



TESTIMONY BEFORE THE UNITED STATES CONGRESS

House Science, Space, & Technology Committee
Subcommittee on Investigations & Oversight

DATA FOR DECISION-MAKING

Responsible Management of Data
During COVID-19 and Beyond

AVIK S. A. ROY

President, The Foundation for Research on Equal Opportunity

September 23, 2020

The Foundation for Research on Equal Opportunity (FREOPP) is a non-partisan, non-profit, 501(c)(3) organization dedicated to expanding economic opportunity to those who least have it. FREOPP does not take institutional positions on any issues. The views expressed in this testimony are solely those of the author.

INTRODUCTION

The performance of U.S. public health infrastructure during the COVID-19 pandemic has been, in large part, a problem of data. Gaps in the collection, reporting, and analysis of data have driven many of the critical policymaking challenges faced at the federal, state, and local levels.

- ***Long-term care facilities.*** 42% of all U.S. deaths from COVID-19 have taken place in long-term care facilities, such as nursing homes and assisted living facilities, that house 0.6% of the U.S. population.¹ However, as recently as May, 11 states were not reporting long-term care deaths as a distinct category.² Reporting of this data only improved after the Centers for Medicare & Medicaid Services began requiring that nursing homes report such deaths directly to CMS.
- ***Misleading data from coronavirus PCR testing.*** The most reliable method of detecting an active SARS-CoV-2 infection (i.e., a “case”) is through amplification of coronaviral RNA in a sample taken from a nasal or throat swab using the *reverse transcriptase polymerase chain reaction* method (RT-PCR). There is considerable evidence that misuse and/or misinterpretation of RT-PCR test results has led policymakers to misapprehend COVID-19 case data and thereby to enact overly aggressive economic restrictions.³
- ***Delivering potentially life-saving treatments to critically ill patients.*** In May, Remdesivir was approved by the FDA for emergency use after clinical trials showed that the drug reduced mortality in hospitalized COVID-19 patients. Antiquated data systems at the CDC had made it difficult to route supplies of the drug to regions facing spikes in hospitalized patients, limiting patients’ ability to benefit from this new treatment. Earlier in the pandemic, similar problems had made it difficult to route supplies of personal protective equipment and ventilators to facilities in need.

Congress has been trying to upgrade America’s public health surveillance infrastructure for more than a decade, with limited results. Fortunately, we are finally seeing progress on multiple fronts, and more progress will be possible as the pandemic subsides.

LTC FACILITIES: 42% OF COVID-19 DEATHS, BUT 0.6% OF THE POPULATION

The most underappreciated aspect of the novel coronavirus pandemic is its effect on a specific population of Americans: those living in nursing homes and assisted living facilities.

