

National Commission to Transform Public Health Data Systems



Roadmap to the white papers and supplemental materials

White Paper	Brief Overview
Preamble	This paper explains the Commission, the foundational white papers, defines terms, describes the Commission approach, and outlines the guiding principles of a modern, equity-oriented public health data system and what that system does. This includes a brief summary of past and present data modernization initiatives.
Why? The focus on equity in the modern public health data system	This paper focuses on the philosophy around equity and what equity means in the context of the entire objective of the transformed public health data system. This paper explores issues of voice, power, use of data, and so forth in the context of equity orientation.
What? The data in the modern public health data system	This paper focuses on the content of the data—what the United States is tracking (objectives, measure areas), what type of data should the nation track, how should the United States handle new types of data and new volumes and varieties of data, and so forth.
How? The design of the modern public health data system	This paper focuses on the system infrastructure including governance, legal and privacy issues, operability, coordination, data stewardship, ethics and so forth that guides the modern public health data system.
Who and What Next? The sector actions to transform the public health data system	This paper includes nods to the sector contributions needed to transform the public health data system. What needs to happen now with the chance of sustainability? What might be risks to transformation?
Supplemental Materials	Table: Federal Government Activities Related to Equity and the Public Health Data System; Glossary of Working Definitions for Terms Used in White Papers

Transforming Public Health Data Systems

Opportunities and Paths
Forward

Preamble

This white paper is part of a series of papers for use by the National Commission to Transform Public Health Data Systems, an independent commission convened by the Robert Wood Johnson Foundation (RWJF). These papers were used to inform Commission discussions and recommendations. The papers represent insights from literature review, a diverse set of stakeholder perspectives, and inputs from RWJF grantee projects related to the topic of the Commission. The white papers were developed in support of the Commission by the RAND Corporation, a nonprofit research organization, within its division, RAND Social and Economic Well-Being. Corresponding authors are [Anita Chandra](#) and [Laurie Martin](#).

**National Commission to
Transform Public Health
Data Systems**

May 2021

Table of Contents

A Motivation and an Urgent Opportunity	2
Commission Charge	3
A Note on Terms: Public Health Data and the Public Health Data System, Modern and Equity-Oriented	3
The Commission White Papers: A Brief Overview	5
The Commission Approach	7
Guiding Principles for the Future Public Health Data System	7
Principle 1	8
Principle 2	10
Principle 3	11
Principle 4	12
Principle 5	13
Principle 6	14
Principle 7	14
The Remaining Papers	16
Appendix A. Key public health data modernization and innovation efforts, including a sample of CDC linked efforts on modernization and health equity	16
Acknowledgements	19
References	19

A Motivation and an Urgent Opportunity

The COVID-19 pandemic along with calls for racial justice and concerns about widening economic gaps were referenced as the triple pandemic in 2020, stressing institutions and magnifying fractures in the systems meant to address these challenges.¹ One of those systemic issues that has been highlighted for months, but has been enduring, is the design, quality, utility and stewardship of public health data. In a March 2021 article for *The Atlantic*, Meyer and Madrigal poignantly noted, “To avoid another data calamity, our public health system must expend as much energy on understanding the present as it does on modeling the future. Governing through a pandemic—or any emergency—is about making the least-bad decisions with the best information available.”² Through a series of [executive orders](#), the Biden administration has advanced the importance of developing more robust federal datasets that can measure equity better than what was observed during the pandemic.³ Federal datasets could be disaggregated by demographic characteristics and provide foundational information that helps address inequities.

While the COVID-19 crisis has illuminated deep inconsistencies and incompleteness in information that is used nationally and locally regarding who is being infected and how people are faring, such failures of basic public health surveillance are not the only gaps in the current public health data system. There are questions about whether public health data and the system that surrounds it are offering enough data on the *health* of the nation (versus its level of “sickness”), the systemic problems that drive health inequities, and the tools that target priorities for public health action. In this discussion of gaps is a request for a modern **public health data system** that is more than simply a collection of individual data points and is defined as the actors and sectors *with* data and agency to make decisions to advance the health and well-being of a community, population, and/or nation equitably.

The cracks in the public health data system are not only a “bad news” story. The current discussion of the public health data system offers a fresh opportunity to create a system that can tell a better story of national health commitment. Because 2020 was characterized by a confluence of mental distress, an examination of racial inequity, and difficult decisions about how the United States cares for those often the most vulnerable (e.g., children, seniors, those with chronic health conditions), the public health data system can use this time of overwhelming health strains to take the leap (or the “moonshot”) to demonstrate national health values differently through new approaches to data.

Given this context and against the backdrop of its vision to build a Culture of Health in America, along with commitment to addressing health equity and advancing well-being, the Robert Wood Johnson Foundation (RWJF) has convened a **National Commission to Transform Public Health Data Systems** to catalyze an agenda and recommendations to effectively and sustainably address these challenges in the public health data system. While it is clear that the nation has significant work to do in the public health *data* system, it is also important to consider that there is need for a revitalized and re-envisioned public health system as well, upon which the new data system can be situated. Though this Commission is focused on public health *data* systems, the potential of these meeting deliberations to propel that larger public health system discussion, inclusive of how data are used to guide decisions and promote action, is noted.

Commission Charge

The purpose of the **National Commission to Transform Public Health Data System** is to provide concrete and actionable recommendations that will move the current American public health data system into a modern, equity-oriented public health data system that can provide data, information, and insight to improve the health and well-being of all in the United States, with attention to issues of equity. The Commission should set the vision and principles for that public health data system, helping to operationalize the ambitions of these characteristics—modern and equity-oriented. The Commission represents thought leaders in public health data and the issues surrounding the content, structure, and use of data; the equity and ethical considerations associated with the public health data system; and insight about how a modern public health data system should function and be situated in a larger, national context of policies, practices, and decision-making. Led by Dr. Gail Christopher of the National Collaborative for Health Equity, the Commission will build on insights from these white papers, which have been developed by the RAND Corporation with input from experts, grantees to RWJF, and other stakeholders.

This preamble and subsequent white papers, further described in the next sections, are intended to provide a foundation for Commission deliberations and inform the agenda and recommendations. Some elements of these papers may inform the final Commission report, but these white papers are for use by the Commission only and should be viewed as working reports.

A Note on Terms: Public Health Data and the Public Health Data System, Modern and Equity-Oriented

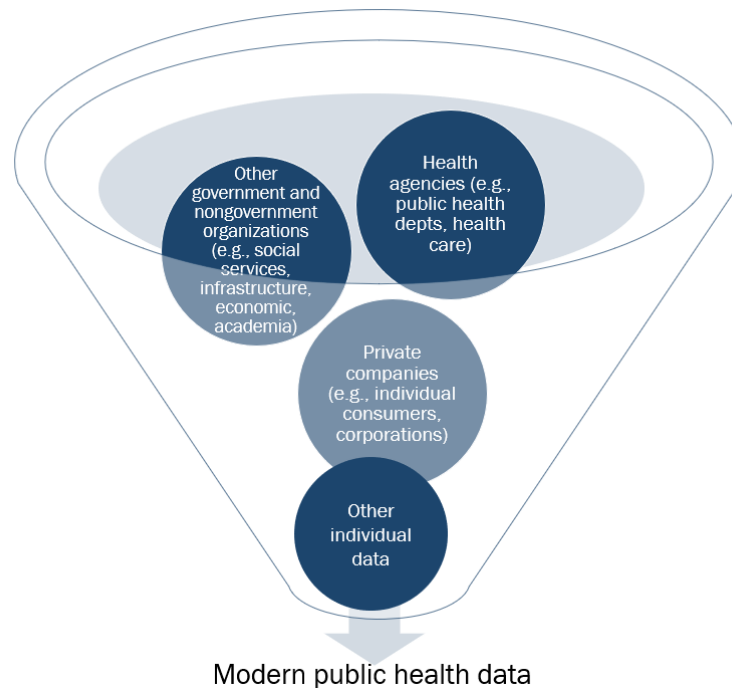
Before going further, it is important to offer working definitions of what public health data and the public health data system entail. The National Commission takes the fullest and foundational definition of the public health system (i.e., defined as “all public, private, and voluntary entities that contribute to the delivery of essential public health services within a jurisdiction”)^a and does not limit this conceptualization to the state or local public health department. Given that context, we offer these longer, working definitions of public health data and the public health data system.

- **Public health data** are any data that can be used to understand, assess, and monitor the health and well-being of a community, population, and/or nation, with the purpose of prospectively determining health assets and needs and retrospectively capturing intergenerational trajectories of health over time. **Figure 1** provides a visual depiction of the key sources of public health data in a modern, public health data system.

^a See CDC Foundation, “What is public health?” (<https://www.cdcfoundation.org/what-public-health>). See CDC, “The public health system,” (www.cdc.gov/publichealthgateway/publichealthservices/essentialhealthservices.html).

- As noted briefly earlier, the **public health data system** is the system or systems of actors and sectors with data and agency to make decisions to advance the health and well-being of a community, population and/or nation, with consideration of how opportunities to be healthy are equitably distributed. The public health data system crosses federal, regional, and local boundaries. For the purposes of the white papers, we use the term “public health data system” for ease and clarity, though we may consider nested public health systems.

Figure 1. Who's generating public health data?



As noted earlier, the Commission is focused on what it means to create a **modern, equity-oriented** public health data system (through vision, principles, and components). For the purposes of these white papers:

- **Modern** is used to describe a data system that is agile, can connect multiple forms of data and translate those data into *information* (defined as useable knowledge) and *insight* (defined as information contextualized in history and structures that can support decision-making), and is fully equipped for the health opportunities and challenges of the 21st century.
- **Equity-oriented** means a public health data system that centers equity, or *the fair and just access to the opportunity to be healthy*,⁴ in the content and structure of data; the engagement of diverse stakeholders in the access to and use of data for decision-making with attention to issues of power, marginalization, and justice; and in how data are used to address systemic inequities.

The Commission White Papers: A Brief Overview

The Commission will convene to prepare the agenda and recommendations for the future public health data system between May and June 2021. This preamble and the four other white papers in this set are meant to serve as foundational documents to describe the current landscape, emerging innovations, and the opportunities for transforming and transitioning into the next era of the public health data system. The white papers are organized into four areas, briefly described in **Table 1**. As noted at the end, this paper organization was informed by a notional set of guiding principles from a RAND-led review between October 2020 and February 2021, which included an environmental scan of literature about data modernization and transformation, a review of peer-reviewed and gray reports on public health data transformation with attention to recent initiatives (**Appendix A** provides summary information on key initiatives but not an evaluation of those initiatives), and a series of stakeholder interviews (N=112) with leaders in health (public health and health care), in data (health and non-health data), in areas germane to social determinants of health (e.g., social services, infrastructure, environment, law), and in technology. The review of past and present data modernization initiatives was only meant to summarize and characterize those efforts but does not include an evaluation of any initiative. The stakeholder interviews were particularly focused on insights from experts in domains such as these:

- Vision for public health data, and gaps and challenges in current public health data
- Content and types of public health data
- Access to and use of public health data
- Public health data governance and stewardship
- Operational elements of public health data (e.g., features of interoperability, legal issues, privacy)
- Role of data providers (public and private sector) and users

Issues of equity were probed within and across all these domains.

In addition to the RAND-led data collection described here, the RAND team integrated insights from a set of RWJF grantees, who received grants during this same period to deepen their understanding of public health data system issues in these domains above, to surface promising practices and case examples of innovation in public health data development and use, and to identify systemic changes to facilitate a modern, equity-oriented public health data system. The Commission's Supplemental Materials includes a report that summarizes a series of expert panels reflecting on transforming public health data systems through the lens of diverse population experiences. Most grantees are continuing their efforts through 2021, but these white papers offer some of the early insights from their work, emerging as of April 2021.

Table 1. Roadmap to the white papers, including the preamble

White Paper	Brief Overview
Preamble	This paper explains the Commission, the foundational white papers, defines terms, describes the Commission approach, and outlines the guiding principles of a modern, equity-oriented public health data system and what that system does. This includes a brief summary of past and present data modernization initiatives (Appendix A).
Why? The focus on equity in the modern public health data system	This paper focuses on the philosophy around equity and what equity means in the context of the entire objective of the transformed public health data system. This paper explores issues of voice, power, use of data, and so forth in the context of equity orientation.
What? The data in the modern public health data system	This paper focuses on the content of the data—what the United States is tracking (objectives, measure areas), what type of data should the nation track, how should the United States handle new types of data and new volumes and varieties of data, and so forth.
How? The design of the modern public health data system	This paper focuses on the system infrastructure including governance, legal and privacy issues, interoperability, coordination, data stewardship, ethics and so forth that guides the modern public health data system.
Who and What Next? The sector actions to transform the public health data system	This paper includes nods to the sector contributions needed to transform the public health data system. What needs to happen now with the chance of sustainability? What might be risks to transformation?

