Evaluating Data Types: A Guide for Decision Makers using Data to Understand the Extent and Spread of COVID-19 (2020)

DETAILS

15 pages | 8.5 x 11 | PDF
ISBN 978-0-309-67858-2 | DOI 10.17226/25826

GET THIS BOOK

FIND RELATED TITLES

CONTRIBUTORS

National Academies of Sciences, Engineering, and Medicine

SUGGESTED CITATION

Visit the National Academies Press at NAP.edu and login or register to get:

– Access to free PDF downloads of thousands of scientific reports

– 10% off the price of print titles

– Email or social media notifications of new titles related to your interests

– Special offers and discounts

EVALUATING DATA TYPES
A Guide for Decision Makers using Data to Understand the Extent and Spread of COVID-19

Authors: Adrian E. Raftery, University of Washington*
             Janet Currie, Princeton University*
             Mary T. Bassett, Harvard T.H. Chan School of Public Health**
             Robert Groves, Georgetown University**

*Member of SEAN Executive Committee
**Co-Chair of SEAN Executive Committee and Member of Standing Committee on Emerging Infectious Diseases and 21st Century Health Threats

The National
Academies of

SCIENCES
ENGINEERING
MEDICINE

EXECUTIVE SUMMARY

This rapid expert consultation was produced by the Societal Experts Action Network (SEAN), a project of the National Academies of Sciences, Engineering, and Medicine with support from the National Science Foundation. Its aim is to enable leaders such as you to gain insight into the strengths and weaknesses of the data on the COVID-19 pandemic in your community by applying five criteria to seven types of data available to support decision making. **By understanding these characteristics, you can work with the data type best-suited to the question at hand, and use the data you have to inform your decisions most effectively.**

The seven data types are: the number of confirmed cases, hospitalizations, emergency department visits, reported confirmed COVID-19 deaths, excess deaths, fraction of viral tests that are positive, and representative prevalence surveys (including both viral and antibody tests). The five criteria are: representativeness; bias; uncertainty, and measurement and sampling error; time; and space. The importance of any of these five criteria depends on the nature of the decision being made, and each data type has different strengths and weaknesses.

Each data type represents a piece of the puzzle, and when used in combination, the various types form a clearer picture of how the disease is spreading and its severity. Because any single data type is likely to yield an under- or over- estimate of the extent and spread of the disease, it is important to consider multiple data types and be cautious in relying on estimates without considering sources of bias. The key implications for decision makers are summarized in Box 1 below.

BOX 1
KEY IMPLICATIONS FOR DECISION MAKERS

Confirmed Cases: This measure is readily available, but is likely to be an underestimate of total persons with the disease. As the volume of testing expands, this measure should be more useful.

Hospitalizations: These data are typically available quickly, but reflect only the most severe cases of infection and patients who were exposed to the virus several weeks before admission.

Emergency Department Visits: Data on visits may be available at the local level in close to real time and are most useful in the early stages of an outbreak or to assess resurgence, though it is important to remember that patients with symptoms were exposed up to 2 weeks earlier.

Reported Confirmed COVID-19 Deaths: These data reflect the state of the outbreak several weeks previously because of the long course of infection.

Excess Deaths: Compared with the other data reviewed, excess deaths are the best indicator of the mortality impacts of the pandemic. Excess deaths also reflect the state of the outbreak several weeks previously given the long course of infection.

Fraction of Viral Tests That Are Positive: These data may not be an adequate measure of prevalence, depending on testing criteria. If mainly symptomatic people are tested, these data are expected to overestimate the true community prevalence. The proportion of positive tests is expected to decline as testing expands to include mildly symptomatic and asymptomatic people.

Representative Prevalence Surveys: Data from these surveys represent the best strategy for understanding the prevalence of a disease in any given population (workplace, nursing home, etc.) at a specific point in time.

INTRODUCTION

Fortunately, more information about how COVID-19 is affecting the nation is now available, but as is so often the case, the information comes in various forms and is not always complete. The purpose of this rapid expert consultation is to help decision makers, especially at state and local levels, better understand and evaluate the strengths and limitations of the various data types being used as indicators of the extent and spread of COVID-19 in their communities. This enhanced understanding can lead to more informed decisions on critical issues that depend on those indicators, such as when to lift social distancing restrictions, allow public gatherings, or reopen businesses. Drawing on relevant literature and expert judgment, this rapid expert consultation describes the considerations that apply in using the available data while taking account of their limitations. It reviews in turn:

1. Seven data types used as indicators for evaluating the course of COVID-19 in a community or population
2. Five criteria against which the reliability and validity of these data types can be assessed
3. Cautions to consider in making decisions with imperfect data
4. Specific limitations and cautions that apply to data on COVID-19

This rapid expert consultation addresses the assessment of the seven data types and the implications of those assessments for decision making; it does not recommend specific policy actions.

