

Testimony for the Record
Submitted to the House Committee on Energy and Commerce
Subcommittee on Health for the Hearing
“Empowered by Data: Legislation to Advance Equity and Public Health”

June 24, 2021

Beth I. Blauer, J.D.
Associate Vice Provost for Public Sector Innovation
Johns Hopkins University

Chairwoman Eshoo, Ranking Member Guthrie and members of the Subcommittee, thank you for inviting me to participate in today’s hearing and for dedicating your time to an examination of the role of data in health outcomes and the capacity of data to assist in the dismantling of entrenched structural racism.

For the last six years I have led a center at Johns Hopkins University (JHU) focused on building the capacity of local leaders to use data to improve outcomes. For the last 17 months, I have been the data lead for Johns Hopkins University’s Coronavirus Resource Center. I have seen the very best and the very worst of data use. I am pleased that this Congress is recognizing the role of data not only in improving outcomes but also in contributing to health disparities and their devastating outcomes. The bills that are the subject of today’s hearing go a long way to realign federal resources with interventions that are proven, measurable, and focused on ending multigenerational health and wellbeing disparities.

To begin, we must acknowledge the role data has played in exacerbating health disparities and invest in data collection methods to prevent further widening of gaps. Whether it’s the historic use of data to create redlined maps to determine where we deploy community-based resources, or inadvertently incentivizing counterproductive policing behavior to produce better metrics/outputs, data has a complicated history in the public sector, at best. I am honored to be here today to describe how the Center advises policymakers at the local, state and national levels to use these same tools to dismantle the systems and the approaches that have created current disparities and improve government performance.

COVID-19 Tracking at JHU

Take COVID-19 data reporting, for example. In the last 18 months, government at all levels from municipalities to federal agencies, did something incredibly remarkable. These governments built data collection efforts, shared data, and made real time decisions that were based on near time data. Never before has the nation endeavored to realize a coordinated effort around data sharing, data informed decision making, and collective outcome measurement at such scale. Local and state governments quickly built a capacity for looking at data around the clock, deploying resources throughout their communities, measuring the effects of their

interventions, using every possible lever to stall the spread of this disease including closing businesses and schools.

These efforts were not without their challenges. States scrambled to understand how to define data, how to share data with the public, and how to communicate policy decisions that were based in data. The confusion that quickly surrounded these efforts was due in large part to a lack of unified data collection and reporting standards. The result? Many States defined Covid cases, deaths, and testing differently and these definitional variances made comparing conditions difficult, made fluid borders confusing, and made policies difficult to enforce. As difficult as this effort was, imagine if we didn't have real time access to this data. Imagine if we accepted a years-long delay getting this data and responding to this devastating virus?

So, what has this last year plus taught us about using data to collectively solve problems?

At the start of 2020, Johns Hopkins launched the Coronavirus Resource Center, which quickly became a trusted resource for millions of viewers worldwide. International health organizations, the federal government, and institutional leaders faced challenges in mobilizing to get data collected or shared given the unprecedented circumstances. Early in 2020, an interdisciplinary group of researchers at JHU began to manually collect international health data that was often released in unstructured formats, press releases, public announcements and government dashboards. Over the course of weeks, JHU developed a methodology for scraping public data and encouraged state and local governments to share their data in standardized formats. JHU data scientists set up internal governance and created definitions with the help of public health and medical experts. Our team articulated standard collection methodologies and openly shared this process with the public. At first, we built data visualizations for fellow researchers but quickly learned that the public was urgently looking to source this data, as well.

By mid-April of 2020 we had millions of viewers visiting our site every day and by January 2021, we had accrued more than a billion views. Our audience included news outlets, local governments, and everyday people that were making deeply personal decisions about how they would navigate their public lives. Families all of a sudden faced with weighing the safety of going to the grocery store or sending children back to school needed reliable, understandable and localized information. The backdrop to our entire pandemic experience was and continues to be a hunger for sound, publicly-available data. Now imagine if we had access to the same level of data but about the health of our communities, the quality of our air, the safety of our communities, the social determinants of our health? Not data that takes more than a year to process like our current collection methods for the very comorbidities that complicated pandemic response. Not data that is only collected for the purposes of reporting on big block grants, but the type of data that can help policymakers immediately invest in interventions that work, in the places that need those investments the most. The type of data that can immediately be used to retire broken systems and help reallocate scarce resources to systems that are proven to be effective.