The Commission Approach

The Commission will use these foundational white papers and meeting deliberations to:

- Chart a vision and set of principles for the modern, equity-oriented public health data system.
- Outline a set of recommendations for key sectors: government actors (federal, state, local) as well as civil society, principally philanthropy and community organizations (public and private) to help catalyze action towards this vision.
 - Balance comprehensiveness and efficiency in order to put forth **transformation-minded and system-minded^b** recommendations for the sectors described above, which can be acted upon, starting in 2021. There will be subsequent efforts pursued by RWJF and others to build on the Commission recommendations as well.
- Factor in implementation science in the construction of recommendations, as well as duration and sector responsibilities in the formulation of recommendations. However, the Commission is not developing detailed implementation guidance.

Guiding Principles for the Future Public Health Data System

As noted earlier, this preamble outlines a proposed set of seven guiding **System Principles (Table 2)** for a public health data system, which meets the working definition of the *public health data system*, noted earlier, as well as characteristics of *modern* and *equity-oriented*. These principles emerged from the landscape analysis (including initiatives summarized in **Appendix A**), interviews, and RWJF grantee inputs that informed the organization and content of the white papers. We describe each principle in further detail below. The System Principles are organized by the core functions of the public health data system (i.e., what should the system do), who should be considered in using and acting upon findings from the data system, and what the data system should contain. The Commission will review these System Principles as a springboard for discussions and deliberations, ideally finalizing and ratifying the set of System Principles that undergird the next public health data system. Combined with setting a vision, these principles also should guide Commission recommendations.

^b System-minded means “seeing the system as a system, understanding system dynamics, and then shaping intentions, decisions, and actions in terms of this new perspective.” Transformation-minded means “shaping intentions in terms of deeply transformed system behavior, and *framing success* in terms of system level transformation, shaping the new types of *agency and action* that will actually result in the system veering towards altered states.” Adapted from an interview with Banny Banerjee, Stanford University ChangeLabs, from an article by Leonard Teichert, *Medium* 2019.

Table 2. Seven guiding system principles for the modern public health data system

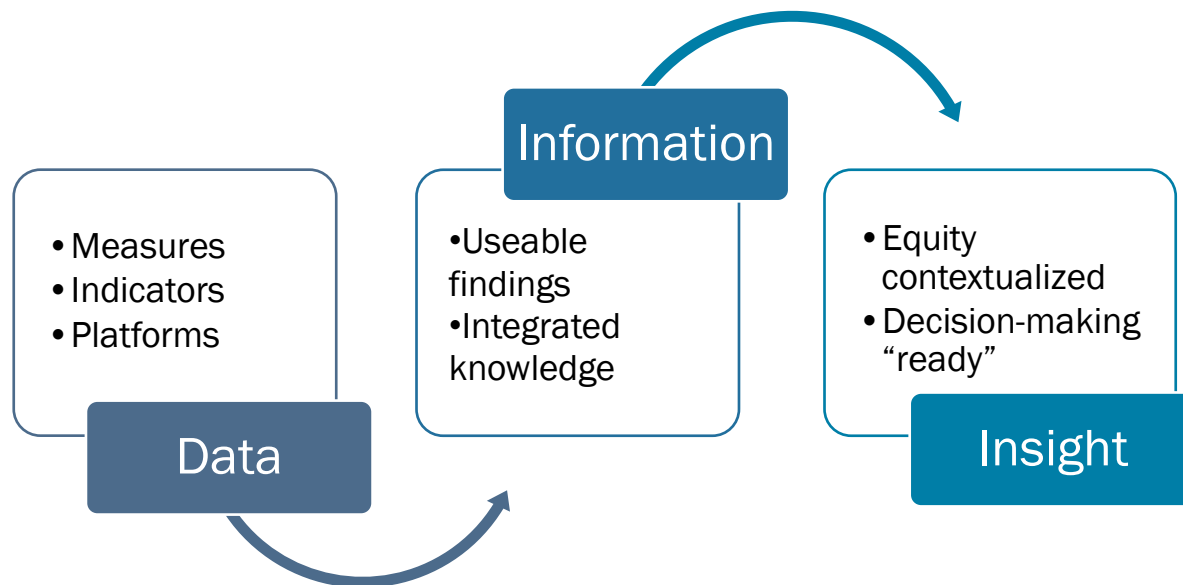
Principles
Principle 1: The public health data system must be more than a collection of numbers; it also must have both sense-making and decision-making functions.
Principle 2: The public health data system must be able to support both <i>steady state</i> health action and response to <i>emergent</i> conditions, including the ability to integrate new information about changing conditions.
Principle 3: The public health data system must support agency among a diversity of actors and sectors.
Principle 4: The public health data system needs to consider all that drives public health change across the five waves of public health. ^c
Principle 5: The public health data system must ensure that data represent a mix of leading and lagging indicators, as well as a balance of data on positive health and well-being.
Principle 6: The public health data system must not simply describe health phenomenon and trends, but also capture the processes by which health outcomes have occurred.
Principle 7: The public health data system needs to balance comprehensiveness and parsimony in data.

Principle 1: The public health data system must be more than a collection of numbers; it also must have both sense-making and decision-making functions

One of the first principles of a public health data system is that it must not only generate data but provide the foundation for *information* and *insight*.⁵ Stakeholders have noted that the current public health data system is not fully serving national or local interests with respect to having information which can prospectively inform decision-making to improve the nation’s health. **Figure 2** provides a graphic of what this *data to information to insight* function of the public health data system should entail. In a public health data system, *data* are aspects to be tracked or sensed, *information* is data analyzed into useable knowledge, and *insight* is information contextualized in systems and history and ready for decision-making. Further, data or information alone will not transform outcomes. As noted in their chapter on health information systems, Stansfield et al. outlined that data are characteristics of people and things, analysis of data enables the identification of patterns, thereby creating information, but it is in the use of information that we get recommendations, rules for action, and ultimately insights.⁶

^c The five waves of public health refer to five stages or eras of public health transformation (Davies et al., 2014), from public health’s origins in hygiene through biomedical and clinical advances to more social and cultural determinants of health.

Figure 2. A modern, equity-oriented public health data system must have the ability to generate outputs in three areas: data, information, and insight



There are two parts of this *data to information to insight continuum*: data sense-making (or the cognitive processes by which people make meaning from data and experiences) and data for decision-making. In a public health data system approach, it is key to include consideration of how users make sense of data and act based on the data, not only on the data itself.

The term **data sense-making** usually means the connection of data to a *frame*, or an explanatory structure that puts items in relationship to each other.⁷ It is well understood from cognitive theory that there are many ways that people process data, in terms of (a) “a reconceptualization of the *data use cycle*, which indicates how data may lead to action; (b) *attribution theory*, which posits that motivation to act is associated with individuals’ perceptions of the causes of outcomes; and (c) *sense-making theory*, which contends with how individuals make meaning of their experiences.”⁸ Acknowledging these interconnections and how cognition and *framing* can help people use data, make meaning of data, and apply these data is an important, yet often overlooked, feature of a data system. A modern, equity-oriented public health data system must fully embed the tools and functions of sense-making to realize positive gains in health outcomes, by informing decisions, ensuring that certain considerations, particularly around equity, are made with the data, and help with prioritization of actions and responses, specifically when dealing with emergent issues like a pandemic. How data are combined with a frame, how data are set on a platform, how data may be linked to other information, who is involved in sense-making, and how data are contextualized are all part of a data system that can support sense-making and impact.⁹

Connected to data sense-making is **data for decision-making** or data for impact. The notion of “data driven decision-making” has become commonplace across sectors. In education, data driven

decision-making has become core to school and district accountability.¹⁰ In health, data for decision-making also has permeated discussions of data surveillance, the creation of data dashboards linked with knowledge management systems, and predictive analytics and tying ‘big data’ to action.¹¹ But despite these advancements, sometimes data driven decision-making in health is hampered by an inability to link information in ways that can inform public health action and challenges to who has access to data for action. There can be structural, ethical, legal, and cultural reasons impeding data linking and data access, all factors that threaten this data to information to insight continuum. As such, the United States has not fully realized the potential for proactive public health decision-making.¹²

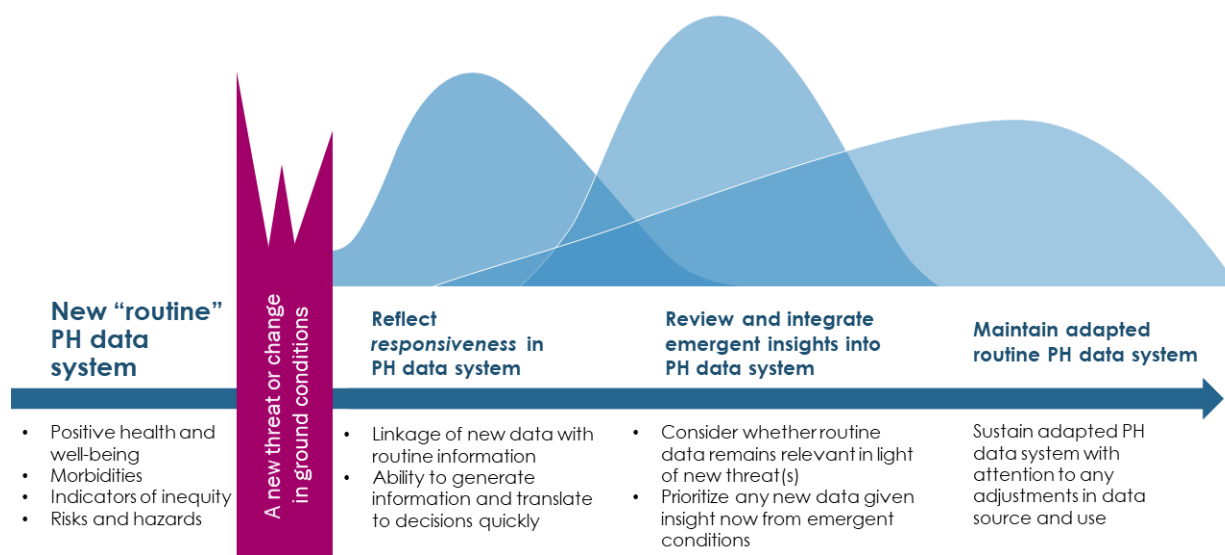
Principle 2: The public health data system must be able to support both *steady state* health action and respond to *emergent* conditions, including the ability to integrate new information about changing conditions.

The COVID-19 pandemic has challenged the agility of data systems to take in new information, to pivot quickly in using real-time data for decisions, and to balance speed against notions of making data “bulletproof.” For instance, data modernization efforts at the Centers for Disease Control and Prevention (CDC) have been on the front lines of this challenge during the pandemic, trying to determine the best ways to disseminate information to the public and local public health departments for transparency and ongoing surveillance as well as to inform strategies and investments, while also recognizing that not all of the data coming in has the level of scrub or rigor often attributed to the usual scientific review of data reports.

The pandemic is the latest disaster that is requiring the public health data system to revisit how it processes information for ongoing health improvement activities (e.g., tracking morbidity) and can surge when an acute event, such as a pandemic or natural disaster, happens.¹³ While stakeholders argue for the need to have that kind of flexibility built into the current public health data system, they acknowledge that this kind of agility is rare. The lack of such agility will become increasingly problematic as the nation deals with new infectious diseases and other conditions and events for which we have limited warning, and even less of a blueprint to follow in response and recovery efforts.¹⁴

Figure 3 presents a notional way to consider how a modern public health data system should operate when a new threat emerges. This figure articulates the features of the transformed public health data system that can move from routine reporting to absorb new insights. This includes being able to link new data with routine health information, integrating insights from the new threat into the system to inform action, and then determining how to sustain any new data in an adapted system going forward. This figure outlines a much more dynamic public health system, able to update (e.g., like computer software) and evolve.

Figure 3. A transformed public health data system must be agile, responsive, and adaptive to emerging conditions



Principle 3: The public health data system must support agency among a diversity of actors and sectors.

A modern, equity-oriented public health data system should mobilize the full range of actors and sectors that influence health outcomes, help with the frame of health priorities, and help to catalyze action. Given the array of influences on health from all sectors, having a public health data system that can be used by a wide range of sectors is important. In the Why? paper on equity orientation in the public health data system, we note that the current public health data system is neither oriented enough towards end use, nor structured in a way that ensures equitable access and use of data to amplify the voice and leadership of those who are historically underrepresented in public health decision-making.