Specific features of the disease and response to the pandemic have implications for understanding this assessment of data types. According to the Centers for Disease Control and Prevention (CDC) (2020a), the incubation period for COVID-19 is thought to be up to 14 days, with a median time of 4–5 days from exposure to onset of symptoms and with deaths indicating infection from several weeks previously. This long incubation period and progression of infection, as well as the possibility of asymptomatic cases, has implications, discussed below, for interpreting the different data types. Also, determining both the prevalence of COVID-19 and deaths from the disease depends on the availability and accuracy of testing. In the early days of the pandemic, viral tests were rationed, and it was difficult for people to get tested. Viral tests have become more widely available, but are still available mainly to people with symptoms. Antibody tests have also become more widely available, but are of variable quality. The utility of antibody tests depends on the sensitivity and specificity of the assays, and current testing at this point could result in relatively more false-positive and fewer false-negative results.[1] Some demographic groups, such as the elderly, African Americans, Latinos, and Native Americans, have been disproportionately affected by the virus, suggesting that data for these groups may deserve particular attention. Data collection should include relevant information to allow examination of such disparities, which at present is frequently missing.

---

[1]According to the CDC (2020), evidence "suggests that the presence of antibodies may decrease a person's infectiousness and offer some level of protection from reinfection. However, definitive data are lacking, and it remains uncertain whether individuals with antibodies (neutralizing or total) are protected against reinfection with SARS-CoV-2, and if so, what concentration of antibodies is needed to confer protection….pending additional data, the presence of antibodies cannot be equated with an individual's immunity from SARS-CoV-2 infection." Moreover, "the utility of tests depends on the sensitivity and specificity of the assays....In most of the country, including areas that have been heavily impacted, the prevalence of SARS-CoV-2 antibody is expected to be low, ranging from <5% to 25%, so that testing at this point might result in relatively more false-positive results and fewer false-negative results." See https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/antibody-tests-guidelines.html.

1. DATA TYPES USED TO EVALUATE THE COURSE OF COVID-19

The following types of data on the extent and spread of COVID-19, some of which are highly correlated with each other, are being used to inform decision making:

- **Number of confirmed cases** (positives from diagnostic/viral tests) as indicators of total COVID-19 cases
- **Hospitalizations** (and ICU beds occupied) as a measure of strain on the hospital system and the numbers of severe cases
- **Emergency department visits** as a measure of patient-initiated care seeking and numbers of people with similar syndromes, such as influenza-like illnesses, which can be an indicator of clinically important COVID-19–type illness
- **Reported confirmed COVID-19 deaths** as the basis for estimating deaths associated with COVID-19
- **Excess deaths** (all causes) over prior comparable time periods as a measure of the total number of deaths that may be directly or indirectly attributable to COVID-19
- **Fraction of viral tests that are positive** as a measure of the total number of currently infected persons
- **Representative prevalence surveys** (including both viral and antibody tests) administered to a representative sample of a defined population to estimate the percentage of persons in that population either currently or formerly positive for COVID-19

Given the rapid evolution of understanding of the virus that causes COVID-19, additional data types are emerging. For instance, surveillance of wastewater to detect the virus that causes COVID-19 could provide information to communities about the virus's reemergence, and some researchers are using cell phone data to track compliance with social distancing guidelines.

2. CRITERIA FOR ASSESSING THE RELIABILITY AND VALIDITY OF THE DATA TYPES

The utility of data for decision making is affected by many factors, including the burden of collecting, cleaning, and interpreting the data across sources. Also, data collection and models tend to improve over time, so their assessment will also need to be updated regularly.[2] Meanwhile, decision makers must use the data that are available while understanding their limitations. To this end, the following five criteria can be considered:

- **Representativeness:** Does the reporting population represent the population of interest? Does each person in the population have an equal chance of being measured?
- **Bias:** Are there systematic factors that could cause the values reported to be overestimates or underestimates of the actual values? Is there a difference between what is reported and what one wants to measure?
- **Uncertainty, and Measurement and Sampling Error:** Is there uncertainty due to small sample sizes; that is, do small sample sizes cause unstable numbers? Have people been measured twice? Do tests produce accurate results?
- **Time:** What is the time lag in reporting the numbers? Are the numbers consistently updated, or are there time gaps in delivery of the data? Do time lags differ across sources? Has the

---

[2]This document does not specifically review models, but the data types reviewed are typically the inputs to models. Thus, understanding the characteristics of the data inputs can inform understanding of models and similar forecasting tools related to the course of the pandemic.

nature of measurement changed over time in a way that impacts reported estimates? Are events recorded on the day they occurred or the day they were reported?

