



# Perspective

## Failing the Test — The Tragic Data Gap Undermining the U.S. Pandemic Response

Eric C. Schneider, M.D.

**A**s the United States navigates one of the most serious pandemics in history, much of the country has been shut down to prevent devastating local outbreaks that threaten lives and can

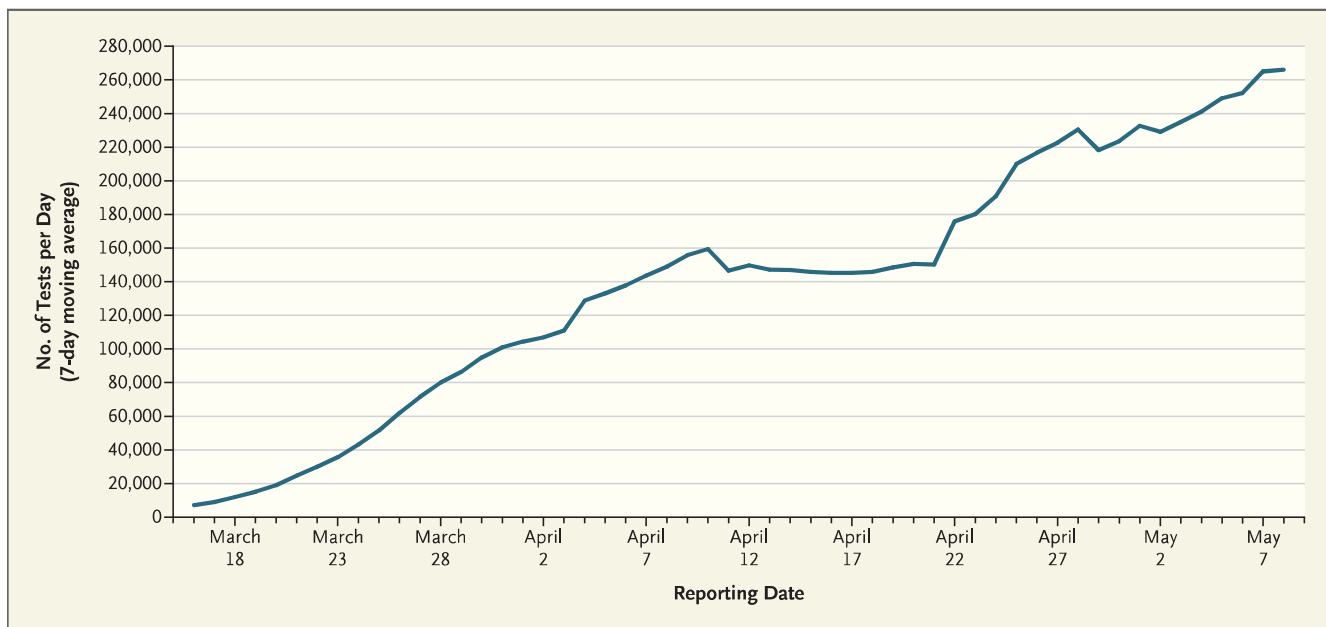
overwhelm hospitals. A breakdown in the federal disaster response delayed state and local responses, allowing SARS-CoV-2 to spread rapidly in New York, New Jersey, Michigan, Louisiana, and other states. Only astute early interventions in Seattle and the San Francisco Bay Area seem to have stemmed a potential tide of cases and deaths. Covid-19 has taken more American lives in 1 month than the Vietnam War claimed over 8 years. Other countries, such as Australia, South Korea, Germany, Singapore, and Taiwan, managed to contain the virus early and are working hard to keep it suppressed as they reopen their economies.

Tragically, the United States, unable to match other countries'

response, has tallied the most cases and deaths in the world — and recent data suggest that those tallies are underestimates. Why has the U.S. response been so ineffectual? One key answer is testing, which has been a cornerstone of Covid-19 control elsewhere. U.S. testing to identify people infected with SARS-CoV-2 has been slow to start and to this day has not sufficiently ramped up. Testing was delayed in January and February as the Centers for Disease Control and Prevention (CDC) distributed faulty test kits, then failed to approve a working test developed by the World Health Organization or those developed by local public health laboratories. Since March, the number of tests per day has never reached

the number needed because of shortages of reagents, materials, and personal protective equipment (see graph).

Having failed to test early enough to contain outbreaks, the country has fallen back on two mitigation strategies: accelerating drug and vaccine development and an unprecedented strategy of nonpharmacologic interventions (NPIs) involving draconian school and business closures, stay-at-home orders, and physical distancing. Drugs and vaccines are extremely unlikely to alter the early course of the pandemic. In the short term, only NPIs have slowed the spread of disease. Yet NPIs carry a heavy economic price as well as their own health burdens, as people fail to receive care for other conditions or suffer mental health consequences from isolation, unemployment, and sudden poverty. Whether NPIs are maintained or not, serious health consequences appear inescapable.



Daily SARS-CoV-2 Tests Performed in the United States from March 16, 2020, through May 8, 2020.

Data are from the COVID Tracking Project (<https://covidtracking.com/api>).