Much of public health data is stored by the public health department and/or health care organizations and is not as fully available or accessible for the commons, meaning the diversity of stakeholders comprising the broader public health system, as defined earlier. This lack of accessibility can prohibit the fullest expression of health equity, or the fair and just access to the opportunity to be healthy, because without data, it is difficult for constituents to act in ways that will advance health equity, as noted in Principle 1.

We know that a sense of agency, or the “feeling of generating and controlling actions in order to influence events in the outside world,” is key in other aspects of health behavior and decision-making, but agency has not always been a factor in how we design and build the public health data system.¹⁵ It is problematic to have a data system that does not support that ability by all actors and sectors to take data and transform into *information* and *insight*, as well as use those data to

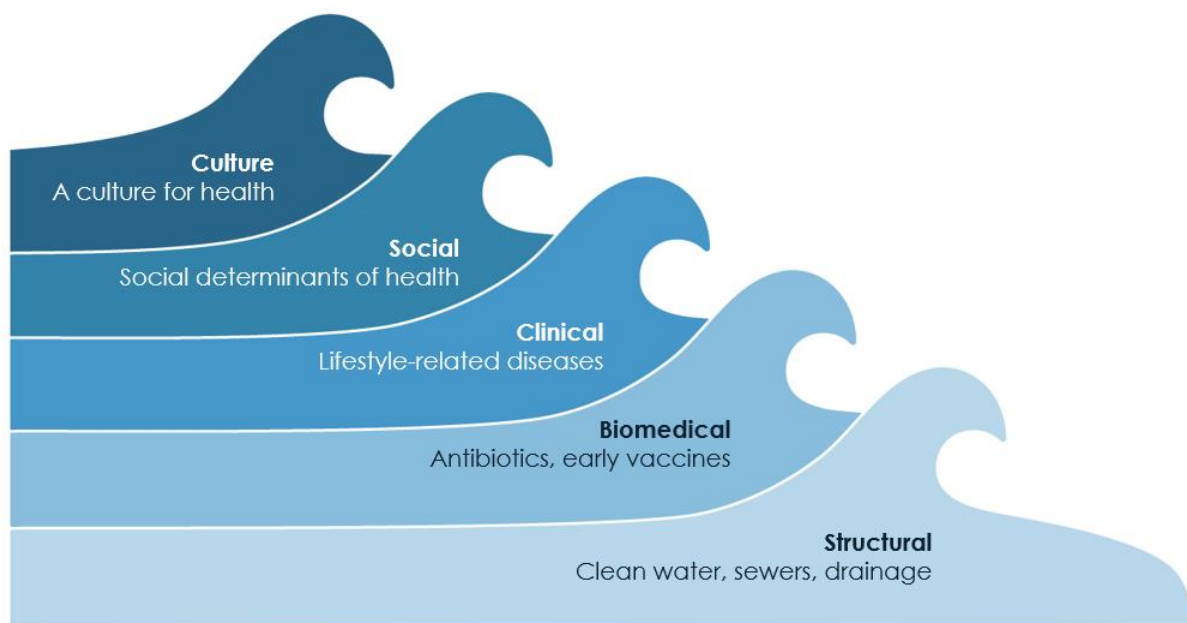
influence one's environments and functioning. The modern, equity-oriented public health data system needs to consider its design, features, and other characteristics that set the conditions to support agency across diverse sectors. For instance, a public health data system should ensure open access where possible, provide supporting materials to engage with the data in meaningful ways, link to policy and programmatic solution options to aid immediate public health action, and have a design that invites access and usability.

Principle 4: The public health data system needs to consider all that drives public health change across the five waves of public health.

In a 2014 article, Davies et al. articulated the importance of the public health system moving towards a culture for health, shaped by shared values to advance a healthy society. This would be the next wave of public health's actions, after the increasing push to recognize the influence of social determinants of health (**Figure 4**).¹⁶ This framework outlined the key chapters of public health evolution from its roots in hygiene and structural changes. The framework then described how public health improved the health of the populations through biomedical and clinical advancements, and how public health established linkages between factors outside of traditional health settings, such as transportation and the economy, and health outcomes. But Davies and others noted that in order for there to be true mobilization of sectors to advance health and to create consistent conditions for a healthy community, public health also needed to focus on the cultural variables that prioritize health in narrative and decision-making.

In the context of the public health data system, these five waves or chapters in public health must be represented not only in the choice of data and what data are being monitored, but also in terms of how clinical, environmental, and social variables of health are integrated for a fuller picture of the health and well-being of the nation and of communities. Including measures that represent these five public health waves help to situate the action and impact functions of the system, as noted in Principle 1. A modern public health data system that prioritizes biomedical outcomes with less focus on social determinants of health misses the core influences on health. While some data sets have braided these two types of data, and collection of social determinants of health is now more common, it is still rare for public health data to factor in additional contextual factors that support health (i.e., that capture a culture for health), such as civic support of health actions, community mindset around health, and investments in health. The RWJF Culture of Health data have advanced some of this holistic thinking about health and well-being, but these ideas have not yet been fully integrated in the next generation of the public health data system.¹⁷

Figure 4. Public health data system should reflect aspects of public health stages



Adapted from Davies, 2014

Principle 5: The public health data system must ensure that data represent a mix of leading and lagging indicators, as well as a balance of data on positive health and well-being.

In addition to the completeness of what is tracked in the public health data system, and the equity orientation, the modern public health data system must balance measures that capture foresight (or indicators of what is occurring or emerging in terms of health in the United States) with hindsight (or indicators that capture performance and can help evaluate the state of health in the United States).¹⁸ To date, the public health data system has not always balanced leading (predicting future state) and lagging indicators (performance today or what has occurred). As such, the system does not always have the information needed for early insight on health trends and conditions, which may worsen down the line, in ways that could inform anticipatory action. The design of indicators must take into account decision and action timelines.

Further, the current public health data system disproportionately focuses on negative health outcomes (i.e., mortality and morbidity), with far less attention on positive health and well-being measures (e.g., hope, civic engagement, prosocial behaviors). While Healthy People 2030 embraces more focus on individual well-being, the broader public health data system has not fully integrated these variables as part of consistent health surveillance. Without such information, the nation cannot track aspects of thriving and flourishing, which are central to health and related social and economic outcomes.¹⁹ Further, the lack of these data inhibits strengths-based policies and interventions that could be enacted to advance health, well-being, and health equity.

Principle 6: The public health data system must not simply describe health phenomenon and trends, but also capture the processes by which health outcomes have occurred.

Related to Principle 4, the public health data system has not been rooted enough in capturing the full extent of factors that drive health, from structural and clinical to social and cultural, and it also can be argued that the public health data system does not embrace a full equity orientation. That is, the current public health system does not fully track historical, systemic, and structural factors that inform the “causes of the causes,” or explain the processes by which health inequities have occurred or persisted. Without those data, it is very difficult to translate data into *information* and *insight* with an equity grounding, and to develop policy and programmatic action that considers context and systemic reform. The COVID-19 pandemic highlighted the inability of the current public health data system to capture systematic information about the pandemic experience for subgroups of Americans, as well as the inability to assess upstream factors, particularly systemic racism, which exacerbated the health, social, and economic impacts of COVID-19.

A future public health data system must include space for monitoring the processes of inequity, history, cumulative risk, and cumulative trauma. It is important to capture health impacts by subgroup in a data equity strategy, but it is not enough.²⁰ The modern public health data system must go both upstream and look at intergenerational inequities. For instance, as noted with research on diseases of despair, the public health data system had not been able to fully assess growing hopelessness in American communities in ways that also could get ahead of the problem.²¹ Work on community allostatic load notes that the current public health data system has not captured historical conditions that contribute to cumulative stress (e.g., histories of discrimination) or can intensify acute threats (e.g., community violence, natural disaster) when they occur.²²

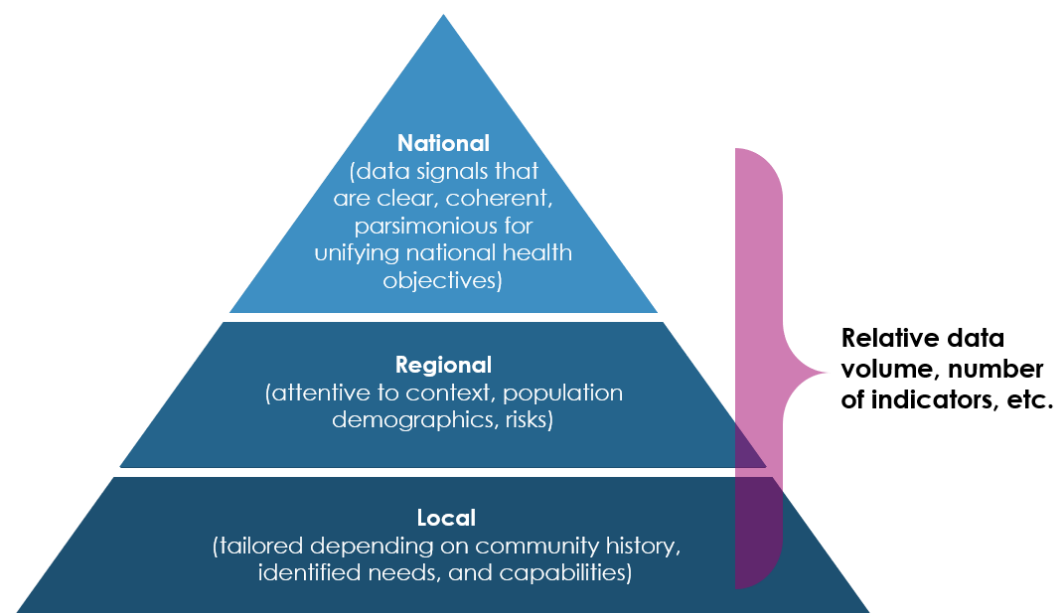
Principle 7: The public health data system needs to balance comprehensiveness and parsimony in data.

Another core principle of the modern, equity-oriented public health data system is the need to offer data that can signal response and action, while not overwhelming the user with so many data points that the system becomes unusable. On the other hand, while parsimony can be valued, it also is important that the public health data system reflects the multidimensionality of health (e.g., physical, social, emotional, civic, economic) and the blend of upstream drivers and downstream outcomes of health. This tension of comprehensiveness versus parsimony must be considered in the future public health data system.

This set of tradeoffs is not new but becomes a challenge of increasing complexity when factoring in the growing volume and variety of data available to inform health as well as increasing understanding of the myriad of factors that influence health (**Figure 1**). Researchers and practitioners are continuing to explore the balance of “measure what matters” versus a growth in standard quality of care measurement, for instance, working to differentiate data that crisply and succinctly demonstrates health value externally versus more detailed data that can support internal operational improvements.²³ The characteristic of *multiresolution*, or the ability to zoom in and out to capture data and information resolution, also becomes an important feature of the public health data system. With the explosion of real-time data, data coming from the Internet of Bodies (an extension of the Internet of Things that connects the human body through a network of devices), and other sources of data, the parsimony question looms even larger.²⁴ How much data is too much data? Which data can be converted into useable measures? Which of those data have the signal that is required for *information* and *insight*?

Another issue relevant to this principle is aligning the balance of parsimony and comprehensiveness with the needs of various levels (e.g., national, regional, local) of public health. If we consider the national, regional, and local capture of public health data, questions emerge regarding how much data is enough for a national picture of health versus data to support localized decision-making? How much data are manageable at each level? **Figure 5** offers one potential frame for how these levels of data and data volume could be considered in a tiered, national public health data system. In this **Figure 5**, we assert that the national data picture should be the most streamlined to support a simple and shared American narrative and a unifying and inspiring health objective. These points are further explored in the Why? and What? papers.

Figure 5. Structure of who holds public health data should be tiered with consideration of parsimony



The Remaining Papers

As outlined in **Table 1**, the remaining papers go deeper on the issues raised in the public health data system principles. The papers are meant to outline a set of considerations emerging from literature and stakeholders, offering insights and examples from efforts underway in the United States. Taken together as a set, the papers should prime the deliberations of the National Commission.