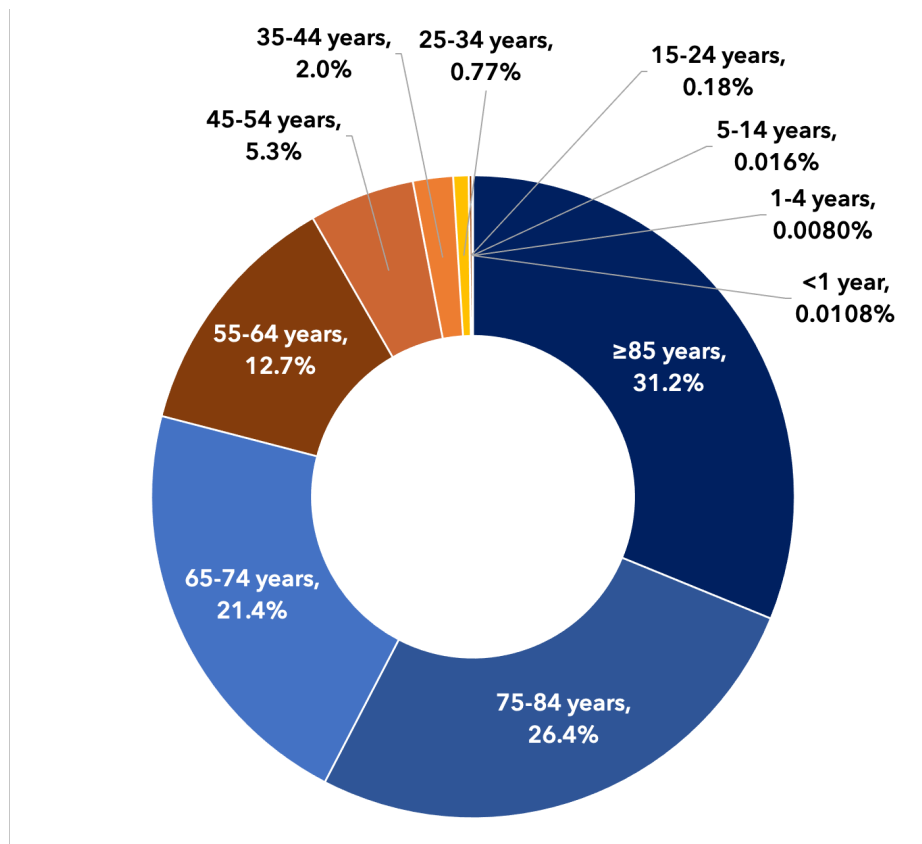
¹ G. Girvan and A. Roy, Nursing Homes & Assisted Living Facilities Account for 42% of COVID-19 Deaths. The Foundation for Research on Equal Opportunity. 2020 May 7: https://freopp.org/the-covid-19-nursing-home-crisis-by-the-numbers-3a47433c3f70?source=collection_home---1-----0-----; accessed September 21, 2020.

² A. Roy, The Most Important Coronavirus Statistic: 42% of U.S. Deaths Are From 0.6% Of The Population. *Forbes*. 2020 May 26: <https://www.forbes.com/sites/theapothecary/2020/05/26/nursing-homes-assisted-living-facilities-0-6-of-the-u-s-population-43-of-u-s-covid-19-deaths/#232a01f074cd>; accessed September 21, 2020.

³ A. Mandavilli, Your Coronavirus Test Is Positive. Maybe It Shouldn’t Be. *The New York Times*. 2020 Aug 29: <https://www.nytimes.com/2020/08/29/health/coronavirus-testing.html>; accessed September 21, 2020.

The disease caused by SARS-CoV-2 affects the elderly far more severely, on average, than younger individuals. 79% of U.S. deaths from COVID-19 have occurred in individuals aged 65 and over, whereas 1% have occurred among those under 35. Among those who are elderly, deaths are concentrated even further among those living in long term care facilities.