- **Space:** Do the numbers cover all geographic areas of interest? Are areas of particular interest covered? Do all areas use the same measurement and classification system? Do the indicators count persons outside the given jurisdiction?

Table 1 shows the seven data types listed above against the five criteria for assessing their reliability and validity. Check marks indicate that a data type generally meets a criterion, while the triangles denote the need for caution, meaning that the questions listed above under a criterion should be asked to better understand the quality of the data.

3. MAKING DECISIONS WITH IMPERFECT DATA: CAUTIONS TO CONSIDER

Decisions must be made in critical situations even when there is uncertainty about the best available data. It is important for decision makers to be aware of the strengths and weaknesses of the data they receive. This requires that a decision maker rely on the data available to the extent that they promote better decision making, while being mindful of the following cautions:

- **Small case counts:** Counts based on small numbers of cases tend to be unstable and of limited utility for decision making.
- **Time lag between the occurrence of an indicator and its reporting:** Data tend to become more complete over time, so that counts must generally be revised (e.g., deaths on weekends are often reported on the next working day). A second problem is that data on deaths, for example, reflect infections that occurred some time ago and thus need to be interpreted in that context.
- **Overestimation and underestimation:** Given two indicators, one of which may result from systematic overestimation and the other from systematic underestimation, it is good to use both to guide a decision. For example, the proportion of positive tests in a sample of people with active symptoms will be an overestimate of the true prevalence of disease in the population, while the number of confirmed cases as a proportion of the population will likely be an underestimate.
- **Disproportionate impact:** Because averages can obscure disproportionate impacts, using averages as the basis for decisions may affect some individuals in the relevant population more than others. In view of the disproportionate impact of COVID-19 on some groups, it is important to consider the numbers for specific groups based on age, location, race/ethnicity, socioeconomic status, and other factors.
- **Importance of qualitative data:** Quantitative data may provide a limited picture of a situation. Thus in some cases, qualitative (non-numerical) data can be a valuable supplement. Before such data are used, however, it is important to consider how representative they are, just as one would do with quantitative data.
- **Transparency:** Ensuring the open availability of data improves transparency and accountability. It is important to share data with the public and to develop feedback mechanisms so that communities can be engaged to inform responses to the data.

Table 1: Assessment of Data Types by Criteria for Reliability and Validity

	Representativeness	Bias	Uncertainty, Measurement & Sampling Error	Time	Space
Number of confirmed cases	⚠	⚠	☑	⚠	⚠

Key Implication for Decision Making: This measure is readily available, but is likely to be a substantial underestimate of the prevalence of the disease in a population given that most people with COVID-19 are asymptomatic, and even among those who are symptomatic, not all are tested. As the volume of testing expands to include populations with less severe symptoms and asymptomatic individuals, this measure will be increasingly useful for determining the prevalence of COVID-19.

Hospitalizations	⚠	⚠	☑	⚠	⚠

Key Implication for Decision Making: Data on hospitalizations are typically available quickly at the local level, although the completeness of reporting may vary from day to day. These data reflect only the most severe cases of infection, but changes in the number of hospitalizations likely reflect similar changes in the total number of infections within a community. Note patients requiring hospitalization were exposed several weeks previously.

Emergency department visits	⚠	⚠	☑	⚠	⚠

Key Implication for Decision Making: In some jurisdictions, data on emergency department (ED) visits are available at the local level in close to real time. The reason for the visit can be reported either as a syndrome (e.g., "influenza-like illness") or as a specific diagnosis (e.g., "COVID-19"). These data are most useful in the early stages of an outbreak or to assess resurgence, though it should be noted that patients with symptoms were exposed up to 2 weeks earlier.

Reported deaths	⚠	⚠	⚠	⚠	⚠

Key Implication for Decision Making: Reported COVID-19 deaths are affected by the accuracy of cause-of-death determinations and reflect the state of the outbreak several weeks previously because of the long course of COVID-19 infection. Sometimes lags in reporting of data also occur.

Excess deaths	☑	⚠	☑	☑	☑

Key Implication for Decision Making: Compared with the other data reviewed here, excess deaths are the best indicator of the mortality impacts of the pandemic. However, because of the possibility of death misclassification, these data represent a mix of confirmed COVID-19 deaths and deaths from other causes.

Fraction of viral tests that are positive	⚠	⚠	⚠	⚠	⚠

Key Implication for Decision Making: These data may not be an adequate measure of prevalence, depending on testing criteria. If mainly symptomatic people are tested, this figure is expected to overestimate the true community prevalence. The proportion is expected to decline as testing expands to include mildly symptomatic and asymptomatic people.