Appendix A. Key public health data modernization and innovation efforts, including a sample of CDC linked efforts on modernization and health equity

Initiative	Brief Description and Website*
3D Commission	The 3D Commission brings together research and stakeholders to advance a common language between the fields of data science, health determinants, and policy and practice decision-making. https://3dcommission.health/about
Data Across Sectors for Health (DASH)	DASH works to build local capacity, an evidence base, and a movement towards multi-sector data-sharing between housing, healthcare, education, public safety, economic development, behavioral health, and other sectors. https://dashconnect.org/about-dash/
Google-Apple Exposure Notification Framework	Google and Apple created a system to support contact tracing. Once users opt in, the system generates a random ID for each smartphone device, which changes every 10-20 minutes to address privacy concerns. Phones exchange IDs via Bluetooth and periodically check all the random IDs associated with positive COVID-19 cases. Note that this is a temporary effort focused on COVID-19. https://www.google.com/covid19/exposurenotifications/
Gravity	Gravity convenes broad stakeholder groups in identifying and harmonizing social risk factor data with regards to food insecurity, housing instability and quality, and transportation access for interoperable electronic health information exchange. https://www.hl7.org/gravity/
HITECH and Meaningful Use	The Health Information Technology for Economic and Clinical Health (HITECH) Act established the meaningful use of interoperable EHR as a national goal. The meaningful use concept was based on 5 pillars: <i>Improving quality, safety, efficiency, and reducing health disparities</i> <i>Engage patients and families in their health</i> <i>Improve care coordination</i> <i>Improve population and public health</i> <i>Ensure adequate privacy and security protection for personal health information</i> https://www.cdc.gov/ehrmeaningfuluse/introduction.html
HL7 PH Working Group	HL7 standards have been developed to serve the needs of clinical care of individual patients. The Public Health Work Group supports the HL7 mission to

	<p>create and promote its standards by helping to assure that HL7 V2 and V3 models, messages, documents, and services address the requirements of the many public health agencies, both governmental and non-governmental.</p> <p>https://www.hl7.org/Special/committees/phcr/overview.cfm</p>
Leading Edge Acceleration Projects (LEAP)	<p>LEAP funds innovative projects to advance interoperable health IT, including the creation of new standards, methods, and tools, such as application programming interfaces (APIs) and mobile applications (apps), to improve care delivery and advance research capabilities.</p> <p>https://www.healthit.gov/buzz-blog/health-innovation/leap-in-health-it-program-overview-2020-and-2019-awards</p>
Public Health Community Platform	<p>This Public Health Community Platform is the platform that resulted in the Digital Bridge (see next) and established the concept of bidirectional data flow and knowledge management to build a cloud-based system connecting EHRs to public health data for a select initial set of notifiable conditions.</p> <p>https://www.astho.org/Programs/Informatics/Public-Health-Community-Platform/Factsheet/</p>
Reportable Condition Knowledge Management System	<p>The Reportable Condition Knowledge Management System is a portal to enhance real-time disease surveillance. It includes an interface for public health agencies to input, manage, and edit jurisdictional data; a database repository; and a decision support service (DSS) that healthcare reporters can use to determine if a potential case is reportable, and to which jurisdiction(s).</p> <p>https://www.cste.org/group/RCKMS</p>
Various CDC Originated Reporting Systems	
APHL Informatics Messaging Service (AIMS)	<p>AIMS is a cloud-based messaging platform to assist in interoperability, security, visualization, hosting, and exchange of public health data between the CDC, regional labs and hospitals, State Health Information Exchanges, and 50 state public health jurisdictions. Examples of data currently exchanged through AIMS include: Aggregated Influenza test result data from public health laboratories to CDC, immunization data exchange among several public health jurisdictions, Electronic Case Reporting (eCR) between providers and jurisdictions across the United States, etc.</p> <p>https://www.aphl.org/programs/informatics/pages/aims_platform.aspx</p>
BioSense	<p>BioSense was launched in early 2000s to establish a national health surveillance system to combat bioterrorism-related illnesses. BioSense has since grown to include alerts for disease outbreaks and other hazardous events. Although it started as a CDC-based program, state and local health departments now use it to send data to the CDC and other partners. De-identified health data, including chief complaint, diagnosis codes, patient characteristics, and location are integrated into a shared platform which the public health community can access to conduct syndromic surveillance to monitor public health.</p> <p>https://www.cdc.gov/nssp/how-sys.html</p>
CDC Data Modernization Initiative	<p>The first comprehensive strategy to modernize data, technology, and workforce capabilities to support public health surveillance, research, and decision-making.</p> <p><i>Addresses the following questions:</i></p> <p><i>How do we improve timeliness and quality of data?</i></p> <p><i>How can we better coordinate data activities and systems?</i></p> <p><i>How can we reduce burden on data partners?</i></p> <p><i>How can we integrate emerging technologies more effectively?</i></p> <p>https://www.cdc.gov/surveillance/surveillance-data-strategies/data-IT-transformation.html</p>
eCR Now	<p>A strategic initiative that allows for rapid adoption and implementation of eCR for COVID-19. eCR (Electronic Case Reporting) is the automated generation and</p>

	transmission of case reports from the electronic health record (EHR) to public health agencies for review and action. https://www.cdc.gov/coronavirus/2019-ncov/hcp/electronic-case-reporting.html
Immunization Gateway (IZ)	The Immunization Gateway is a portfolio of components that share a common IT infrastructure. These components support the exchange of immunization data between immunization information systems (IISs), provider organizations, and consumer applications. The IZ Gateway can streamline time- and resource-intensive data exchange onboarding. It also replaces multiple one-to-one connections with centralized routing. https://www.cdc.gov/vaccines/covid-19/reporting/iz-gateway/overview.html
Immunization Information Systems (IIS)	In the IIS, people receive vaccinations from a number of sources, the records are sent to the IIS, and the IIS provides records to patients and authorized professionals. https://www.cdc.gov/vaccines/programs/iis/downloads/basics-immun-info-sys-iis-508.pdf
Public Health Information Network (PHIN)	PHIN includes tools and resources to help increase the capacity of public health agencies to electronically exchange health data and information, including a standardized vocabulary system, a messaging system between CDC and public health agencies, and a repository of information about organizations and jurisdictions important to public health programs. PHIN provides Object Identifiers (OIDs) for use within the public health community. https://www.cdc.gov/phinf/
Syndromic Surveillance	This is CDC's system of syndromic surveillance as reported by EHRs and collected via interoperable secure messaging systems for Meaningful Use such as AIMS. https://www.cdc.gov/nssp/overview.html
Example CDC Health Equity Initiatives	
CDC Indicators of Health Equity (CIHEP)	This effort is led by the Office of Minority Health and Health Equity, in collaboration with leaders from other Centers, Institutes, and Offices (CIOs). The goal is to develop a parsimonious set of indicators as well as analytic studies and supporting products and help the CDC focus on health outcomes and determinants in a way that will drive action to reduce health disparities/inequities and advance health equity.
National Center for Chronic Disease Prevention and Health Promotion (NCCDPHP) Social Determinants of Health (SDOH) Measurement Work Group	The goal of this effort is to identify existing and/or develop new programmatic measures to guide NCCDPHP's work in the SDOH Measurement framework foundational domains and indicators: Poverty; Employment, Education, Housing, Neighborhood & Environment; Nutrition Security; Access to Healthcare; and Social Support.

**The description comes verbatim or slightly paraphrased from website information. This table is only meant to summarize basic descriptions of initiatives with no assessment of impact.*

Acknowledgements

Each paper in this series of white papers has been reviewed by Mahshid Abir of the RAND Corporation and the University of Michigan and Karen Smith of Public Health Strategies. We appreciate these reviewers for their insights and contributions. We also thank Burness for their editorial and communication support for each paper.

References

1. *Congressional Briefing: Ending the Triple Pandemic: Advancing Racial Equity by Promoting Health, Economic Opportunity and Criminal Justice Reform*, Online: Dialogue 4 Health, 2020. <https://www.tfah.org/story/ending-the-triple-pandemic-congressional-briefing/>
2. Meyer R, Madrigal AC, "Why the Pandemic Experts Failed: We're still thinking about pandemic data in the wrong ways.," *The Atlantic*, 2021. <https://www.theatlantic.com/science/archive/2021/03/americas-coronavirus-catastrophe-began-with-data/618287/>
3. Biden JR, *Executive Order On Advancing Racial Equity and Support for Underserved Communities Through the Federal Government*, Washington, D. C., House, The White, The White House, 2021. <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/20/executive-order-advancing-racial-equity-and-support-for-underserved-communities-through-the-federal-government/>
4. Braveman P, Arkin E, Orleans T, Proctor D, Plough A, *What Is Health Equity? And What Difference Does a Definition Make?*, Princeton, N. J.: Robert Wood Johnson Foundation, 2017.
5. Koesten L, Kathleen G, Groth P, Simperl E, "Talking datasets – Understanding data sensemaking behaviours," *International Journal of Human-Computer Studies*, Vol. 146, 2021, p. 102562.
6. Stansfield SK, Walsh J, Prata N, Evans T, Jamison DT, Bremen JG, Measham AR, Alleyne G, Claeson M, Evans DB, Jha P, Mills A, Musgrove P, "Information to Improve Decision Making for Health," *Disease Control Priorities in Developing Countries*, Washington, D. C.: The International Bank for Reconstruction and Development / The World Bank, 2006. <https://pubmed.ncbi.nlm.nih.gov/21250312/>
7. Klein G, Phillips JK, Rall EL, and Peluso DA, "A data-frame theory of sensemaking," in Hoffman, R. R., ed., *Expertise out of context: Proceedings of the Sixth International Conference on Naturalistic Decision Making*, New York, N. Y.: Lawrence Erlbaum Associates Publishers, 2007, pp. 113-155.
8. Bertrand M, Marsh JA, "Teachers' Sensemaking of Data and Implications for Equity," *American Educational Research Journal*, Vol. 52, No. 5, 2015, pp. 861-893.
9. Ferguson G, Mathur S, Shah B, "Evolving from information to insight," *MIT Sloan Management Review*, Vol. 46, No. 2, 2005, p. 51.
10. Wohlstetter P, Datnow A, Park V, "Creating a system for data-driven decision-making: Applying the principal-agent framework," *School effectiveness and school improvement*, Vol. 19, No. 3, 2008, pp. 239-259.
11. Milovic B, Milovic M, "Prediction and decision making in health care using data mining," *International Journal of Public Health Science*, Vol. 1, No. 2, 2012, pp. 69-78.
12. Wu H, LaRue EM, "Linking the health data system in the U.S.: Challenges to the benefits," *International Journal of Nursing Sciences*, Vol. 4, No. 4, 2017, pp. 410-417.
13. Kubo T, Yanasan A, Herbosa T, Buddh N, Fernando F, and Kayano R, "Health Data Collection Before, During and After Emergencies and Disasters-The Result of the Kobe Expert Meeting," *Int J Environ Res Public Health*, Vol. 16, No. 5, 2019, p. 893.
14. Aung E, and Whittaker M, "Preparing routine health information systems for immediate health responses to disasters," *Health Policy and Planning*, Vol. 28, No. 5, 2013, pp. 495-507.
15. Moore JW, Fletcher PC, "Sense of agency in health and disease: a review of cue integration approaches," *Consciousness and Cognition*, Vol. 21, No. 1, 2012, pp. 59-68.
16. Davies SC, Winpenny E, Ball S, Fowler T, Rubin J, and Nolte E, "For debate: a new wave in public health improvement," *Lancet*, Vol. 384, No. 9957, 2014, pp. 1889-1895.
17. RWJF, "Building A Culture of Health," 2021. As of March 15: <https://www.rwjf.org/en/cultureofhealth.html/en/cultureofhealth.html>

18. Gray P, Hovav A, "From hindsight to foresight: Applying futures research techniques in information systems," *Communications of the Association for Information Systems*, Vol. 22, No. 1, 2008, p. 12.
19. Richmond CAM, Ross NA, Egeland GM, "Social support and thriving health: a new approach to understanding the health of indigenous Canadians," *American Journal of Public Health*, Vol. 97, No. 10, 2007, pp. 1827-1833.
20. Team, IHI Multimedia, "Create the Data Infrastructure to Improve Health Equity," 2019. As of March 15: <http://www.ihl.org/communities/blogs/create-the-data-infrastructure-to-improve-health-equity>
21. Brignone E, George DR, Sinoway L, Katz C, Sauder C, Murray A, Gladden R, Kraschnewski JL, "Trends in the diagnosis of diseases of despair in the United States, 2009–2018: a retrospective cohort study," *BMJ Open*, Vol. 10, No. 10, 2020, p. e037679.
22. Chandra A, Cahill M, Yeung D, Ross R, *Toward an Initial Conceptual Framework to Assess Community Allostatic Load: Early Themes from Literature Review and Community Analyses on the Role of Cumulative Community Stress*, Santa Monica, Calif.: RAND Corporation, 2018. https://www.rand.org/pubs/research_reports/RR2559.html.
23. Meyer GS, Nelson EC, Pryor DB, James B, Swensen SJ, Kaplan GS, Weissberg JI, Bisognano M, Yates GR, Hunt GC, "More quality measures versus measuring what matters: a call for balance and parsimony," *BMJ Qual Saf*, Vol. 21, No. 11, 2012, pp. 964-968.
24. Lee M, Boudreaux B, Chaturvedi R, Romanosky S, Downing B, *The Internet of Bodies: Opportunities, Risks, and Governance*, Santa Monica, Calif.: RAND Corporation, 2020. https://www.rand.org/pubs/research_reports/RR3226.html

Transforming Public Health Data Systems

Why? The focus on equity in
the modern public health
data system

This white paper is part of a series of papers for use by the National Commission to Transform Public Health Data Systems, an independent commission convened by the Robert Wood Johnson Foundation (RWJF). These papers were used to inform Commission discussions and recommendations. The papers represent insights from literature review, a diverse set of stakeholder perspectives, and inputs from RWJF grantee projects related to the topic of the Commission. The white papers were developed in support of the Commission by the RAND Corporation, a nonprofit research organization, within its division, RAND Social and Economic Well-Being. Corresponding authors are [Anita Chandra](#) and [Laurie Martin](#).