Figure 1. Share of COVID-19 Fatalities by Age Bracket



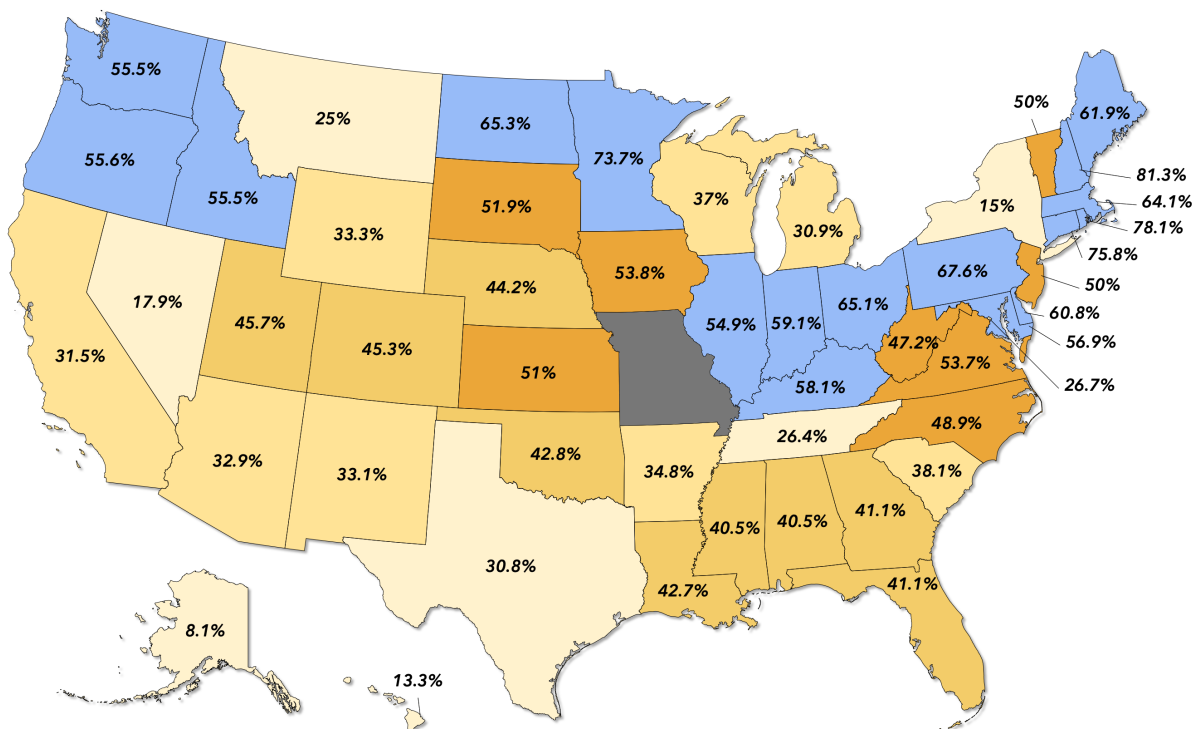
79% of U.S. COVID-19 deaths have occurred in people over 65. Those under 35 years of age represent 1 percent of deaths. (Sources: CDC, FREOPP analysis)

Nursing homes serve disproportionately poor individuals, with a large number of Medicaid enrollees. Vulnerable seniors residing in such long-term care facilities represent 42 percent of U.S. COVID-19 fatalities, while residents of such facilities only account for 0.6 percent of the total U.S. population.⁴

⁴ G. Girvan and A. Roy, Nursing Homes & Assisted Living Facilities Account for 42% of COVID-19 Deaths. The Foundation for Research on Equal Opportunity. 2020 May 7: https://freopp.org/the-covid-19-nursing-home-crisis-by-the-numbers-3a47433c3f70?source=collection_home---1-----0-----; accessed September 21, 2020.

In part this is due to disastrous decisions taken by some state governors to force nursing homes to accept COVID-infected patients who had been discharged from a hospital, including New York, New Jersey, and Michigan.⁵ This catastrophic policy helped spread COVID-19 in long-term care facilities, leading to needless deaths and additional hospitalizations that we then asked our health care personnel to take on.

Figure 2. COVID-19 Deaths in Long-Term Care Facilities as a Share of Total COVID-19 Deaths (as of August 31, 2020)



0.6% of Americans live in long-term care facilities that account for 42% of all COVID-19 deaths. In some states, this tragedy was compounded by policies that forced nursing homes to accept patients infected with the novel coronavirus SARS-CoV-2. (Source: G. Girvan and A. Roy, FREOPP.org)

Importantly, these policy mistakes were driven by poor data. In the spring, too many policymakers were unaware of the fact that residents of long-term care facilities were especially vulnerable to COVID-19, and did not take adequate precautions to limit the spread of the coronavirus in these facilities. For example, New York only began to track

⁵ A. Roy, The Most Important Coronavirus Statistic: 42% of U.S. Deaths Are From 0.6% Of The Population. *Forbes*. 2020 May 26: <https://www.forbes.com/sites/theapotheary/2020/05/26/nursing-homes-assisted-living-facilities-0-6-of-the-u-s-population-43-of-u-s-covid-19-deaths/#232a01f074cd>; accessed September 21, 2020.

nursing home deaths in mid-April. Michigan only began reporting such deaths in mid-May. These states and others should have done more to route testing supplies to LTC facilities early on, and restrict visitation from relatives in affected communities.

