doi: 10.1016/j.lana.2022.100403

63. Garcia-Carreras B, Hitchings MDT, Johansson MA, Biggerstaff M, Slayton RB, Healy JM, et al. Accounting for assay performance when estimating the temporal dynamics in SARS-CoV-2 seroprevalence in the U.S. *Nat Commun.* (2023) 14:944. doi: 10.1038/s41467-023-37944-5
64. Ng DL, Goldhof GM, Shy BR, Levine AG, Balcerak J, Bapat SP, et al. SARS-CoV-2 seroprevalence and neutralizing activity in donor and patient blood. *Nat Commun.* (2020) 11:4698. doi: 10.1038/s41467-020-18468-8
65. Self WH, Tenforde MW, Stubblefield WB, Feldstein LR, Steingrub JS, Shapiro NI, et al. Decline in SARS-CoV-2 antibodies after mild infection among frontline Health care personnel in a multistate hospital network - 12 states, April-august 2020. *MMWR Morb Mortal Wkly Rep.* (2020) 69:1762-6. doi: 10.15585/mmwr.mm6947a2
66. Si Y, Covello L, Wang S, Covello T, Gelman A. Beyond vaccination rates: a synthetic random proxy metric of Total SARS-CoV-2 immunity Seroprevalence in the community. *Epidemiology.* (2022) 33:457-64. doi: 10.1097/EDE.0000000000001488
67. Covello L, Gelman A, Si Y, Wang S. Routine hospital-based SARS-CoV-2 testing outperforms state-based data in predicting clinical burden. *Epidemiology.* (2021) 32:792-9. doi: 10.1097/EDE.0000000000001396
68. Kissler SM, Kishore N, Prabhu M, Goffman D, Beilin Y, Landau R, et al. Reductions in commuting mobility correlate with geographic differences in SARS-CoV-2 prevalence in new York City. *Nat Commun.* (2020) 11:4674. doi: 10.1038/s41467-020-18271-5
69. US Department of Health and Human Services. *COVID-19 Reported Patient Impact and Hospital Capacity by State Timeseries (RAW)*. (2023). Available at: <https://healthdata.gov/Hospital/COVID-19-Reported-Patient-Impact-and-Hospital-Capacity-62h-syeh>.
70. US Department of Health and Human Services. *Activities Deemed Not to Be Research: Public Health Surveillance* (2018). Available at: <https://www.hhs.gov/ohrp/regulations-and-policy/requests-for-comments/draft-guidance-activities-deemed-not-be-research-public-health-surveillance/index.html>.
71. Gobeille V. *Liberty mutual insurance co*. S.Ct. p. 936 (2016).
72. Accorsi EK, Britton A, Fleming-Dutra KE, Smith ZR, Shang N, Derado G, et al. Association between 3 doses of mRNA COVID-19 vaccine and symptomatic infection caused by the SARS-CoV-2 omicron and Delta variants. *JAMA.* (2022) 327:639-51. doi: 10.1001/jama.2022.0470
73. Link-Gelles R, Levy ME, Gaglani M, Irving SA, Stockwell M, Dascomb K, et al. Effectiveness of 2, 3, and 4 COVID-19 mRNA Vaccine Doses Among Immunocompetent Adults During Periods when SARS-CoV-2 Omicron BA.1 and BA.2/BA.2.12.1 Sublineages Predominated - VISION Network, 10 States, December 2021-June 2022. *MMWR Morb Mortal Wkly Rep.* (2022) 71:931-9. doi: 10.15585/mmwr.mm7129e1
74. Lewnard JA, Hong VX, Patel MM, Kahn R, Lipsitch M, Tartof SY. Clinical outcomes associated with SARS-CoV-2 omicron (B.1.1.529) variant and BA.1/BA.1.1 or BA.2 subvariant infection in Southern California. *Nat Med.* (2022) 28:1933-43. doi: 10.1038/s41591-022-01887-z
75. Shah MM, Joyce B, Plumb ID, Sahakian S, Feldstein LR, Barkley E, et al. Paxlovid associated with decreased hospitalization rate among adults with COVID-19 - United States, April-September 2022. *MMWR Morb Mortal Wkly Rep.* (2022) 71:1531-7. doi: 10.15585/mmwr.mm7148e2
76. Reinhart A, Brooks L, Jahja M, Rumack A, Tang J, Agrawal S, et al. An open repository of real-time COVID-19 indicators. *Proc Natl Acad Sci USA.* (2021) 118:2118. doi: 10.1073/pnas.2111452118
77. Covid Crisis Group. *Lessons from the COVID war: An investigative report*. New York: Public Affairs Books (2023). 352 p.
78. Infectious Diseases Society of America. *The Leaders in Epidemiology, Antimicrobial Stewardship and Public Health Fellowship*. (2023). Available at: [www.idsociety.org/professional-development/fellows-in-training-career-education-center/leap-fellowship/leap-fellowship/](http://www.idsociety.org/professional-development/fellows-in-training-career-education-center/leap-fellowship/leap-fellowship/).