**National Commission to
Transform Public Health
Data Systems**

May 2021

Table of Contents

Introduction	2
Equity Purpose of Public Health Data and the Public Health Data System.....	3
Applying an Equity Lens from Data Collection Through End Use and Decision-making.....	5
1. Types of equity	6
2. The current state of equity in public health data collection and analysis	10
3. Equity and the role of power, voice, and agency in the public health data system	16
Conclusion.....	19
Acknowledgements	19
References	19

Introduction

Too often, in data modernization efforts, there is a disproportionate focus on form and sometimes function but less on an overriding mission or purpose. Without a central purpose around which a system can coalesce, it becomes difficult to make choices about data inputs, information outputs, and how insights are gleaned for action and social change (see Figure 2, preamble). In stakeholder discussions, there were significant concerns that the United States has no single goal for health that is coherent and unifying. Some cited Gross Domestic Product (GDP) as an exemplar to be adapted for health; even with its limitations, “[GDP] functions as a comprehensive scorecard of a given country’s economic health.”¹ Further, as explained in the preamble (Principle 1), a wide range of stakeholders with interests in a modern public health data system collectively asserted the need for the future public health data system to be more than a collection of numbers but to support both sensemaking and decision-making functions.

As described in the preamble, equity and equity orientation should not be a feature but must be the primary purpose of the modern public health data system. These Commission white papers utilize the Robert Wood Johnson Foundation (RWJF) definition of health equity (and its broader equity implications) to be: *Health equity means that everyone has a fair and just opportunity to be as healthy as possible. This requires removing obstacles to health such as poverty, discrimination, and their consequences, including powerlessness and lack of access to good jobs with fair pay, quality education and housing, safe environments, and health care.*² However, it is unclear how equity should center in the modern public health data system, what it means to have an equity orientation, and why starting with equity offers a new roadmap for configuring a public health data system that is distinct from what we currently have.

This white paper sets the stage for the purpose of the next public health data system, the philosophical and pragmatic underpinnings of holding equity centeredness, and what this means for measurement, data collection, data analysis and end use. Themes of equity thread through the remaining papers in further detail associated with the content of public health data (What?); the design of the public health data system (How?); and the implementation of the public health data system (Who and What Next?). As noted in the preamble, this paper is informed by stakeholder input, literature review, and grantee insights. The examples and exhibits included in this paper are meant to be exemplar rather than exhaustive.

Equity Purpose of Public Health Data and the Public Health Data System

In their seminal blog on public health 3.0, DeSalvo and Benjamin (2016) rightly asserted that public health leaders must serve as the chief health strategists for the community; that public health departments must be strengthened, including through more robust cross-sector collaboration; that funding for public health must be enhanced and modified to link the social determinants of health; and that there is a strong need for timely, reliable and granular public health data that can guide work on prevention and equity.³ While these authors along with many others have crystallized important points about the future of public health, including improving public health data and taking advantage of new forms of technologies to augment data, what is notably missing is a well-understood, singular, national health goal that centers equity. *For instance, should the public health data system “moonshot,” as referenced in the preamble, emphasize a nation that is actively dismantling systemic health injustice?* The public health field has been hampered by criticisms that it lacks a clear and embraceable purpose, and that improving health and well-being is not galvanizing enough or too broad to be understood. Could something different work better when centering equity?

Take, for instance, the goal of Healthy People 2030 (HP2030) to improve the health and well-being of people in the United States, with underlying objectives pointing to the need to reduce disparities and achieve health equity as part of this goal. These kinds of objectives are not uncommon in public health, and their importance has certainly been underscored by recent discussions elevating health inequities as untenable, particularly in light of COVID-19.⁴⁻⁶ While the Healthy People objectives and others like them are laudable, there are perhaps two challenges with national goals like this that may be helpful as the Department of Health and Human Services (HHS) Secretary Advisory Committee on HP2030 and other national efforts roll forward.⁷ First, in the context of data system principles, such as supporting sensemaking and decision-making (Principle 1, preamble) and balancing data parsimony (Principle 7, preamble), this goal can read vague for national action and shared motivation. Second, this kind of goal does not fully pinpoint health equity as central. For instance, questions loom with respect to how much equity is truly a throughline of these national scorecards, such that equity is understood as **both** process and outcome; whether the components of equity such as procedural and distributive equity can be assessed; and the extent to which history, cumulative stresses and structures are actually monitored as part of the public health data system.

When applying an *equity lens*^a to the purpose of public health data and the public health data system, significant questions emerged from both literature and stakeholder inputs regarding whether equity is honestly reflected in the public health data system’s organizing mission (**Table 1**). For instance, in the context of COVID-19 vaccine discussions, the concept of equitable access had to newly consider whether the age thresholds for early access should be benched to

^a An **equity lens** is defined in these white papers as any approach for analyzing the equity impact of actions on people or places, with attention to those historically marginalized, as well as the process of addressing any structural or systemic barriers to the opportunity to be healthy (*adapted from multiple sources*).

differential life expectancies by racial/ethnic groups (i.e., earlier access to vaccines for groups with lower life expectancies).⁸ If equity had been centered in public health data systems in the first place, would this debate about what constitutes equitable access have ensued?

Table 1. Equity as a centering purpose in public health data: Stakeholder identified issues/questions

Sample of typical public health objectives	Sample questions about equity within objectives
Improving health and well-being	What does <i>fair access to opportunities to be healthy</i> mean in the context of the multiple dimensions of well-being (e.g., social, environmental, civic) given systemic barriers present in those dimensions?
Reducing disease burden	How should burden be calculated given differential life experience and cumulative risks?
Reducing health costs	How do different historical influences on health factor into health cost calculations?
Improving responsiveness to emerging infectious disease/syndromic surveillance	How should inequitable capacities of local health systems factor into what data are tracked to aid disease response decisions?

Applying an Equity Lens from Data Collection Through End Use and Decision-making

Exhibit 1 presents proposed equity principles for an equity-oriented public health data system developed by the Joint Workgroup convened by the CDC Foundation, and for use by Centers, Institutes, and Offices (CIOs) within the Centers for Disease Control and Prevention (CDC).⁹ As noted in **Exhibit 1**, these principles align well with other themes raised in this paper and the **System Principles** for the broader public health data system in the preamble. The CDC example demonstrates how one part of the public health data system may approach equity orientation going forward.

Exhibit 1: Draft principles for equity orientation in public health data systems for use by CDC Centers, Institutes, and Offices (CIOs)

The CDC Foundation, working with partners in CDC, have proposed principles for equity-oriented data collection and use across the CDC. The principles offer a framework of equity measures to be built into existing and potential new cross-agency data collection, analysis, and dissemination, and developing actionable recommendations for how CDC and the CDC Foundation could promote cross-agency actions to transform data use to drive health equity.⁹ These draft principles include:

1. **Community inclusion:** Proactively seek to include communities' interest in design considerations when collecting, analyzing, using, and sharing the results of CDC's data.
2. **Limited burden:** Ensure that data collection and use policies minimize undue burden for the affected populations and do not exceed the community gain.
3. **Equitable access:** A multi-tiered approach to data access should first begin with a clear delineation of practical and legal data availability. Any open data should carefully consider a range of perspectives and its release and access should be granted to a diverse range of stakeholders.
4. **Disaggregation:** Data needs to be disaggregated by several categories including, but not limited to, race, ethnicity, age, sexual orientation, and gender.
5. **Transparency:** The data should be publicly available and shared from a trusted and official source.

The next sections go further by describing different dimensions of equity and how these must be considered in a modern, equity-oriented public health data system across sectors and actors. This includes how certain types of equity are integrated; how equity currently factors into measurement choices and approaches to data collection; and how equity and issues of power, voice, and agency are part of design and end use decisions. Each section offers key considerations for the public health data system.

1. Types of equity

1.1 Key Action: A modern public health data system should be more intentional about factoring in different types of equity to truly make progress on health equity.

There are several types of equity, but rarely have these aspects of equity been explicitly acknowledged in the choice of public health data and the design of the public health data system. Taking a page from ecosystem services (or the human benefits provided by the natural environment), we use this framework (**Figure 1**) to outline key equity considerations for the content of the modern public health data system. Equity orientation in the public health data system must outline how the parameters of equity are set (e.g., who is included in decision-making, who are the targets of the effort in terms of *for whom* to improve equity); why equity is the focus (versus a “do no harm” model, for instance, that does not consider differential population needs and histories); who counts as the subject of equity and how are generational and historical considerations taken into account; and then given that, what is the content of the equity (i.e., what counts as a matter of equity). This includes:

- Procedural equity: the perceived fairness of processes and procedures to make decisions
- Distributive equity: how social welfare and need is balanced
- Contextual equity: how pre-existing social conditions influence equity

Figure 1. An equity framework

Adapted from McDermott, M., Mahanty, S. and Schreckenberger, K. (2011) ‘Defining Equity: A framework for evaluating equity in the context of ecosystem services’ available at <http://redd-net.org/resource-library/defining-equity-a-framework-for-evaluating-equity-in-the-co>.



1.1a. Consideration: A lack of procedural equity, or the perceived fairness of approaches and procedures to make decisions, has hampered a sense of trust between the public and public health leaders.

Procedural equity addresses how the concept of fairness is included in approaches and policies once the equity parameters are set (i.e., for whom are we improving equity). People must believe in the justness of the system, with attention to trust, ethics, voice, and participation. The issue of trust in public health is well-documented as influencing everything from adherence to medication regimes and behavioral change interventions to the uptake of health risk communications.¹⁰ At the start of the COVID-19 pandemic, the public generally rated their trust in public health leaders as high, but as the pandemic continued, the nation's trust in public health declined due, in part, to political influences and concerns about data gaps and misinformation.¹¹ We know that transparency and engagement of the public in decision-making can be helpful antidotes to trust concerns, particularly in the times of uncertainty that a pandemic or another emergent threat presents.¹²

However, procedural equity is not simply about trust. This type of equity includes components such as representation, as well as voice and participation in decision-making. In the context of public health data, these components are meaningful, yet not systematically pursued. Representation and inclusion have become even more resonant during COVID-19, stemming from a lack of granularity in key demographic categories, such as race/ethnicity or disability status, hampering an ability to respond effectively to the needs of certain subgroups. In a blog on data and health justice, Nancy Krieger notes to “ensure that COVID-19 work is grounded in health justice, we must generate and publicly report data on how it affects different populations and social groups and use a health equity lens to examine how the pandemic is exacerbating inequities.”¹³

The issue of voice and participation in decision-making is a persistent problem in public health, and the public health data system is no exception. By **voice**,¹⁴ we reference the inclusion of the perspectives, ideas, and lived experiences of those impacted by public health decisions. By **participation**,¹⁵ we reference the active role of those impacted by public health decisions in the actual process of arraying decision options and providing meaningful input on those decisions.