Some states continue to produce misleading data regarding nursing home cases and fatalities. For example, New York counts deaths of nursing home residents that occur in a hospital as hospital deaths, not as nursing home deaths, resulting in a significant undercount of the severity of the pandemic in long-term care facilities.⁶

MISLEADING DATA FROM COVID-19 PCR TESTING

There are three principal categories of COVID-19 tests available in the U.S. Antibody tests, sometimes called serology tests, detect the presence of anti-SARS-CoV-2 antibodies in a patient's serum, indicating that a patient has likely been infected by the coronavirus in the recent past. Antigen tests detect the presence of certain viral proteins in a patient sample, indicating an active infection. RT-PCR tests detect the presence of viral genetic material in a patient sample.

While we do not know how many tests of each type are being performed in the U.S.—itself an important gap in COVID-19 data—a critical problem has emerged with RT-PCR testing, which is considered to be the most accurate and reliable method for detection of an active infection.

An investigation by Apoorva Mandavilli of the *New York Times* found that a critical piece of data is missing from most PCR test results: the cycle threshold, or Ct, needed to detect SARS-CoV-2 RNA.⁷

RT-PCR works by amplifying a targeted sequence of viral RNA. Each cycle of amplification roughly doubles the amount of viral genetic material in a given sample. Hence, ten cycles of amplification yields a roughly 1,000-fold amplification ($2^{10} = 1,024$), and 20 cycles yields a roughly million-fold amplification. (The precise amount of amplification can vary based on a number of factors.)

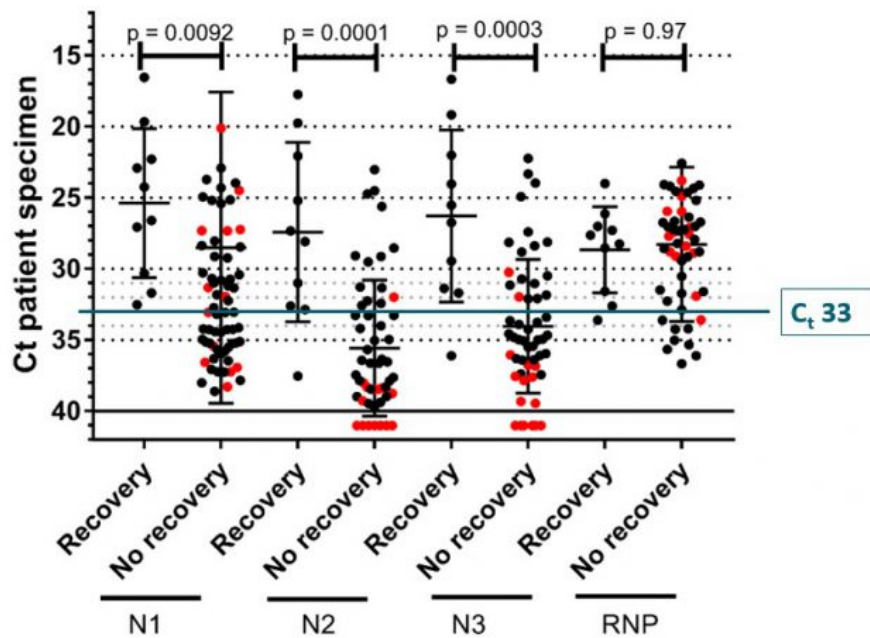
Much like turning up the volume on a quiet musical recording, the amount of viral genetic material in a given sample is inversely proportional to the number of amplification cycles needed to detect it. That is: the more amplification required, the fewer copies of the virus are in the sample. Higher concentrations of viral particles in a patient's serum may correlate to more severe illness and/or a higher probability of viral transmission, though in the case of SARS-CoV-2 this is not well understood.

⁶ J. Vielkind et al., In Worst-Hit Covid State, New York's Cuomo Called All the Shots. *The Wall Street Journal*. 2020 Sep 11: <https://www.wsj.com/articles/cuomo-covid-new-york-coronavirus-de-blasio-shutdown-timing-11599836994>; accessed September 21, 2020.