Prevalence surveys (representative)	☑	☑	☑	⚠	☑

Key Implication for Decision Making: Representative prevalence surveys are the best strategy for understanding the prevalence of a disease in any given population at a specific point in time. Such surveys can be undertaken for specific populations (e.g., workplace, nursing home, jails and prisons). Although they require undertaking a special study rather than using routinely collected data, many public health agencies have this capacity. There will be some time lag involved, however, in mounting and interpreting such a survey.

☑   Data source usually meets this criterion.

⚠   Data source may or may not meet the criterion, and questions related to that criterion should be asked.

4. SPECIFIC LIMITATIONS AND CAUTIONS REGARDING COVID-19 DATA

This section applies the five criteria described in section 2 to the seven data types commonly used to make COVID-19 policy decisions as outlined in section 1. Decision makers should use the data available to them, as they represent some of the best indicators currently available, while being explicit about their limitations and highlighting questions that should be asked of those providing the data.

Number of Confirmed Cases

Implications for decision making: This measure is readily available, but is likely to be a substantial underestimate of the prevalence of the disease in a population given that most people with COVID-19 are asymptomatic, and even among those who are symptomatic, not all are tested. As the volume of testing expands to include populations with less severe symptoms and asymptomatic individuals, this measure will be increasingly useful for determining the prevalence of COVID-19.

- **Representativeness:** The number of confirmed cases per 1,000 people per week, month, or year (i.e., the rate of confirmed cases) is not representative of actual prevalence in the population because of limited testing capacity and the widespread lack of testing of asymptomatic individuals. The reported number is likely to be a substantial underestimate of infected persons by a factor of as much as 10 or more, although this factor is likely to decline over time as testing becomes more widespread (Bedford et al., 2020; Johndrow et al., 2020).

- **Bias:** The number of confirmed cases is an underestimate because in addition to the limitations of testing noted above, many people lack access to testing and are less likely to seek it out, and only those with sufficiently severe symptoms are tested. This may be a particular problem for those who lack health insurance, who live in underresourced relative to more affluent areas, or who may avoid seeking testing because of fear (e.g., undocumented immigrants) (Borjas, 2020). An additional problem arises if testing is more intensive in virus hot spots. Well-done surveys of representative samples of people can help in understanding the magnitude of this problem. Contact tracing and testing may also be helpful in identifying previously unconfirmed cases.

- **Uncertainty, and measurement and sampling error:** Sampling error due to small numbers of cases is likely to be a much smaller problem than bias. If multiple positive tests are reported for the same person over time, the number of positive tests divided by the base population could possibly produce an overestimate of the actual number of cases.

- **Time:** Confirmed cases are usually reported daily, but these reports contain consistent errors, such as underreporting on weekends, and it may take several days for test results to be confirmed. The underestimation of prevalence is likely to be consistent over short time periods, and so the trend in confirmed cases can be a good indicator of short-term trends in prevalence. However, there are problems of comparability over longer time scales, because the extent of underestimation tends to decline over time as more people are tested.

- **Space:** Confirmed cases tend to be reported by hospitals or other official testing sites and thus should be available at a fine-grained geographic scale. However, the bias in confirmed cases and rates of testing may differ among different areas, so these data may not be comparable across space. In addition, some localities include suspected cases and some do not. Accounting for what is being reported, different testing strategies or types of tests, and cases in which people seek treatment in counties where they do not reside may help explain some of these differences, thus making comparisons more useful.

Hospitalizations

Implications for decision making: Data on hospitalizations are typically available quickly at the local level, although the completeness of reporting may vary from day to day. These data reflect only the most severe cases of infection, but short-term changes in the number of hospitalizations

likely reflect similar changes in the total number of infections within a community. Note that patients requiring hospitalization were exposed several weeks previously.

- **Representativeness:** Data on current hospitalizations tend to be relatively complete, and thus representative of the *hospitalized population* of COVID-19 patients given that people hospitalized with COVID-19 symptoms are much more likely to be tested than the general public. Eventually, all hospital discharges are reported to state authorities. Of course, there are still disparities in access to care, even hospitalization, based on race/ethnicity, nativity, and socioeconomic status (Azar et al., 2020).

- **Bias:** Underestimates may arise from the misdiagnosis of patients with COVID-19, which occurred early in the epidemic and may still be occurring, but less often.[3] For example, some patients diagnosed with pneumonia may actually have COVID-19. However, diagnosis is improving over time.