There is growing evidence about the benefits of having people involved in health decision-making because the health field confronts difficult trade-offs of resources and values, and values are often driven by ethics (see **Exhibit 2**) and social and cultural experiences.¹⁵ And yet, despite this recognition, full public participation in public health decision-making is not common. This problem has been particularly acute among those groups that have been historically marginalized and disproportionately affected by health issues and negative health exposures.¹⁶ If there were full voice and participation in decision-making in the modern public health data system, this would include meaningful inputs on measures selection, how data are collected, how data are disaggregated, how data are represented, and how information and insight are drawn from the data (Figure 2, preamble).¹⁷ Ideally, with more attention to procedural equity, data or digital colonialism, concerns that the private sector in particular is harvesting data without reciprocal public benefit could be less of an issue.¹⁸

Exhibit 2: Ethics and empathy, race/ethnicity, and public health data

The Urban Institute has been leading a landscape scan of the role of ethics and empathy in public health data.¹⁹ Data disaggregation can offer insight about systemic racism and patterns of oppression, but because data are often not presented this way or able to be disaggregated as constructed, “researchers are forced to choose between using imprecise methods to estimate race (such as the predominant race of the individual’s zip code) or forgoing disaggregation altogether.”¹⁹ While there are methodological innovations around imputation that can help with this disaggregation, there are areas of ethical risk that must be considered. Through this landscape scan including stakeholder input and literature review, the Urban research team is already identifying key insights and gaps. This includes:

- **Gaps in the methods and ethics literature.** There is not enough literature on imputing race and ethnicity and on practical methods for incorporating community engagement principles in this analytic work.
- **Gap between stakeholders’ motivating philosophy and approach.** Equity-driven stakeholders identify a need for disaggregated data, but do not consistently see imputation as the most viable or useful approach doing so, even when some researchers do.
- **Strong concerns about accuracy, representativeness, and visibility.** Good data are required to support analytic model development and to ensure accuracy, but Urban has identified concerns about the quality of existing data for these purposes and whether imputation can accurately represent smaller, more integrated, or more dispersed populations (such as American Indians, Alaska Natives, Asians, Native Hawaiians, and Pacific Islanders).
- **Community variation makes establishing guardrails difficult.** Establishing guardrails for using community data from vulnerable populations is complex, but ethical standards also must be tailored to the specific needs of the communities implicated in the research and the data.
- **Structure and the social and historic construction of race.** Imputed race allows for understanding of disparities but does not fully address the fact that race is a social and historical construction. It is critical that research and policy lift up the correlated structural factors when reporting on race.

1.1b. Consideration: *Limited attention to distributive equity, or how social welfare and need is balanced, makes it difficult for the public health data system to fully support sense-making and decision-making functions.*

The concept of distributive equity focuses on allocation and resource management decisions, with attention to the balance of costs, risks, and benefits. Usually, distributive equity considers how decisions are made, and how benefits are distributed based on dimensions of need and social benefit.

It is unclear how much the concepts of distributive equity are embedded in the design of the current public health data system, but as the modern public health data system grapples with how to center equity and how to support forward-leaning public health action, it is useful to examine which data are used and how data are arrayed to inform allocation decisions. To date, many of the public health data dashboards organize information into some combination of dimensions of clinical outcomes, the social and economic environment, and risk behaviors. But most datasets and data platforms do not organize data in ways that align measures and indicators in a distributive equity framework. For instance, this can mean organizing data to cluster indicators about the proximal drivers of a particular health outcome (e.g., insurance status, quality of health services, prenatal care, discrimination) with the health outcome (e.g., maternal mortality). This allows decision-makers to consider the investments needed in all of those drivers together to influence an outcome and what is then possible given resource constraints. The next public health data system could offer information and ultimately insights that help with tradeoffs for realizing concepts, such as targeted universalism, and efforts to address historical inequities.

1.1c. Consideration: *Contextual equity, which considers how pre-existing social conditions and generational impacts influence equity, has not been fully embedded yet in the public health data system.*

Contextual equity is the backdrop of both procedural and distributive equity because this form of equity accounts for the political, economic, social, and intergenerational factors in which populations engage with society, its systems, and its benefits. This includes contextual variables such as access (e.g., access to capital) and power (e.g., in this context, the ability to gain and maintain access to resources).²⁰ **Exhibit 3** offers an example of how vaccination rates and vaccination inequities, with respect to access, can be better understood in the context of social determinants of health.

Despite its importance, in the context of the public health data system, contextual equity has tended to receive less attention both in the content and type of data, which data are tracked, and how those data are translated into public health action. It is rare that variables such as the accumulation of risk exposures, the legacy of injustice, and systemic barriers are factored into how measures of health are calculated or how data are used to justify or explain certain types of public health action. As some communities pursue a greater focus on applying an equity lens to local budgets and policymaking (e.g., Tacoma, Louisville), standing up Offices of Equity, and/or appointing Chief Equity Officers, the public health data system needs to meet that momentum by considering how data aligns with new policy frameworks. For instance, like the Drexel example in **Exhibit 3**, the California

Equity Index was used to drive equitable response and vaccine distribution, by linking the Healthy Places index and other information with vaccination allocations.²³

Without intentional pursuit of information about pre-existing inequities, it is difficult to take typical public health data and use it in a way that can lead to equitable action, particularly over time and across generations. As such, transforming public health data must include consideration of how the translation of data into information and insight has a historical explanation included in the interpretation.

Exhibit 3: Tracking outcomes and contextual inequities in the COVID-19 pandemic

The Drexel Urban Health Collaborative has been conducting a series of projects on the COVID-19 pandemic and tracking inequities, including characterizing inequities in various COVID-19 related outcomes, such as incidence, hospitalizations, mortality, and testing and vaccinations in Big Cities Health Coalition (BCHC) cities; identifying city, neighborhood and other factors associated with health and health inequities in COVID-19 outcomes across cities; and evaluating public health policies aimed at minimizing the impact of the COVID-19 pandemic and assessing policy effects on COVID-19 inequities.^{21,22} The Collaborative has developed this [tool](#) to link factors such as spatial heterogeneity and COVID-19 outcomes. The equity framework used to drive analyses includes inequities between populations (e.g., occupational differences), between neighborhoods (e.g., redlining policies), and between cities (e.g., economic and social conditions).

2. The current state of equity in public health data collection and analysis

2.1 Key Action: Given the multiple dimensions of equity, revisit how equity is presented in current equity-based datasets.

Over the last few years, there has been a growth in the number of U.S. communities promoting equity indicators. This progress represents, in many ways, a new recognition of the importance of equity. As public health pursued biomedical advancements and greater understanding of the social determinants of health (see Figure 4, preamble), the interest in closing gaps and addressing health *disparities*, grew exponentially. In Healthy People 2020, disparities were defined as:

“... a particular type of health difference that is closely linked with economic, social, or environmental disadvantage. Health disparities adversely affect groups of people who have systematically experienced greater social or economic obstacles to health based on their racial or ethnic group, religion, socioeconomic status, gender, age, or mental health; cognitive, sensory, or physical disability; sexual orientation or gender identity; geographic location; or other characteristics historically linked to discrimination or exclusion.”²⁴

This disparities definition began to include the language of systems and obstacles. But as the nation approaches the expectations of Healthy People 2030, the broader definition of health equity (noted on page 1) has been embraced to not bypass the value of addressing health disparities, but to capture more of the factors that serve as the “causes of the causes,” or the upstream, systemic, and historical factors that continue to drive those health disparities. Further, health equity is often intertwined with the concepts of health justice, meaning that health must be paired with an understanding of ethics and basic human rights.²⁵ Health justice also includes analysis of jurisprudential and legislative actions that influence health outcomes.

2.1a. Consideration: *Current equity indicators frameworks from fields outside of health may help to inform health equity going forward.*

A brief review of current equity indicators efforts is illuminating, both in terms of the progress being made to capture a wider array of factors that influence health outcomes, but also in terms of how much those indicators still miss some aspects of health justice and specifically procedural, distributive, and contextual equity.

Other equity indicator efforts outside of health provide important templates for potential evolution in health equity indicators. For instance, the National Academy of Sciences recently produced a guidebook on building educational equity indicators for states and school districts. **Table 2** highlights key features of educational equity indicators, which should attend to the structural and systemic sources of inequity (e.g., item 4).

Table 2. Features of educational equity indicators²⁶

Crucial characteristics of educational equity indicator systems
1. Measure multiple dimensions of educational outcomes and opportunities.
2. Align with students’ developmental trajectories and with schooling levels that intersect with important developmental transitions.
3. Measure disparities among salient, well-defined population groups.
4. Measure contextual and structural disparities that educational systems must confront and counteract.
5. Include measures that are comparable across time and place.
6. Measure equity at multiple geographic and organizational scales (classrooms, schools, districts, states).
7. Produce frequent, readily interpretable, high-level summary statistics, in addition to more nuanced statistics.
8. Rely on credible evidence about the validity and reliability of the measures.
9. Include a plan for continuous research and improvement to reflect evolving education and child developmental sciences.

The National Equity Atlas also includes a multidimensional approach to equity, capturing variables in these areas²⁷:

- Demographics (e.g., racial generation gap)
- Economic vitality (e.g., income growth)
- Readiness (e.g., disconnected youth)
- Connectedness (e.g., housing burden)
- Economic benefits (e.g., racial equity in income)

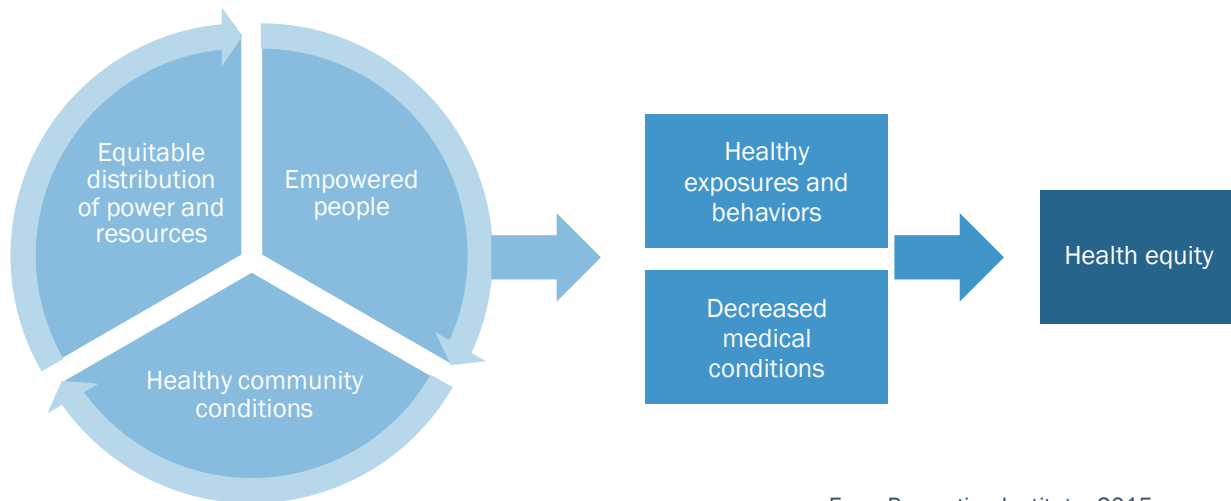
The Atlas offers some indicators that speak to notions of intergenerational equity, such as the racial generation gap (or how much investment is happening in youth of color). This measure can aid considerations of distributive and contextual equity.

The Equality Indicators effort²⁸, promoted by the City University of New York Institute for State and Local Governance, provides insight that can be useful for health equity. Six cities have incorporated the Equality Indicators framework, though most cities have transitioned *Equality* into *Equity* Indicators frameworks. For instance, Dallas includes dimensions of Economic Opportunity, Education, Housing and Neighborhood, Justice and Government, Public Health, and Transportation and Infrastructure. However, when drilling into the public health dimension, there are limits in how much the measures selected actually speak to health equity and not simply disparities in health outcomes. Access to health care, risk behaviors, population health, and maternal and child health variables are included, but there is very little attention to notions of fairness or justice, or the relative distribution of health assets for subgroups. As such, while equity indicators like Dallas' make important progress in sorting public health data against social, economic, and environmental factors such as justice and government, there are still limitations in the extent to which the health data itself use a full equity lens.

2.1b. Consideration: *Health and well-being indicators are identifying equity but are still limited in expressions of equity dimensions.*

As noted in fields outside of health explicitly, there have been important moves in embracing equity concepts in measurement philosophy and in the selection of some indicators. In 2015, the Prevention Institute put forth a useful analysis of the trajectories of *health* equity, as noted in **Figure 2.**²⁹

Figure 2. Trajectories of health equity



From Prevention Institute, 2015

Further, efforts like the HOPE initiative of the National Collaborative for Health Equity, represent tremendous advancement in how the myriad of structural drivers of health outcomes is considered.³⁰ However, when reviewing current health equity measurement and health equity indicators efforts, it is unclear how much the *equitable distribution of power and resources* and *empowered people* are actually being measured, despite its importance for procedural equity. Further, given the importance of historical and intergenerational factors in considerations of distributive and contextual equity, the current health equity datasets do not quite capture those factors either.

Take, for example, the Rhode Island Health Equity Indicators project (**Table 3**). The effort is significant in terms of how it braids social, economic, and environmental variables with traditional measures of health care, and it is emblematic of many state and local efforts in health equity. It is notable that the indicators include aspects of community resiliency and community trauma. But, like many other measurement dashboards of its type, there is somewhat less attention to the full range of factors that may inform procedural or distributive equity (e.g., over time measures, cumulative risk, historical determinants), nor are data presented in a way that can more directly guide equity-based public health action.