79. Infectious Diseases Society of America. *ID/EIS joint fellowship - New Opportunity!* (2023). Available at: <https://www.idsociety.org/professional-development/fellows-in-training-career-education-center/joint-id-eis-fellowship/id-eis-fellowship/>.
80. Yang W, Kandula S, Huynh M, Greene SK, Van Wye G, Li W, et al. Estimating the infection-fatality risk of SARS-CoV-2 in new York City during the spring 2020 pandemic wave: a model-based analysis. *Lancet Infect Dis.* (2021) 21:203-12. doi: 10.1016/S1473-3099(20)30769-6
81. Greene SK, McGough SF, Culp GM, Graf LE, Lipsitch M, Menzies NA, et al. Nowcasting for real-time COVID-19 tracking in new York City: an evaluation using reportable disease data from early in the pandemic. *JMIR Public Health Surveill.* (2021) 7:e25538. doi: 10.2196/25538
82. Data. (2023). Available at: <https://data.org/initiatives/epiverse/>.
83. R Epidemics Consortium. *RECON*. (2023). Available at: <https://www.repidemicsconsortium.org/>.
84. Gouglas D, Christodoulou M, Hatchett R. The 100 days Mission—2022 global pandemic preparedness summit. *Emerg Infect Dis.* (2023) 29:21142. doi: 10.3201/eid2903.221142
85. Elliott P, Whitaker M, Tang D, Eales O, Steyn N, Bodinier B, et al. Design and implementation of a national SARS-CoV-2 monitoring program in England: REACT-1 study. *Am J Public Health.* (2023) 113:545-54. doi: 10.2105/AJPH.2023.307230
86. Pulliam JRC, van Schalkwyk C, Govender N, von Gottberg A, Cohen C, Groome MJ, et al. Increased risk of SARS-CoV-2 reinfection associated with emergence of omicron in South Africa. *Science.* (2022) 376:4947. doi: 10.1126/science.abn4947
87. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med.* (2020) 382:1199-207. doi: 10.1056/NEJMoa2001316
88. Juliano AD, Reed C, Guh A, Desai M, Dee DL, Kutty P, et al. Notes from the field: outbreak of 2009 pandemic influenza A (H1N1) virus at a large public university in Delaware, April-may 2009. *Clin Infect Dis.* (2009) 49:1811-20. doi: 10.1086/649555
89. Kimball A, Hatfield KM, Arons M, James A, Taylor J, Spicer K, et al. Asymptomatic and Presymptomatic SARS-CoV-2 infections in residents of a long-term care skilled nursing facility - King County, Washington, march 2020. *MMWR Morb Mortal Wkly Rep.* (2020, 13) 69:377-81. doi: 10.15585/mmwr.mm6913e1
90. Lefferts B, Blake I, Bruden D, Hagen MB, Hodges E, Kirking HL, et al. Antigen test positivity after COVID-19 isolation - Yukon-Kuskokwim Delta region, Alaska, January-February 2022. *MMWR Morb Mortal Wkly Rep.* (2022) 71:293-8. doi: 10.15585/mmwr.mm7108a3
91. Lipsitch M, Santillana M. Enhancing situational awareness to prevent infectious disease outbreaks from becoming catastrophic. *Curr Top Microbiol Immunol.* (2019) 424:59-74. doi: 10.1007/82\_2019\_172
92. CDC. *Mpox Technical Reports*. (2021). Available at: <https://www.cdc.gov/poxvirus/mpox/cases-data/technical-report.html>.
93. UKHSA. *Investigation of SARS-CoV-2 variants: technical briefings*. (2021). Available at: <https://www.gov.uk/government/publications/investigation-of-sars-cov-2-variants-technical-briefings>.
94. Feldman JM, Bassett MT. Variation in COVID-19 mortality in the US by race and ethnicity and educational attainment. *JAMA Netw Open.* (2021) 4:e2135967. doi: 10.1001/jamanetworkopen.2021.35967
95. Department of Health and Social Care U. *Technical report on the COVID-19 pandemic in the UK: A technical report for future UK Chief Medical Officers, Government Chief Scientific Advisers, National Medical Directors and public health leaders in a pandemic*. (2023). Available at: <https://www.gov.uk/government/publications/technical-report-on-the-covid-19-pandemic-in-the-uk>.