- **Uncertainty, and measurement and sampling error:** There is relatively little sampling error in this measure, apart from the misdiagnosis bias just mentioned.

- **Time:** As noted, hospitalizations are usually reported in a relatively timely manner, although they may not be reported every day, and aggregation of data from individual hospitals can be unsystematic. The data are likely to be consistent over the relatively short periods of time that are important for decision making. However, they may not be consistent over longer time periods because of the previously discussed potential for misdiagnosis of patients early on in the pandemic. Changes may also have occurred over time in the severity of patients being admitted and lengths of stay, which may depend, for instance, on how crowded a hospital is.

- **Space:** There may be differences across space in the severity of patients being admitted to hospitals. Moreover, the way illnesses are diagnosed and coded varies from hospital to hospital and across cities and states and over time. In areas where hospitals have reached capacity, it is important to track transfers to other hospitals, especially from areas with limited facilities.

Emergency Department Visits

Implications for decision making: In some jurisdictions, data on emergency department (ED) visits are available at the local level in close to real time. The reason for the visit can be reported either as a syndrome (e.g., "influenza-like illness") or as a specific diagnosis (e.g., "COVID-19") (Henning, 2004). These data are most useful in the early stages of an outbreak or to assess resurgence, though it should be noted that patients with symptoms were exposed up to 2 weeks earlier.

- **Representativeness:** Most EDs report visit data, and nearly 75 percent of ED visits nationally are captured by the National Syndromic Surveillance System (Hartnett et al., 2020). These data reflect those who use EDs for their health care needs. Depending on local availability of and barriers to accessing both primary and ED care, specific groups may be either under- or overrepresented in ED visit data. Consideration should be given to data representativeness by such characteristics as race/ethnicity, income, and nativity. However, those with the most severe disease may seek care in the ED regardless of these considerations.

- **Bias:** Diagnoses made in the ED may be modified subsequently and may underestimate or overestimate actual COVID-19 cases, especially given time lags in processing of tests. Diagnoses of syndromes, such as "influenza like illness," are based on International Classification of Diseases (ICD) coding, which may be incomplete at any given time or may be driven by considerations of reimbursement or other nonclinical factors.

---

[3]It is of course possible that overestimation could occur, as would be the case if someone who is positive for COVID-19 has been hospitalized for a different reason (e.g., heart attack). Multiple factors contribute to a person's state of health, and there may be some differentiation in how hospitals classify the reason for hospitalization. That said, the potential for such overestimation is less concerning than the underestimation described above in terms of assessing public health risks.

- **Uncertainty, and measurement and sampling error:** There is relatively little sampling error in this measure as it is a fairly complete count of visits in most jurisdictions, apart from the problem of provisional diagnoses noted above. In some cases, data on ED visits may be combined with data on hospitalizations to give a count of people using a facility in a particular period. In this case, care should be taken to account for possible double counting of patients who appeared in the ED and were hospitalized later. Given uncertainty about diagnosis and the possibility that many people with such syndromes as "influenza-like illness" do not have COVID-19, it may be useful to compare these data with data from prior years.
- **Time:** As with hospitalizations, a key advantage of ED visit data is their availability and relative timeliness at the local level, though it will take longer for the data to be transmitted to national databases. Such syndromes as "influenza-like illness" are not specific diagnoses, and all ED diagnoses are provisional. Users of data on time trends should consider that there may be improvement in diagnosis over time.
- **Space:** As in the case of hospitalizations, using data on ED visits to draw inferences about differences in COVID-19 prevalence across places is unlikely to be possible. The extent to which the local population utilizes the ED is likely to vary with population characteristics, as discussed above. Moreover, the way illnesses are diagnosed and coded varies from hospital to hospital and across cities and states, as well as over time.

Reported Deaths

Implications for decision making: Reported COVID-19 deaths are affected by the accuracy of cause-of-death determinations and reflect the state of the outbreak several weeks ago because of the long course of COVID-19 infection. Sometimes lags in reporting of data also occur.