This Congress has an opportunity to capitalize on this public demand for data, the financial investment in data infrastructure and the newly minted analytic skill that has emerged during the COVID-19 pandemic and improve upon the systems to provide accountability, accessibility, consistency, equity, and sustainability. The bills under review today are an encouraging step in this direction.

The data provided by local, state and national COVID-19 dashboards have led to major policy decisions across all levels of government and influenced the behavior and decisions of many Americans. The average American can now access COVID data, investigate hotspots, find information about vaccination locations and infection rates, and browse pertinent articles directed at the lay public. Due in large part to these efforts, test positivity rates are now dramatically decreasing across the country. While lower infection rates in nearly every State are cause for measured optimism, it is crucial that we maintain the infrastructure developed during the pandemic, both to monitor the continued spread of COVID-19 and to leverage in the fight against the many public and social health crises that our communities face. The Federal government can encourage growth in these collection efforts and reverse the concerning trend of divestment in data infrastructure that has begun in more than half of the states in this country.

We Need Data Standards

Pandemic data collection is not flawless and our national data strategy could be improved through data standardization. The pandemic revealed that the systems that feed these dashboards are underdeveloped and often arduous. Across the many agencies of the federal government there is no one common language to help ease the often-confusing process of collecting program level data. Definitions vary in the simplest of terms, multiple agencies with different ways of collecting addresses, enrollment, the definition of recipient or household variables that make understanding conditions in our communities particularly challenging. This will not be solved by one agency, this is an interagency dilemma that requires a centralized administrative focus. COVID-19 data collection again proves illustrative in this context.

We know federal data standards work when they are established. Hospitalization data has some of the best temporal and geographic resolutions in the realm of COVID-19 data due to standardized data reporting methods mandated by the Centers for Medicare & Medicaid Services. The U.S. Department of Health and Human Services releases the hospitalization data it receives from the states daily. This daily reporting of data allowed for early identification of the shift in trends before the COVID-19 resurgence in the Spring. This equipped local leaders to enact new policies to curb the rise in cases, and to request aid from the federal government. This is one of the only cases where federal standards were established for pandemic data, and they work well for the purposes of governance and decision making, but the standards were not well planned, were rolled out with confusing guidance and added additional burden to already strapped health systems.

Today you are discussing H.R. 976, the "Ensuring Transparent Honest Information on COVID-19 Act." This bill would require States to provide many critical data on either a daily or weekly basis. The consistency across States established by the ETHIC Act would allow data scientists to perform in-depth analyses on COVID-19 spread, the efficacy of policy measures and shutdowns, disease resurgence, and much more. This bill could be strengthened by requiring public health data reporting for States moving forward as we have now seen the potential opportunity for near time health data collection and the damage when public health data is not standardized, centralized, and consistently updated.

Without federal mandates, States are now beginning to reduce their data collection efforts and reporting cadences. As of this writing, 32 States and territories were no longer providing COVID-19 data updates every single day as they previously did. Delays in reporting updated public health data could result in delayed identification of resurgence and variants, limiting the potential response of the federal government. The ability of the government to rapidly respond to everchanging data during this pandemic has saved countless lives, and we should aggressively take advantage of this infrastructure, deepen the data skills of program leaders, and begin to widen our regular data collection efforts to align to the important business of measuring public health writ large.

Some states have argued that dashboards are no longer necessary since the CDC tracks pandemic data. While the CDC does aggregate COVID-19 data, it is important to note that the CDC is limited to using data filed through official reporting channels, which are not publicly available and often tied to regulatory frameworks that are not conducive to real-time response. This can result in a serious reporting lag as well as data discrepancies when State updates are out of sync with official federal channels. While the CDC has been an invaluable leader and resource throughout the pandemic, this does highlight the need for robust data collection and reporting systems at the local and state levels to better inform and give context to federal data reported by the CDC. This is not the time to slow data collection and reporting; instead, we must solidify and unify data systems to end COVID-19, address the multifaceted challenges of public health our communities were facing in pre pandemic times, and prepare for the next public health crisis.