96. Mena GE, Martinez PP, Mahmud AS, Marquet PA, Buckee CO, Santillana M. Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. *Science.* (2021) 372:5298. doi: 10.1126/science.abg5298
97. Lundberg DJ, Wrigley-Field E, Cho A, Raquib R, Nsoesie EO, Paglino E, et al. COVID-19 mortality by race and ethnicity in US metropolitan and nonmetropolitan areas, march 2020 to February 2022. *JAMA Netw Open.* (2023) 6:e2311098. doi: 10.1001/jamanetworkopen.2023.11098
98. Presidential COVID-19 Health Equity Task Force. *Final Report and Recommendations*. (2021). Available at: [https://www.minorityhealth.hhs.gov/assets/pdf/HETF\\_Report\\_508\\_102821\\_9am\\_508Team%20WIP11-compressed.pdf](https://www.minorityhealth.hhs.gov/assets/pdf/HETF_Report_508_102821_9am_508Team%20WIP11-compressed.pdf)
99. Jones CP. Invited commentary: "race," racism, and the practice of epidemiology. *Am J Epidemiol.* (2001) 154:299-304. doi: 10.1093/aje/154.4.299
100. Bassett MT, Chen JT, Krieger N. Variation in racial/ethnic disparities in COVID-19 mortality by age in the United States: a cross-sectional study. *PLoS Med.* (2020) 17:e1003402. doi: 10.1371/journal.pmed.1003402
101. NYC Department of Health and Mental Hygiene. *Racial Inequities in COVID-19 Hospitalizations During the Omicron Wave in NYC*, p. 17. (2022). Available at: <https://www.nyc.gov/assets/doh/downloads/pdf/covid/black-hospitalizations-omicron-wave.pdf>.
102. The White House. *National HIV/AIDS Strategy for the United States: Updated to 2020*. (2015). Available at: <https://files.hiv.gov/s3fs-public/nhas-update.pdf>.
103. Gardner EM, McLees MP, Steiner JF, Del Rio C, Burman WJ. The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection [research support, N.I.H., extramural review]. *Clin Infect Dis.* (2011) 52:793-800. doi: 10.1093/cid/ciq243
104. Ogedegbe G, Ravenell J, Adhikari S, Butler M, Cook T, Francois F, et al. Assessment of racial/ethnic disparities in hospitalization and mortality in patients with COVID-19 in new York City. *JAMA Netw Open.* (2020) 3:e2026881. doi: 10.1001/jamanetworkopen.2020.26881
105. Circle of Blue. *Infographic: America's Septic Systems*. (2015). Available at: <https://www.circleofblue.org/2015/world/infographic-americas-septic-systems/>.
106. Kirby AE, Walters MS, Jennings WC, Fugitt R, LaCross N, Mattioli M, et al. Using wastewater surveillance data to support the COVID-19 response - United States, 2020-2021. *MMWR Morb Mortal Wkly Rep.* (2021) 70:1242-4. doi: 10.15585/mmwr.mm7036a2
107. Diamond MB, Keshaviah A, Bento AI, Conroy-Ben O, Driver EM, Ensor KB, et al. Wastewater surveillance of pathogens can inform public health responses. *Nat Med.* (2022) 28:1992-5. doi: 10.1038/s41591-022-01940-x
108. Mina MJ, Metcalf CJE, McDermott AB, Douek DC, Farrar J, Grenfell BT. A global immunological observatory to meet a time of pandemics. *eLife.* (2020) 9:9. doi: 10.7554/eLife.58989

109. Jacobs D, McDaniel T, Varsani A, Halden RU, Forrest S, Lee H. Wastewater monitoring raises privacy and ethical considerations. *IEEE Trans Technol Soc.* (2021) 2:116–21. doi: 10.1109/TTS.2021.3073886
110. Savi MK, Yadav A, Zhang W, Vembar N, Schroeder A, Balsari S, et al. A standardised differential privacy framework for epidemiological modelling with mobile phone data. *PLOS Digit Health.* (2023) 2:e0000233. doi: 10.1371/journal.pdig.0000233
111. Rader B, Gertz A, Iuliano AD, Gilmer M, Wronski L, Astley CM, et al. Use of at-home COVID-19 tests - United States, august 23, 2021 -march 12, 2022. *MMWR Morb Mortal Wkly Rep.* (2022) 71:489–94. doi: 10.15585/mmwr.mm7113e1
112. Salomon JA, Reinhart A, Bilinski A, Chua EJ, La Motte-Kerr W, Ronn MM, et al. The US COVID-19 trends and impact survey: continuous real-time measurement of COVID-19 symptoms, risks, protective behaviors, testing, and vaccination. *Proc Natl Acad Sci USA.* (2021) 118:4118. doi: 10.1073/pnas.2111454118
113. Harruff RC, De Silva M, Lubin M. Medical examiner surveillance of the COVID-19 pandemic in King County, Washington, USA. *Am J Forensic Med Pathol.* (2023) 44:157–65. doi: 10.1097/PAF.0000000000000839
114. Mwananyanda L, Gill CJ, MacLeod W, Kwenda G, Pieciak R, Mupila Z, et al. Covid-19 deaths in Africa: prospective systematic postmortem surveillance study. *BMJ.* (2021) 372:n334. doi: 10.1136/bmj.n334
115. Lu FS, Nguyen AT, Link NB, Molina M, Davis JT, Chinazzi M, et al. Estimating the cumulative incidence of COVID-19 in the United States using influenza surveillance, virologic testing, and mortality data: four complementary approaches. *PLoS Comput Biol.* (2021) 17:e1008994. doi: 10.1371/journal.pcbi.1008994