Without testing, the response will continue to fall short. Shortages of test materials have forced a narrow local testing strategy dedicated to managing the care of hospitalized patients and preventing health care workers from transmitting Covid-19. As state government officials and business leaders search for an exit from NPIs and study the success of other countries, they are realizing that testing, contact tracing, and isolation of people who test positive will be essential to successfully reopening economies. The most recent congressional rescue package featured \$25 billion for testing.

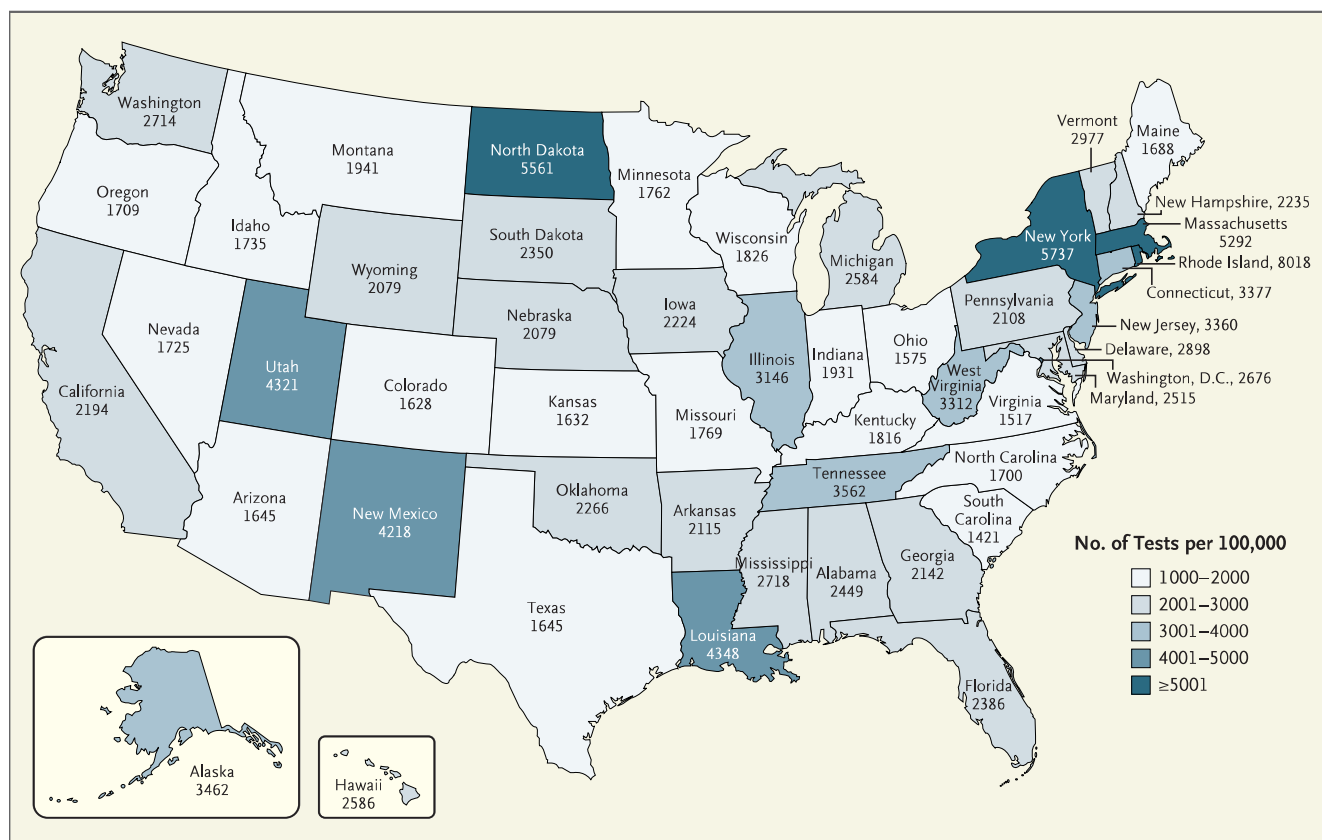
To date, efforts to bolster testing have focused on operational issues: whether testing capacity is adequate, why shortages and supply chain failures are so pervasive, and how to scale up testing to the massive numbers needed to mitigate the U.S. epidemic. Yet offering more tests is not a strategy in and of itself. If enough

tests were available, we would still need to answer a fundamental question: What decisions are the results meant to inform? Testing has many purposes beyond diagnosis and protection of health care workers. Testing data are needed to manage all aspects of a pandemic. For instance, they are a cornerstone of epidemic forecasting models, which are sorely needed to reveal the future demand for care, including the timing of case surges and the magnitude of required emergency medical services, hospital staff, hospital beds, ventilator equipment, and mortuary services. Without good testing data, forecasters have to rely on guesswork and assumptions.

During this pandemic, model forecasts have ranged from tens of thousands to more than 2 million deaths during the initial months of the U.S. epidemic.<sup>1,2</sup> This variation is not surprising. Modeling is difficult, and a paucity of the facts required to inform models is problematic. Precise

facts about the virus, its transmissibility, clinical course, and lethality are only just beginning to emerge. Few facts are known about the effectiveness of physical distancing and other NPIs, which depend on unpredictable human behavior. Modelers make up for missing facts by including assumptions. Critiques of the models have centered on the assumptions used and their influence on results: as the refrain goes, models are only as good as their assumptions.