Data standards and a strong cross government data governance approach would be welcomed by local leaders. Over the course of this pandemic, we at JHU have been in constant communication with Governors and Mayors, and health departments as they have navigated their role in data collection, data use and data sharing. Leaders at all levels were desperate for guidance on everything with simple definitions to positivity calculations. Enforcing data standards would be a welcome challenge to the leaders in our country who have unleashed the power of data use in their communities to good effect.

Demographic Data is in Disarray

State and federal demographic data does not align. State reports to the CDC do not always match the state-reported data on their own dashboards. Inconsistencies in categorization between states make data incomparable and can obfuscate the disproportionate effects that the pandemic has had on people of color. Some states even lack standardization within their own health departments, using different demographic categories between testing data and vaccination data. Without standards there is no way to analyze available data to locate vulnerable populations. Our current system does not allow for simple analyses such as determining whether black people or white people are more likely to die from COVID-19 or have access to life saving vaccines.

Local politicians have been tasked with protecting their citizens, but they are not equipped with the complete datasets they desperately need to make critical decisions. Due to inconsistencies and lack of communication between local municipalities, states, and the federal government, local leaders are forced to make decisions without the full power of data behind them, which is unacceptable during a public health crisis but should always be unacceptable. As of now, local leaders can only design targeted outreach strategies by guessing where hotspots of unvaccinated individuals exist through comparisons of incomplete demographic data on people vaccinated with census-level population data.

Additionally, the lack of comprehensive demographic data cloaks the disproportionate impact the virus is inflicting on minority communities and prevents local leaders from designing relief programs. This dynamic is not unique to COVID. States often only provide demographic data in an aggregated form, preventing local governments and the public from accessing detailed demographic information for their areas. This information would inform policy decisions of city and county leaders by more accurately identifying hotspots, opportunities to invest in public health assets, and track disease outbreaks in near real-time. The data can and should be publicly available, disaggregated, randomized, individualized, and anonymized, so that policymakers and the public can make safe, informed decisions and identify immediate risks.

Make the Data Public

Data collected by the government and funded by the taxpayers should be public. While I applaud that many of these bills require data collection in a manner that is anonymized, disaggregated, and stratified by race, ethnicity, age, sex, geographic region, they do not all provide a plan for public dissemination of the data. These data will be high-quality, high-resolution, versatile tools for research, public health planning and pandemic preparedness.

The JHU COVID efforts were purposefully public in every possible way. We have learned the value of opening up data to the public. Obviously not everyone will use raw data, but often local views of the data help identify hidden clues that can unlock stronger and healthier communities. Mayors and county administrators have relied on the public release of data to get a better sense of the impact the pandemic is having on their communities. The complete lack of intergovernmental data sharing agreements deepened local leaders' dependence on public

data. Just as money follows the well-worn paths of the large block grant programs from the federal government to individuals, so should the data. Data on who is most in need, insights on the efficacy of best practice, and the outcomes programs were designed to achieve. Data, in compliance with privacy protections afforded in our laws, should be shared with policy makers and with the public to build trust, to change our behavior and to incentivize participation in the solutions we seek.

Government will not be able to turn the tide on the social determinants of health alone, it requires deep coordination with government at all levels and public engagement in the most intimate of ways. Our center at JHU continues to work to build the capacity of local leaders to use more data as they examine their own practices and architect on the ground strategies to deliver better outcomes for people. Our communities can see the difference and the pandemic is another example of how government data shaped our outcomes. Over the last year we wore our masks, we stayed home, we sacrificed, and now we are beginning to recover. Imagine if we applied the same level of focus to the generational challenges our communities have endured for far too long.

I would now be pleased to take any questions you may have.

The opinions expressed herein are my own and do not necessarily reflect the views of The Johns Hopkins University.