⁷ A. Mandavilli, Your Coronavirus Test Is Positive. Maybe It Shouldn't Be. *The New York Times*. 2020 Aug 29: <https://www.nytimes.com/2020/08/29/health/coronavirus-testing.html>; accessed September 21, 2020.

Nonetheless, research by the U.S. Centers for Disease Control & Prevention found that above 33 cycles of amplification (i.e., a cycle threshold, or Ct, of 33), the probability of detecting viral particles that are capable of replicating is very low.⁸

Figure 3. PCR Cycle Thresholds Required to Detect Replication-Competent SARS-CoV-2



Replication-competent SARS-CoV-2 is rarely found above Ct 33. CDC data indicates that replication-competent coronavirus was recovered from patient samples where SARS-CoV-2 was detected in 33 amplification cycles or fewer. (Source: CDC)

Similarly, research from the Robert Koch Institute—the German equivalent of the CDC—found a “loss of cultivability in cell culture [of SARS-CoV-2] corresponded to a Ct value > 30.”⁹ The European Centre for Disease Prevention and Control warns that “a high Ct value (e.g. >35) could be due to...contamination of reagents” and re-tested with a “second gene

⁸ Centers for Disease Control & Prevention. Duration of Isolation and Precautions for Adults with COVID-19. 2020 Sep 10: <https://www.cdc.gov/coronavirus/2019-ncov/hcp/duration-isolation.html>; accessed September 21, 2020.

⁹ Robert Koch Institute. Instructions for testing patients for infection with the novel coronavirus SARS-CoV-2. 2020 Aug 11: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Vorl_Testung_nCoV.html; accessed September 21, 2020.

target.”¹⁰ A meta-analysis of 25 studies published in the pre-print journal medRxiv found that “a cut-off of RT-PCR Ct > 30 was associated with non-infectious samples.”¹¹

While exact thresholds will vary by test, then, it appears that individuals for whom SARS-CoV-2 genetic material can only be detected above 30-35 cycles are unlikely to be at high risk of severe COVID-19 illness, and their test results may even be falsely positive in many cases.

In her investigation of U.S. PCR testing practices, however, the *New York Times*' Mandavilli found that it was most common for laboratories to deploy 40 cycles, with some labs amplifying for 37 cycles. “I’m shocked that people would think that 40 could represent a positive,” said Juliet Morrison, a virologist at the University of California, Riverside. Morrison agreed that a Ct cutoff of 30-35 would be more reasonable.

At Mandavilli’s request, New York state’s Wadsworth Center examined 872 positive PCR test results it had obtained in July, after amplification for 40 cycles. “With a cutoff of 35,” Mandavilli reported, “about 43 percent of those tests would no longer qualify as positive. About 63 percent would no longer be judged as positive if the cycles were limited to 30.” In Minnesota, of 300 positive tests from a state lab, roughly half were positive at 30 cycles or fewer.¹²

Michael Mina, a Harvard epidemiologist, told Mandavilli that the cutoff should be 30 “or even less,” and indicated that 85 to 90 percent of those testing positive in Massachusetts in July at 40 cycles would have been classified as negative at a cycle threshold of 30. “I would say that none of those people should be contact-traced, not one.” Other experts interviewed by Mandavilli were “stunned” or “shocked” to learn of this information.

Obviously, if between 63 and 90 percent of PCR tests are falsely positive, due to a cycle threshold of 40 instead of 30, this has significant implications for the U.S. COVID-19 policy response.

For example, in California’s four-tier system of economic restrictions, “many non-essential indoor businesses are closed” in counties with a test positivity rate of more than 8%, or with more than 7 new cases per 100,000 residents. As of September 21, 2020, the majority of California counties were above one or more of these thresholds.¹³

¹⁰ European Centre for Disease Prevention and Control. Questions and answers regarding laboratory topics on SARS-CoV-2. <https://www.ecdc.europa.eu/en/all-topics-z/coronavirus/threats-and-outbreaks/covid-19/laboratory-support/questions>; accessed September 21, 2020.