- **Representativeness:** Reported deaths from COVID-19 are likely to be an underestimate because of underdiagnosis, as well as variations in testing across locations. The underestimate may be substantial, and will depend in part on whether "probable" deaths as well as deaths "confirmed" via a test are included. For example, Washington State officials estimate that the number of actual COVID-19 deaths in that state may have been three times greater than the reported number because of the lack of testing early in the epidemic (Bellisle, 2020). Moreover, patients who died from another underlying condition (e.g., heart failure) may be misclassified as COVID-19 deaths, biasing estimates upward.
- **Bias:** As noted, reported deaths from COVID-19 are likely to be underestimates, although some positive bias in the case of patients who were already severely ill is possible (see above).
- **Uncertainty, and measurement and sampling error:** There is little sampling error in this indicator. Measurement issues arise from misdiagnosis or uncertainty about true causes of death. Also, race/ethnicity may be misreported or incomplete on death certificates, especially for American Indian/Alaska Native populations (Arias et al., 2016), leading to errors in calculated death rates by race and ethnicity.
- **Time:** Local health authorities initially report deaths quickly, but the final, complete, cleaned data may take time to produce. The quality of diagnosis may be improving over time, leading to inconsistency, particularly between the early and later periods of the pandemic.
- **Space:** All jurisdictions report deaths, but if misdiagnosis varies among areas, then comparability across areas could be compromised. Note that a similar issue occurs with respect to hospitalizations, so the same caveat applies to both sources of data.

Excess Deaths

Implications for decision making: Compared with the other data reviewed here, excess deaths are the best indicator of the mortality impacts of the pandemic. However, because of the possibility of death misclassification noted above, these data represent a mix of confirmed COVID-19 deaths and deaths from other causes. For example, from March 11-May 2, 2020, New York City reported

13,831 confirmed and 5,048 probable COIVD-19 deaths. Additionally, there were a further 5,293 excess deaths that might have been directly or indirectly attributable to the pandemic (Centers for Disease Control and Prevention, 2020b). The percentages of these deaths that occurred in persons infected with COVID-19 or that resulted from indirect impacts of the pandemic are unknown and require further investigation.

- **Representativeness:** Since all deaths are counted, and each decedent's age, gender, residence, and race/ethnicity are known, these are likely the most representative data available other than those from representative prevalence surveys. As noted above, however, race/ethnicity may be misreported or incomplete on death certificates, leading to errors in death rates by race and ethnicity (Arias et al., 2016).

- **Bias:** The main potential source of bias is selection of a comparison period that itself is unrepresentative. This source of bias can be mitigated by using an average of the past several years as a comparison. Also, it is important to keep in mind that the number of excess deaths will be affected by COVID-19 deaths; deaths from other causes that may have been exacerbated by the response to the pandemic (e.g., when patients delayed seeking care because of concerns about contracting COVID-19 at the hospital, or suicides or domestic violence deaths associated with lockdowns increased); and deaths (e.g., due to traffic accidents) that were prevented by lockdowns. A second type of bias that should be considered is related to underlying changes in population composition, due, for instance, to migration or changes in the age structure. Changes in population composition over time, especially at the county level, may affect trends in excess deaths. These underlying trends should be considered when assessing and adjusting data on excess deaths.

- **Uncertainty, and measurement and sampling error:** There will always be some uncertainty about excess deaths. While the total number of deaths is reasonably accurate, it is difficult to calculate "excess deaths" because deaths in each year reflect unique public health phenomena. As a result, computing excess deaths is a statistical procedure that entails comparing current deaths with expected deaths based on historical averages, and the magnitude of the excess will depend on the time period chosen for comparison.

- **Time:** Given data on the number of deaths, excess deaths can be computed in a timely way relative to past numbers of deaths in the same week or month of the year in most local jurisdictions. It takes longer for deaths to be transmitted to states and the federal government. Data on excess deaths are likely to be the most complete data available because they do not depend on accuracy of diagnosis of COVID-19. Hence, these data will also be the most comparable over time. As with reported deaths, however, the final, complete data on deaths, and therefore on excess deaths, may take time to produce.

- **Space:** These data will be complete and comparable across space.

Fraction of Viral Tests That Are Positive

Implications for decision making: These data may not be an adequate measure of prevalence, depending on testing criteria. If mainly symptomatic people are tested, these data are expected to overestimate the true community prevalence. The proportion is expected to decline as testing expands to include mildly symptomatic and asymptomatic people. *Note: Understanding of the accuracy of antibody testing and its utility as an indicator of immunity is still evolving (reference footnote 1 in the introduction).*

- **Representativeness:** The extent to which these data meet this criterion will depend on the extent to which the people tested are representative of the population. Currently, many tests are administered to people who are referred by their doctors or who feel that they may have COVID-19, and the testing may occur in either public or private facilities. The positivity rate among people with symptoms is likely to be biased upward because this population is unrepresentative, consisting of people who are, on average, more likely to have the disease

relative to those without symptoms. If the people tested are not representative of the population, then as the volume of testing increases, the percentage of positive viral tests can be expected to decline even if the true prevalence of the disease in the population remains constant. Thus the fraction of tests that are positive reflects both the prevalence of the disease and the extent of testing and is not separately a reliable measure of either. However, the fraction of positive tests will decline as testing expands to people who are mildly ill or asymptomatic.[4]

- **Bias:** Tests are usually imperfect, and treating them as if they are perfect can lead to large biases. In general, the observed prevalence, or positivity rate, can be misleading unless the test employed is of high quality. The quality of tests is measured by two numbers: sensitivity, or the proportion of people who have the disease and test positive for it (the true positive rate); and specificity, or the proportion of healthy people that test negative (the true negative rate). With a perfect test, both of these numbers would be 100 percent.[5]

- **Uncertainty, and measurement and sampling error:** Both false positives and false negatives can occur. The quality of tests with respect to this criterion is measured by the sensitivity and specificity of the test. When the prevalence of COVID-19 is low, the likelihood that a positive test predicts disease will decline.