The uncertainty of models is beginning to collide with a frustrated, desperate public staring at the catastrophic economic effects of the NPI response: unemployment, loss of income, and closures of businesses that may never return. These harsh facts are increasingly overriding public concern about death-toll forecasts from models with limited correspondence to unfolding reality. People are increasingly calling for ending stay-at-home orders, re-



**Cumulative Rate of SARS-CoV-2 Tests by State (March 16, 2020, through May 8, 2020).**

Data are from the COVID Tracking Project (<https://covidtracking.com/api>) and are based on Census data (estimates of the total resident population and resident population age 18 years or older for the United States and Puerto Rico: July 1, 2019 [SCPRC-EST2019-18+POP-RES], U.S. Census Bureau, Population Division, release date: December 2019).

opening businesses, and returning to life as it was before.

In the rush to discard uncertain forecasts and err on the side of hope, surprisingly little energy has been dedicated to an important fix: replacing models' assumptions with verifiable facts. For drug and vaccine development, solid data play a central role. Well-established protocols determine at every step the data needed, the methods for collecting them, and their analysis. Dozens of sophisticated metrics are recorded and analyzed regarding the safety and efficacy of candidate products and their potential benefits; assumptions and hope play little role in the process.

Yet when it comes to evaluat-

ing NPI use in this pandemic, we seem unable or unwilling to muster the testing data that could inform statistical models and guide our actions. Forecasters and planners desperately need timely testing data. Yet as the absence of comprehensive public data on race and ethnicity revealed, the United States has underfunded and undermined its disease surveillance programs and done a poor job of organizing its 50 state systems for collecting and reporting testing data. The pandemic affects all states, yet states' data are incomplete and uneven at best (see map). The shortcomings are even more puzzling in the light of two decades of bipartisan federal efforts to build measurement and

public reporting systems for health care and implement electronic health records.

The best database on testing for Covid-19 in the United States, created through valiant efforts by news media organizations to fill the gap left by the CDC, contains testing data limited to aggregated counts of the tests done each day, the states where tests were performed, and the number of positive results.<sup>3</sup> The validity and reliability of the data are not fully known. Inspection of the data suggests a patchwork of inconsistent reporting from state and commercial labs. The database lacks basic information about tests such as the characteristics of the people tested, where they were

tested, how they were selected for testing, and what factors led to the decision to test them. Yet these data are the best we have.

That the United States is failing such a simple test of its capacity to protect public health is shocking. Collecting and reporting public health data are not rocket science. Other countries, notably Canada and Belgium, are already reporting nationwide data on testing at the individual level, including individual demographic data (using ranges for each person to protect privacy) and other key attributes for each test.<sup>4</sup> The United States was once a leader in collecting systematic federal data on population health. Now our national disease-tracking effort seems stuck with well-meaning but scattershot efforts by tech companies using cellular phone signals, social media surveys, online searches, and smart thermometers as we try to guess where Covid-19 outbreaks may be lurking. Small one-off studies using convenience samples have popped up to try to fill the vacuum with

basics such as percentages of cases that are asymptomatic and of symptomatic people who seek care. Because of sampling bias, these studies are producing wildly different and nearly uninterpretable results. Estimates are so wide ranging that modelers have little choice but to default back to imprecise assumptions.

In the information age, the United States seems to be swimming in big data. This country has generated many of the world's largest, most innovative, most profitable data companies. Yet when it comes to forecasting the spread of a major pandemic that is killing Americans and wreaking havoc on our economy, we seem oddly lost. With more than 80,000 dead and no end in sight, our national efforts seem feeble and more halting than the 19th-century work of Florence Nightingale in the Crimean War and William Farr in England, where they used systematically collected epidemiologic data and rigorous analysis to save countless lives. Would that our statistical models

had such standardized, systematically collected, and readily reported data to inform them. Reopening state economies without the precision provided by analysis of rigorously reported testing data seems a peculiarly American form of madness.

Disclosure forms provided by the author are available at NEJM.org.

From the Commonwealth Fund, New York.

This article was published on May 15, 2020, at NEJM.org.

1. Ferguson NM, Laydon D, Nedjati-Gilani G, et al. Report 9: impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. London: Imperial College COVID-19 Response Team, March 16, 2020 (<https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf>).
2. Reich Lab COVID-19 Forecast Hub. COVID-19 forecasts: cumulative deaths. May 1, 2020 (<https://reichlab.io/covid19-forecast-hub/>).
3. The Covid Tracking Project home page (<https://covidtracking.com/api>).
4. Berry I, Soucy JR, Tuite A, Fisman D. Open access epidemiologic data and an interactive dashboard to monitor the COVID-19 outbreak in Canada. *CMAJ* 2020;192(15):E420.

DOI: 10.1056/NEJMp2014836

Copyright © 2020 Massachusetts Medical Society.