In designing the modern, equity-oriented public health data system, it is important to consider both how the full expression of equity is represented in current datasets and how data are combined to support that core data system principle (Principle 1) of both sense-making and decision-making. To be sure, the Rhode Island effort includes a significant community engagement process, and understandably, indicator efforts are often hampered by choosing from data that are currently collected. But this example serves an important point with respect to how complete (or incomplete) health equity measurement in public health may be situated and how much farther the field may need to proceed to realize full expressions of equity.

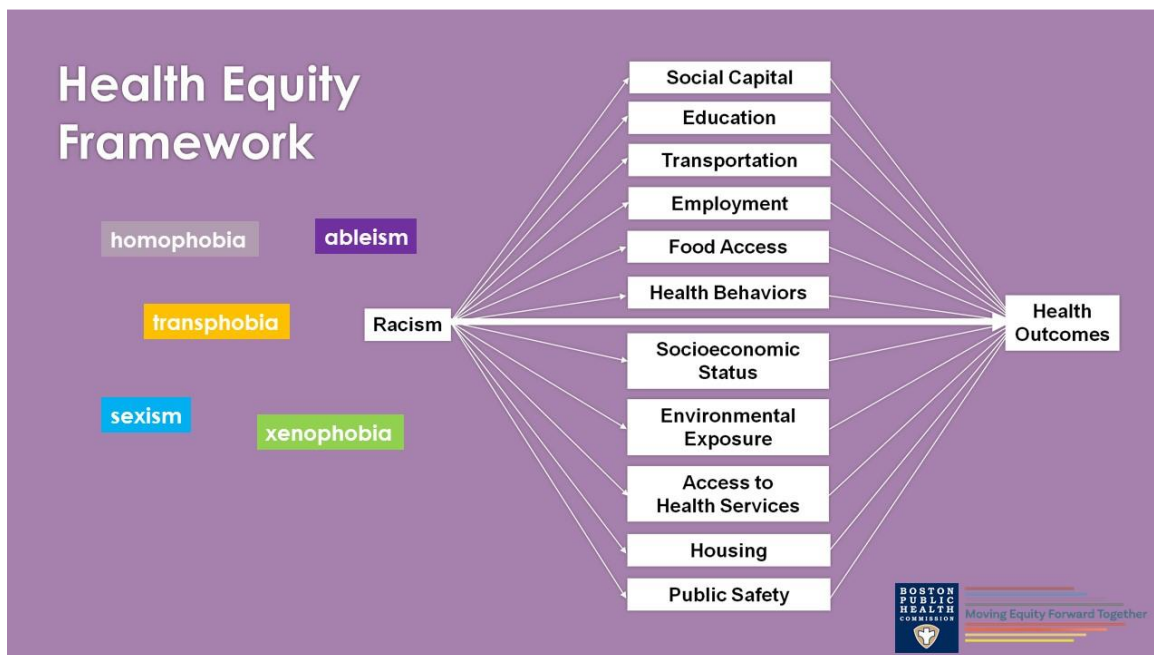
Table 3. Rhode Island health equity measures³¹

Domain	Determinant	Measure	Data Source
Integrated Healthcare	Healthcare Access	Percentage of adults who reported not seeking medical care or dental care due to cost (2 measures)	Behavioral Risk Factor Surveillance System (BRFSS)
	Social Services	Ratio: Number of individuals receiving to number of individuals eligible for SNAP benefits, based on income	Supplemental Nutrition Assistance Program (SNAP), U.S. Census Bureau
	Behavioral Health	Ratio: Number of naloxone kits distributed to number of overdose deaths	RIDOH, Prevent Overdose RI website
Community Resiliency	Civic Engagement	Percentage of registered voters participating in the most recent presidential election	Rhode Island Board of Elections
	Social Vulnerability	Index score that reflects the social vulnerability of communities	Centers for Disease Control and Prevention (CDC) Social Vulnerability Index, Agency for Toxic Substances and Disease Registry (ATSDR)
	Equity in Policy	Ratio: Number of low- to moderate-income housing units to number of low- to moderate-income households	HousingWorks RI, Comprehensive Housing Affordability Strategy
Physical Environment	Natural Environment	Percentage of overall landmass with tree canopy cover	U.S. Department of Agriculture (USDA) Forest Service i-Tree Tools
	Transportation	Index score that reflects the affordability of transportation for renters	U.S. Department of Housing and Urban Development (HUD) Low-Cost Transportation Index
	Environmental Hazards	Number and percentage of children with blood lead levels higher than 5 micrograms per deciliter	RIDOH Environmental Lead Program

Socioeconomics	Housing Cost Burden	Percentage of cost-burdened renters and owners	American Community Survey
	Food Insecurity	Percentage of population who are food insecure	Feeding America
	Education	Percentage of high school students graduating with a regular diploma within four years	Rhode Island Department of Education
Community Trauma	Discrimination	Percentage of adults reporting racial discrimination in healthcare settings in the past 12 months	BRFSS (available 2020)
	Criminal Justice	Number of non-violent offenders under Rhode Island probation and parole (per 1,000 residents age 18 and older)	Rhode Island Department of Corrections, U.S. Census Bureau
	Public Safety	Violent crime rate and non-violent crime rate (per 100,000 people)	Rhode Island State Police Uniform Crime Reports, FBI Uniform Crime Reporting Program

The Boston Public Health Commission Racial Justice and Health Equity effort importantly does go farther in naming the influences of “isms” like racism and sexism in their framework (**Figure 3**) of health equity. Questions remain about how measurement of those factors will be presented, returning to the idea that data is not enough but must translated into useable *information* and contextualized into data-informed *insight* (see preamble).

Figure 3. Boston Health Equity Framework



From Boston Public Health Commission, 2020

3. Equity and the role of power, voice, and agency in the public health data system

3.1 Key Action: The public health data system needs to be actively oriented to facilitate agency by actors and sectors.

As noted in the description of procedural equity, the issue of both representation and power in health decision-making should be a key part of a modern, equity-oriented public health data system. Despite important progress in participation and accountability (e.g., community boards) in some communities, there is significant work ahead. By **power**,³² we mean the authority to shape expectations, decisions, and outcomes in public health. Truth, reconciliation, and racial healing may need to be more central to the actions of public health in designing the next public health data system.³³ The current public health data system is not usually characterized by transparency, support of data access and use that facilitates sustained civic engagement, and deep consideration of the voices of historically marginalized and chronically underrepresented populations. Yet, a core principle of the future public health data system (Principle 3 in the preamble) is to support a sense of agency to use and act upon public health data going forward. This focus on equity in data use and decision-making is even more critical in the context of the greater volume, velocity, and variety of data (often used to characterize ‘big data’), the role of technology in data generation and manipulation (e.g., artificial intelligence, machine learning), and increasing challenges to civic engagement and democracy.

3.1a. Consideration: *The notion of “nothing about us without us” is not fully part of the public health data system yet, but it must be a core tenet.*

In a perspective piece about the role of artificial intelligence (AI), Kalluri wrote, “don’t ask if AI is good or fair, ask it if shifts power.”³⁴ This sentiment targeting AI was focused on ensuring that the tool does not exacerbate inequities through problems such as algorithmic bias, but this theme also can be widened and attributed to all of the components of a modern public health data system. In discussions with stakeholders about the future of public health data, there was an overriding consensus that the power in the public health data system was not oriented properly. In short, stakeholders expressed concerns that have been compounding for years: Data were not representative of community needs and those most impacted by what the data revealed were not in the discussions early and consistently to help drive public health action. Further, data were not “righting” social inequities but rather potentially worsening those inequities by aligning data in ways that do not support equity-promoting policy solutions.

In the context of new forms of data or ‘big data,’ these issues of participation and representation are even more resonant.³⁵ Key ethical issues meriting more comprehensive engagement include fair distribution of benefits and burdens, control and sharing of data, and accountability. For instance, people participation is needed to determine if the benefits of these ‘big data’ accrue to corporations while the burden is actually borne by individual citizens. How private companies hold information as business assets versus supporting community benefit is a key consideration. Further, how data are used for good, whose good, and then how data can be used for harm also is a question of accountability to whom and by whom.³⁶

Given this context, key tenets of community voice emerged from both the literature review and stakeholder inputs (**Table 4**).

Table 4. Proposed tenets of community voice and leadership in the modern public health data system

Tenets of community voice and leadership
Representation from a range of community leaders (e.g., community-based organization leaders, advocates) across sectors (e.g., housing, economic development, environment) is needed for public health data system design, content, and use decisions.
Public health data should be collected in ways that balance intrusion and burden with benefit, with clear goals for those data, which are well-understood by the public
Public health data must be able to be disaggregated in ways that make sense for community context and needs.
Public participation is needed to effectively translate data into information and insight.
Community inputs should be used to inform the array of data and how data are organized for equity-based decisions, including all of the dimensions of equity.
Community leadership is needed to create public health narratives with the data, in ways that can promote proactive public health action.

3.1b. Consideration: *The discussion of the public health data system comes at a time of civil unrest, concerns about democracy, and civic engagement, which could be harnessed for advancing the modern public health data system.*

One of the benefits of this period for advancing a modern, equity-oriented public health data system is the level of civic activation that exists in the United States today. From prior reviews of civic engagement and health, we know that more effort is needed to activate health-related civic engagement (e.g., advocating for health policies) in the country.³⁷ But, the timing of national calls for racial justice and better health response to the pandemic, as well as growing civic action on behalf of addressing widening social and economic inequities, creates an important foundation upon which the efforts of this Commission are well-positioned.

There are important lessons from civic participation science that should be integrated into the equity orientation of the next public health data system. The Kirwan Institute outlines several parameters of equitable and inclusive civic engagement, which can be used for the public health data system discussions:³⁸

- **Embracing the Gifts of Diversity** – *harnessing social capital*
- **Realizing the Role of Race, Power, and Injustice** – *examining power imbalances*
- **Radical Hospitality: Invitation and Listening** – *using intentional inclusion*
- **Trust-Building and Commitment** – *incorporating mutual accountability*
- **Honoring Dissent and Embracing Protest** – *creating conditions that surface and consider conflicts*
- **Adaptability to Community Change** – *openness to changing civic structures*

To date, it can be argued, few of these parameters are included in the design and content of the public health data system. Yet, more purposeful integration of these ideas could attend to the three dimensions of equity-procedural, distributive, and contextual more completely.

Conclusion

The term *equity* has in many ways been the “term du jour,” and equity in health has been no exception. While there has been some progress in how equity is showing up in public health and in public health data, getting from equity as a feature to equity as a center and orientation in the modern public health data system requires a few more steps. First and foremost, the dimensions of equity, including procedural, distributive, and contextual, must be considered in choice, structure, presentation, and use of data. It is quite clear that the tenets of equity and the notion of “causes of the causes” are not yet the throughline of the public health data system, as attention to data on systems, structures, and histories is limited at best. Further, there has been some consideration of equity in data access and use and its ties to meaningful public health action, but there is no standard or consistency in how public health data systems at national, regional, and local levels step beyond simple community engagement to fully representative stakeholder leadership on data choices and decisions.

Acknowledgements

Each paper in this series of white papers has been reviewed by Mahshid Abir of the RAND Corporation and the University of Michigan and Karen Smith of Public Health Strategies. We appreciate these reviewers for their insights and contributions. We also thank Burness for their editorial and communication support for each paper.

References

1. Fernando J. Gross Domestic Product (GDP). Investopedia. Economic Indicators Web site. <https://www.investopedia.com/terms/g/gdp.asp>. Published 2020. Accessed March 30, 2021.
2. Braveman P, Arkin E, Orleans T, Proctor D, Plough A. *What is Health Equity?* Princeton, N.J.: Robert Wood Johnson Foundation; 2017.
3. DeSalvo K, Benjamin G. Public Health 3.0: A Blueprint For The Future Of Public Health. In. *Health Affairs Blog*. Vol 2020; 2016.
4. CDC. Paving the Road to Health Equity. Centers for Disease Control and Prevention. https://www.cdc.gov/minorityhealth/publications/health_equity/index.html. Published 2020. Accessed April 4, 2021.
5. Dover DC, Belon AP. The health equity measurement framework: a comprehensive model to measure social inequities in health. *Int J Equity Health*. 2049;18(36).
6. Improvement IfH. A Measurement Framework For Health Equity. Health Leads. Published 2019. Accessed April 4, 2021.
7. 2020 HP. Secretary's Advisory Committee on National Health Promotion and Disease Prevention Objectives for 2030. Office of Disease Prevention and Health Promotion. <https://www.healthypeople.gov/2020/About-Healthy-People/Development-Healthy-People-2030/Advisory-Committee>. Published 2020. Accessed April 4, 2021.
8. Persad G, Peek ME, Emanuel EJ. Fairly Prioritizing Groups for Access to COVID-19 Vaccines. *JAMA*. 2020;324(16):1601-1602.
9. Smith L. *Project Status Report: Transforming Public Health Data Systems at the Centers for Disease Control and Prevention to Advance Health Equity*. Atlanta, GA: CDC Foundation; 2021.