- **Time:** Comparability across time may be poor unless (1) the sample of people tested is representative of the population, and (2) the test results are adjusted for sensitivity and specificity, as described above.

- **Space:** Comparability across space may also be poor unless the sample is representative and the test results are adjusted. Indeed, it may be worse than comparability across time because the differences in sensitivity and specificity across different testing sites may be considerable.

Representative Prevalence Surveys

For these surveys, a representative sample of people to be tested is selected. The World Health Organization (WHO) has produced a protocol for such surveys for COVID-19 (World Health Organization, 2020). Dean (2020) outlines the advantages and challenges of such surveys, as well as ways to make the most of them. Such surveys can be conducted at the local, state, or national level. Oregon, Indiana, and Ohio have initiated such efforts. Similar surveys are often carried out for social science or market research as well as epidemiological purposes, so the methodology is well established. Several such surveys have been conducted to capture the prevalence of COVID-19, including the COVID-19 Impact Survey that is administering symptom checkers to known representative samples in 18 subnational areas (Wozniak et al., 2020; Vogel, 2020; Joseph and Branswell, 2020).

> *Implications for decision making:* Representative prevalence surveys are the best strategy for understanding the prevalence of a disease in any given population at a specific point in time. Such surveys can be undertaken for specific populations (e.g., workplace, nursing home, jails and prisons). Although they require undertaking a special study rather than using routinely collected data, many public health agencies have this capacity. There will be some time lag involved, however, in mounting and interpreting such a survey. While prevalence surveys in general, such as surveys of health care workers or convenience samples (defined in footnote 7 below) of grocery shoppers, may be useful if replicated over time to measure trends, they are not necessarily representative.

- **Representativeness:** These surveys are representative by design, thus avoiding the lack of representativeness that characterizes many current prevalence estimates based on tests.

---

[4]See also http://freerangestats.info/blog/2020/05/09/covid-population-incidence.

[5]If the test is not perfect, the observed prevalence can be adjusted as follows: Adjusted prevalence = (Observed prevalence + Specificity – 1)/(Sensitivity + Specificity – 1) (Rogan and Gladen, 1978).

Representativeness is usually ensured by taking a random sample, whereas prevalence surveys using convenience samples[7] will generally not be representative.[8] The same problem with lack of representativeness is true of samples consisting of volunteers.

- **Bias:** If prevalence surveys are based on representative samples and if the sensitivity and specificity of the viral tests are known, bias due to errors in the tests can be corrected using well-known statistical formulas. It is important to make these corrections so that unbiased estimates can be obtained; see footnote 1 (Biemer and Lyberg, 2008).
- **Uncertainty, and measurement and sampling error:** If a survey is representative, these data deficiencies can be quantified. However, sampling error can be substantial in small random samples. Also, uncertainty in a prevalence survey will not be well quantified if the survey is not designed to be representative, as when, for example, convenience samples or volunteers are used.
- **Time:** This criterion depends on how quickly the results of a survey can be produced and how often the survey is carried out. If it is not carried out frequently, the results may still be useful to adjust for biases in other data types. For example, data from the American Community Survey are often used to see how representative a given sample may be in terms of the distribution of such demographic characteristics as age, sex, and race.
- **Space:** Spatial completeness is good as long as a representative sample is used, but this may not be the case with convenience samples or volunteer subjects.

CONCLUSION

The COVID-19 pandemic is a reminder, once again, of the importance of evidence and a robust public health data infrastructure. Decision making related to the pandemic requires the use of data often not designed for the task at hand. With greater understanding of the strengths and limitations of these data, decision makers can make better decisions. Continued investment in public health and its data surveillance structures is needed to meet the nation's current and future public health challenges.

SEAN is interested in your feedback. Was this rapid expert consultation useful? Send comments to sean@nas.edu or (202) 334-3440.

---

[7]Convenience samples are constructed from a group of people that are easy to contact or reach, and are not random.