¹¹ T. Jefferson et al., Viral cultures for COVID-19 infectivity assessment—Systematic review. medRxiv. 2020 Sep 3: <https://www.medrxiv.org/content/10.1101/2020.08.04.20167932v3>; accessed September 21, 2020.

¹² J. Olson, Minnesota’s broad COVID-19 testing under microscope. *Minneapolis Star-Tribune*. 2020 Sep 12: <https://www.startribune.com/broad-covid-19-testing-under-microscope/572396572/>; accessed September 21, 2020.

¹³ State of California. Blueprint for a Safer Economy. 20 Sep 2020: <https://covid19.ca.gov/safer-economy/>; accessed September 21, 2020.

Economic restrictions are harmful enough on their own, and have a disproportionate impact on low-income and minority populations.¹⁴ But they are doubly harmful if they are excessive or unwarranted, for example if positivity rates and case rates are being inflated by overamplification of RT-PCR samples. In California’s case, a summer breakdown in the state’s infectious disease monitoring system made it even more difficult to develop an effective policy response.¹⁵

DELIVERING LIFE-SAVING TREATMENT TO SERIOUSLY ILL COVID-19 PATIENTS

Remdesivir, an antiviral drug manufactured by Gilead Sciences, received Emergency Use Authorization from the FDA in May, after a double-blind, randomized, placebo-controlled trial indicated that hospitalized COVID-19 patients treated with remdesivir recovered four days earlier, on average, on treatment, with a 40% reduction in mortality.¹⁶

Gilead expects to be able to meet world demand for remdesivir later this year. However, at present, supply of the drug is subject to manufacturing constraints. Hence, directing scarce remdesivir inventory to facilities with high numbers of hospitalized COVID-19 patients is literally a matter of life or death.

The National Healthcare Safety Network, a surveillance system managed by the CDC, is the traditional source of such data. However, of the 6,200 hospitals in the U.S., only 3,000 regularly report COVID-19-related data through NHSN. This meant that NHSN lacked visibility into the remdesivir needs of more than half of all U.S. hospitals. In addition, NHSN ran on antiquated software and hardware that makes it difficult to modernize or upgrade.

Indeed, on four separate occasions this century, Congress has passed laws requiring HHS and CDC to modernize their situational awareness network capabilities. Until now, none of these statutes have been effective:

- In 2006, Congress passed the Pandemic and All Hazards Preparedness Act (PAHPA), which required HHS to “establish [by 2008] a near real-time electronic nationwide public health situational awareness capability through an interoperable network of systems to share data and information to enhance early detection of, rapid response to, and management of, potentially catastrophic infectious disease outbreaks and other public health emergencies that originate domestically or abroad.” In 2010, the U.S. Government Accountability Office concluded that HHS

¹⁴ A. Roy et al., *A New Strategy for Bringing People Back to Work During COVID-19*. The Foundation for Research on Equal Opportunity. 2020 Apr 14: https://freopp.org/a-new-strategy-for-bringing-people-back-to-work-during-covid-19-a912247f1ab5?source=collection_category--1-----1-----; accessed September 21, 2020.

¹⁵ A. Barry-Jester & A. Hart, *California’s Data Failures Stymie Efforts to Curb the Virus*. Kaiser Health news. 2020 Aug 21: <https://khn.org/news/californias-data-failures-stymie-efforts-to-curb-the-virus/>; accessed September 21, 2020.

¹⁶ J. Beigel et al., *Remdesivir for the Treatment of COVID-19—Preliminary Report*. *New England Journal of Medicine*. 2020 May 22: <https://www.nejm.org/doi/full/10.1056/NEJMoa2007764>; accessed September 21, 2020.

had not done so, nor had it even developed a “comprehensive strategic plan [as] required by PAHPA.”¹⁷

- In **2013**, Congress passed the Pandemic and All Hazards Preparedness Reauthorization Act (PAHPRA), this time requiring HHS to submit a comprehensive strategy within 180 days. This did not happen. GAO in 2017 concluded that HHS had not taken “measurable steps for completing and tracking the status of the activities required by the law.”¹⁸
- In **2019**, Congress passed the Pandemic and All-Hazards Preparedness and Advancing Innovation Act (PAHPAIA), building on the previous mandates. CDC has stated that it has not hired new biosurveillance specialists in accordance with the statute.¹⁹
- In **2020**, Congress passed the Coronavirus Aid, Relief, and Economic Security Act (CARES), which authorized \$1 billion for modernizing public health data infrastructure, and required CDC to develop a plan by April 30. The plan has yet to be published.

The urgent necessity of improving surveillance data, especially from hospitals, first arose during supply constraints related to the distribution of PPE and ventilator equipment. In response to this problem, and in anticipation of supply challenges with emerging therapies, CDC and the U.S. Department of Health & Human Services established HHS Protect, which aggregates data from NHSN as well as from state public health agencies, private vendors, and hundreds of other data sources.

In July, CDC and HHS required that states the report data to HHS Protect do so through a system developed by TeleTracking, a private-sector vendor. The new system enabled public health officials to gain access to COVID-19 data from an additional 3,100 hospitals: 2,000 who report directly to HHS or to state governments; and 1,100 using TeleTracking.²⁰ As a result, HHS Protect now has access to COVID-19 data from over 90% of U.S. hospitals: a substantial improvement from the legacy NHSN system.

While some have claimed that HHS Protect “sidelines” the CDC and the NHSN, CDC Director Robert Redfield has directly contradicted those claims:

¹⁷ U.S. Government Accountability Office. Public Health Information Technology: Additional Strategic Planning Needed to Guide HHS’s Efforts to Establish Electronic Situational Awareness Capabilities. 2010 Dec 17: <https://www.gao.gov/products/gao-11-99>; accessed September 21, 2020.

¹⁸ U.S. Government Accountability Office. Public Health Information Technology: HHS Has Made Little Progress toward Implementing Enhanced Situational Awareness Network Capabilities. 2017 Sep 6: <https://www.gao.gov/products/GAO-17-377>; accessed September 21, 2020.

¹⁹ J. White & D. Badger, In Order to Defeat COVID-19, the Federal Government Must Modernize Its Public Health Data. The Heritage Foundation. 2020 Sep 3: <https://www.heritage.org/health-care-reform/report/order-defeat-covid-19-the-federal-government-must-modernize-its-public>; accessed September 21, 2020.

²⁰ J. Arrieta, Prepared Remarks from HHS Media Call with CDC Director Redfield and CIO Arrieta on COVID-19 Data Collection. HHS Press Office. 2020 Jul 15: <https://www.hhs.gov/about/news/2020/07/15/prepared-remarks-from-hhs-media-call-cdc-director-redfield-cio-arrieta-covid-19-data-collection.html>; accessed September 21, 2020.

As many of you know, CDC operates a system called the National Health Safety Network. This is an important surveillance system in our nation’s hospitals, which focuses on fighting antibiotic resistance.

In April, HHS leaders, with input from CDC, created a new system, called HHS Protect, that allows us to combine data through systems like NHSN, as well as other public and private sources. The data reported from hospitals that went into HHS Protect either came through the NHSN, directly to HHS Protect from the states, or through a system called TeleTracking.

What we have now asked is that, going forward, states provide data from hospitals directly through the TeleTracking system or directly to the HHS Protect system.

First, this reduces the reporting burden—it reduces confusion and duplication of reporting. Streamlining reporting enables us to distribute scarce resources using the best possible data.

TeleTracking also provides rapid ways to update the type of data we are collecting—such as adding, for instance, input fields on what kind of treatments are being used. In order to meet this need for flexible data gathering, CDC agreed that we needed to remove NHSN from the collection process, in order to streamline reporting.²¹

Importantly, the new reporting system requires hospitals to report more details regarding patient demographic and clinical characteristics, so that remdesivir can be sent to hospitals with patients who will benefit from treatment. “No one is taking access or data away from CDC,” said Redfield. “The new infrastructure we have now actually provides our CDC team with easier access to a much broader variety of data sets than they would have without it. Approximately 1,000 CDC experts have, and continue to have access to the raw data collected in HHS Protect—in addition to thousands of other public health professionals across HHS.”