[8]For prevalence surveys, it is sometimes possible to use proxies for representative samples. For instance, in March and April 2020, seroprevalence surveys of health care workers in many major medical centers did not do a bad job of anticipating the local area prevalence. Health care workers are at higher occupational risk, but they are also more affluent than average, so those biases cancelled each other out somewhat. Another example might be a large heterogenous employer in a city that had all its employees tested; this might not be a bad proxy for a truly representative sample for that city.

REFERENCES

Arias, E., Heron, M., and Hakes, J.K. (2016). The validity of race and Hispanic-origin reporting on death certificates in the United States: An update. *Vital and Health Statistics* 2(172), 1-21. Available: https://www.researchgate.net/publication/306079754_The_Validity_of_Race_and_Hispanic-origin_Reporting_on_Death_Certificates_in_the_United_States_An_Update.

Azar, K., Shen, Z., Romanelli, R., Lockhart, S., Smits, K., Robinson, S., Brown, S., and Pressman, A. (2020). Disparities in outcomes among COVID-19 patients in a large health care system in California. *Health Affairs*, 39(7), 1-8.

Bedford, T., Greninger, A.L., Roychoudhury, P., Starita, L.M., Famulare, M., et al. (2020). *Cryptic transmission of SARS-CoV2 in Washington State.* MedRxiv Preprint. Available: https://www.medrxiv.org/content/medrxiv/early/2020/04/16/2020.04.02.20051417.full.pdf.

Bellisle, M. (2020). Washington State's actual coronavirus death toll may be higher than current tallies, health officials say. *Seattle Times*. May 21. Available: https://www.seattletimes.com/seattle-news/health/washington-states-actual-coronavirus-death-toll-may-be-higher-than-current-tallies-health-officials-say.

Biemer, P., and Lyberg, L. (2008). *Introduction to Survey Quality*. New York: Wiley Interscience.

Borjas, G.J. (2020). Demographic determinants of testing incidence and COVID-19 infection in New York City neighborhoods. *NBER Working Paper 26952.* April. Available: https://www.nber.org/papers/w26952.pdf.

Centers for Disease Control and Prevention. (2020a). *Interim clinical guidelines for management of patients with confirmed coronavirus disease (COVID-19).* Available: https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-guidance-management-patients.html.

_____. (2020b). Preliminary estimate of excess mortality during the COVID-19 outbreak—New York City, March 11–May 2, 2020. *Morbidity and Mortality Weekly Report,* 69(19), 603-605.

Dean, N.E. (2020). COVID-19 data dives: The takeaways from seroprevalence surveys. *Medscape*. May 4. Available: https://www.medscape.com/viewarticle/929861.

Henning, K.J. (2004). What is syndromic surveillance? Syndromic Surveillance: Reports from a National Conference, 2003. *Morbidity and Mortality Weekly Report*, 53 (Suppl):7-11.

Hartnett, K.P., Kite-Powell, A., DeVies, J., Coletta, M.A., Boehmer, T.K., Adjemian, J., Gundlapalli, A.V. (2020). Impact of the COVID-19 pandemic on emergency department visits—United States, January 1, 2019–May 30, 2020. *Morbidity and Mortality Weekly Report*, June 3, 2020, early release, 69.

Johndrow, J., Lum, K., Gargiulo, M. and Ball, P. (2020). Estimating the number of SARS-CoV-2 infections and the impact of social distancing in the United States. *arXiv Preprint*. Available: https://arxiv.org/pdf/2004.02605v2.pdf.

Joseph, A., and Branswell, H. (2020). The results of coronavirus 'serosurveys' are starting to be released. Here's how to kick their tires. *STAT,* April 24. Available: https://www.statnews.com/2020/04/24/the-results-of-coronavirus-serosurveys-are-starting-to-be-released-heres-how-to-kick-their-tires.

Rogan, W.J., and Gladen, B. (1978). Estimating prevalence from the results of a screening test. *American Journal of Epidemiology*, 107(1), 71-76. Available: https://pubmed.ncbi.nlm.nih.gov/623091.

Vogel, G. (2020). Antibody surveys suggesting vast undercount of coronavirus infections may be unreliable. *Science,* April 21. Available: https://www.sciencemag.org/news/2020/04/antibody-surveys-suggesting-vast-undercount-coronavirus-infections-may-be-unreliable.

World Health Organization. (2020). *Population-based age-stratified seroepidemiological investigation protocol for COVID-19 virus infection*. Available: https://apps.who.int/iris/handle/10665/331656.

Wozniak, A., Willey, J., Benz, J., and Hart, N. (2020) *COVID Impact Survey.* Chicago, IL: National Opinion Research Center.

ACKNOWLEDGMENTS