THE BENEFITS OF A MODERNIZED COVID-19 DATA SYSTEM

Switching to a new reporting system in the middle of a pandemic is difficult to do, of course, and HHS and CDC have acknowledged that they could have done better in communicating the reporting change to hospitals. But the new system should save lives, and not just due to more effective delivery of remdesivir.

According to CDC Director Redfield, HHS Protect frees up personnel at NHSN “to increase its focus on...nursing home and long-term care facility reporting needs...streamlining the hospital reporting system allows NHSN to concentrate its COVID-19 activity on the high-priority area of protecting the vulnerable in nursing homes.”

In addition, a modernized data aggregation system could compile detailed RT-PCR lab results, especially from large commercial lab companies that are able to report Ct values. A more systematic analysis of Ct values, through HHS Protect, should yield important insights into the true COVID-19 caseload, and discover any correlations between SARS-CoV-2 viral

²¹ R. Redfield, Prepared Remarks from HHS Media Call with CDC Director Redfield and CIO Arrieta on COVID-19 Data Collection. HHS Press Office. 2020 Jul 15: <https://www.hhs.gov/about/news/2020/07/15/prepared-remarks-from-hhs-media-call-cdc-director-redfield-cio-arrieta-covid-19-data-collection.html>; accessed September 21, 2020.

load, illness severity, and transmission. These insights could prove important in the treatment and management of COVID-19, and in reducing the spread of the coronavirus.

Avik S. A. Roy

President and Co-founder, The Foundation for Research on Equal Opportunity



Avik S. A. Roy co-founded The Foundation for Research on Equal Opportunity (FREOPP) in July of 2016. The non-partisan think tank researches policy that affects those below the average median income in the U.S. FREOPP has recently conducted studies on COVID-19 policy entitled “Reopening America’s Schools and Colleges During COVID-19” and “A New Strategy for Bringing People Back to Work During COVID-19.”

Avik has had a career in finance, journalism, and public policy. He worked as an analyst and portfolio manager at the investment firms Bain Capital and JPMorgan Chase before he left to found a healthcare-focused hedge fund. In early 2012, Roy founded Roy Healthcare Research, a health care policy research firm and worked as the principle.

His career in journalism began in 2009 with his blog *The Apothecary*, which focused on healthcare policy and the Affordable Care Act. In 2011 he became a senior fellow at the Manhattan Institute with a focus on health care policy. In 2013 he published his book *How Medicaid Fails the Poor*. Through the Manhattan Institute he authored a proposal for health care entitled *Transcending Obamacare: A Patient-Centered Plan for Near-Universal Coverage and Permanent Fiscal Solvency* and *The Case Against Obamacare*. He worked at Forbes as a contributor in 2010, an opinion editor in 2014, and since 2018 has been a policy editor. He has made appearances as a commentator on Fox News, Fox Business, MSNBC, CNBC, Bloomberg Television, PBS’s Newshour, and HBO’s Real Time with Bill Maher.

His political and policy career includes work on Romney for President in 2012 as part of the Health Care Policy Advisory Group. Avik helped formulate Governor Romney’s national health reform plan and advised the campaign on other health care-related matters, including Medicare & Medicaid reform. Avik was a senior policy advisor for Perry for President in 2015. He led major policy initiatives on economic opportunity for all Americans, reforming Wall Street, and defeating Islamic jihad.

His current policy work, in addition to his role as president at FREOPP, includes membership of the board of advisors for both the National Institute for Health Care Management and at the Center for Advancing the American Dream. Additionally, Mr. Roy is a senior advisor for the working group on health care reform at the Bipartisan Policy Institute. Avik earned a B.S. in biology from MIT and attended the Yale University School of Medicine.