10. Shore DA. Communicating in times of uncertainty: the need for trust. *Journal of Health Communication*. 2003;8(S1):13-14.
11. Udow-Phillips M, Lantz PM. Trust in Public Health Is Essential Amid the COVID-19 Pandemic. *Journal of Hospital Medicine*. 2020;15(7):431-433.
12. Frewer L. Risk perception, social trust, and public participation in strategic decision making: Implications for emerging technologies. *Ambio*. 1999;569-574.
13. Krieger N. COVID-19, Data, and Health Justice. In: *The Commonwealth Fund*. Vol 2021: The Commonwealth Fund; 2020.
14. Casebeer A, Mackean G, Abelson J, Maloff B, Musto R, Forest P-G. Engaging the Public Voice in Health Care Decision-Making. In: McKee L, Ferlie E, Hyde P, eds. *Organizing and Reorganizing. Organizational Behaviour in Health Care*. London: Palgrave Macmillan; 2008.
15. Church J, Saunders D, Wanke M, Pong R, Spooner C, Dorgan M. Citizen participation in health decision-making: past experience and future prospects. *Journal of Public Health Policy*. 2002;23(1):12-32.
16. Wiggins A, Wilbanks J. The Rise of Citizen Science in Health and Biomedical Research. *The American Journal of Bioethics*. 2019;19(8):3-14.
17. Boyd RW, Lindo EG, Weeks LD, McLemore MR. On Racism: A New Standard For Publishing On Racial Health Inequities. In: *Health Affairs Blog*. Vol 2020: Health Affairs; 2020.
18. Couldry N, Mejias UA. Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject. *Television & New Media*. 2019;20(4):336-349.
19. Institute TU. *Preliminary Findings: Ethics and Empathy Landscape Scan [White Paper]*. Santa Monica, CA: RAND Corporation; 2021.
20. McDermott M, Mahanty S, Schreckenber K. *Defining Equity: A framework for evaluating equity in the context of ecosystem services*. REDD-net; 2011.
21. Collaborative DUH. *Tracking equity issues in COVID-19 testing and vaccination access in selected US cities*. Princeton, N. J.: Robert Wood Johnson Foundation; 2021.
22. Collaborative DUH. *Tracking outcomes and inequalities of the COVID-19 pandemic across the 30 US cities that participate in the BCHC*. Philadelphia, P.A.: Robert Wood Johnson Foundation; 2021.
23. Tracking COVID-19 in CA. California State Government. <https://covid19.ca.gov/>. Published 2021. Accessed April 5, 2021.
24. Disparities. HealthyPeople.gov. Foundation Health Measures Archive Web site. <https://www.healthypeople.gov/2020/about/foundation-health-measures/Disparities>. Published 2020. Accessed 2021.
25. Benfer EA. Health Justice: A Framework (and Call to Action) for the Elimination of Health Inequity and Social Injustice. *American University Law Review*. 2015;65(2):275-351.
26. Nielsen N. Developing a System of Equity Indicators: What States, Districts, and their Partners can do. In: *Building Educational Equity Indicator Systems: A Guidebook for States and School Districts*. Washington, D. C.: National Academies Press; 2020.
27. National Equity Atlas. National Equity Atlas. <https://nationalequityatlas.org>. Published 2021. Accessed April 5, 2021.
28. Governance CUoNYIFSL. Equality Indicators. City University of New York Institute for State & Local Governance. <https://islg.cuny.edu/sites/our-work/equality-indicators/>. Published 2021. Accessed April 5, 2021.
29. Davis R. *Measuring What Works to Achieve Health Equity: Metrics for the Determinants of Health*. Oakland, Calif.: Prevention Institute; 2015.
30. Initiative TH. Measures to Advance Health and Opportunity. The HOPE Initiative. <https://www.hopeinitiative.org/>. Published undated. Accessed April 4, 2021.
31. Ausura C. Rhode Island Health Equity Measures. Rhode Island Department of Health. <https://health.ri.gov/data/healthequity/>. Published 2021. Accessed 2021.
32. Gore R, Parker R. Analysing power and politics in health policies and systems. *Global Public Health*. 2019;14(4):481-488.
33. Collaborative N. Truth, Racial Healing & Transformation (TRHT) with Dr. Gail C. Christopher. YouTube2021.
34. Kalluri P. Don't ask if artificial intelligence is good or fair, ask how it shifts power. *Nature*. 2020;583(169).
35. Vayena E, Madoff L. Navigating the Ethics of Big Data in Public Health. In: Mastroianni AC, Kahn JP, Kass NE, eds. *The Oxford Handbook of Public Health Ethics*. New York, N. Y.: Oxford University Press; 2019.
36. Mittelstadt BD, Floridi L. The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts. *Science and Engineering Ethics*. 2015;22(2):303-341.
37. Nelson C, Sloan J, Chandra A. *Examining Civic Engagement Links to Health: Findings from the Literature and Implications for a Culture of Health*. Santa Monica, Calif.: RAND Corporation; 2019.
38. Holley K. *The Principles for Equitable and Inclusive Civic Engagement*. Columbus, OH: Kirwan Institute for the Study of Race and Ethnicity; 2016.

Transforming Public Health Data Systems

What? The data in the
modern public health data
system

This white paper is part of a series of papers for use by the National Commission to Transform Public Health Data Systems, an independent commission convened by the Robert Wood Johnson Foundation (RWJF). These papers were used to inform Commission discussions and recommendations. The papers represent insights from literature review, a diverse set of stakeholder perspectives, and inputs from RWJF grantee projects related to the topic of the Commission. The white papers were developed in support of the Commission by the RAND Corporation, a nonprofit research organization, within its division, RAND Social and Economic Well-Being. Corresponding authors are [Anita Chandra](#) and [Laurie Martin](#).

**National Commission to
Transform Public Health
Data Systems**

May 2021

Table of Contents

Introduction	2
Reimagining Public Health Data.....	3
1. Approach to public health data.....	3
2. Content of public health data	10
3. Volume and coordination of public health data.....	18
4. Precision, granularity, and timeliness of public health data.....	22
Conclusion.....	26
Appendix A. Gaps in Social Determinants of Health Data from Varied National Public Health Data Sources.....	28
Acknowledgements	31
References	31

Introduction

Deteriorating and outdated infrastructure (technological and physical) and a shrinking public health workforce, further exacerbated by the events of 2020 and beyond, have resulted in insufficiencies in the availability of public health data, particularly to identify and track populations affected by multiple stresses, such as COVID-19 and chronic disease. The opportunity described in the preamble offers a chance to reimagine what matters for health in the United States; how the nation uses data as a tool to aid action on inequity; and how the United States uses public health data to capture concepts, such as systemic health injustice, thriving, and resilience.

The purpose of this white paper is to review what data are needed in the modern public health data system. As mentioned in the guiding system principles in the preamble, public health data needs to be more than a collection of numbers; it needs to have an impact and sensemaking function, balance comprehensiveness and parsimony, and place equity at the forefront. Advancing the culture of health, one that is guided by the cultural and social drivers of health, will require a radically different approach to public health data collection and use, including how to capture concepts like civic participation and positive health, among other concepts. To date, public health data have not led public health action, in large part because there is no integrated or coordinated public health data system to support sensemaking and data-informed decision-making.

Reimagining Public Health Data

Given this context, this paper describes the **approach to collecting and reporting** public health data; the **content** of public health data; the **volume and coordination** of data across sectors; and the level of **precision, granularity, and timeliness** of data currently reflected in public health data sources. Each of these four sections includes a brief discussion of gaps and limitations, as well as key actions and considerations to inform a more forward leaning, coordinated, and integrated public health data system that can not only support, but help to modernize public health.¹ Special attention is paid to opportunities to leverage emerging innovations and technology to improve the existing state of data. As noted in the preamble, this paper is informed by stakeholder input, literature review, and grantee insights. The examples and exhibits included in this paper are meant to be exemplar, rather than exhaustive.

1. Approach to public health data

1.1 Key Action: Ensure the public health data system promotes national, regional, and local data alignment across a more parsimonious set of core public health measures, and with attention to consistent use of new forms of public health data.

The general *approach* (or the philosophy about data content and data collaborations) to public health data has been stagnant due to a fragmented and misaligned public health system. Rather than being used to drive changes in public health over time, the data system has been somewhat reactive. This section discusses the general approach to public health data, in accordance with the data system principles defined in the preamble, to promote data that can inform proactive decision-making (Principle 1), emphasize alignment across levels of decision-making (Principle 3), value parsimony in the measures selected (Principle 7), and leverage new forms of data from technology companies and associated sectors (Principles 5 and 6). Some of this data system reactivity may be a result of the lack of a crisp national health vision, a point described in the prior paper. This lack of a focused and clear health vision can result in behaviors that simply add more measures to capture a broad, national health goal of improving health and well-being, without enough attention to what data provide the best and most focused signals for public health action.

Efforts to modernize the public health data system will require prioritizing a smaller set of core national measures to ensure the parsimony needed to proactively make social change on a few key public health efforts, rather than continue to spread the primary focus of public health thin across many areas. For example, if addressing systemic health injustice becomes the priority, then the primary areas of focus and associated measures and indicators should emphasize upstream drivers of inequity before those inequities become disparities. If the priority becomes supporting the positive health of the American people, then the data system must elevate measures and accompanying data that will support that vision. However, the current use of the public health data system in the United States does not proportionately weight measures toward any clearly operationalized priority, whether positive health, addressing inequity, or other forward-leading public health action. As such,

the current public health data system does not fully communicate a common set of values through an alignment and parsimony of national measures.

Although such parsimony may be useful at a national level, it does not preclude the development of regional and local data priorities that can support action. Figure 5 in the preamble provides a notional schematic for these three potential levels of data (national, regional, and local). Having a simplified set of core national measures would not prevent local health departments from augmenting that core data set with data that reflects local needs, particularly when considering local context and end use. On the other hand, adhering to alignment and parsimony at the national level can also help with clarity at the local level (**Exhibit 1**).

Exhibit 1: Leveraging Community Information Exchanges and 2-1-1 data for equitable and inclusive public health data systems during a time of pandemic²⁻⁴

Community Information Exchanges (CIEs) aim to be more proactive and inclusive by promoting a community-led approach to collecting accurate and comprehensive data from individuals bearing most of the health burdens. 2-1-1 San Diego and the CIE National Community Council are working together to develop and advance a national CIE strategic agenda that would promote alignment with other multisector data sharing initiatives and cultivate a culture of health equity through meaningful systems change. Interviews with seven communities early in the planning or implementation stage of this CIE is helping to define the data system equity drivers along 11 dimensions (e.g., informed consent and refusal, security, data stewardship/collection/use) and generate recommendations for next steps including a CIE accreditation process, a policy advocate to highlight the need for community-based organizations' involvement in community solutions, and a call for dedicated or future funding that requires community engagement and a reflection of lived experience in the project plan.

2-1-1 also has played a critical public health role during COVID-19, facilitating testing, enabling vaccine distribution, expanding food delivery by coordinating with food pantries and DoorDash, establishing an Economic Impact Payment Line, and establishing relief funds for impacted individuals and communities. Catalyzed by their efforts in response to COVID-19, **United Way Worldwide** is in the process of developing a **National 2-1-1 Data Platform** to map and aggregate local 2-1-1 data and provide insights about community needs and resources in near real-time.²⁻⁵

1.1a. Consideration: *The existing elements of the public health data system tend to be reactive to emerging health challenges or threats, rather than designed to proactively drive public health action.*

The main function of public health data has remained consistent over the years, primarily focused on population health surveillance, one of the 10 essential public health services.^{a,6} But, public health data do not simply serve the purpose of passive surveillance. Data are used to support other public health services including the selection of public health actions, communication with the public, and the building of partnerships to improve health. Despite these varied and strategic uses of public health data, the content of the data has been augmented over time in ways that are not consistent with proactive support of those key public health services, but rather in ways that are reactive in monitoring health conditions. This often includes the disproportionate use of lagging indicators (meaning data on performance today or what *has* occurred), rather than leading indicators that portend future health needs.

Data systems have been enhanced, for example, as the nation has learned more about key factors that influence health. *Healthy People*, the national strategic management plan that guides health promotion and disease prevention efforts of federal, state, and local public health organizations, and their community partners, demonstrates the steady expansion of public health's primary areas of focus over the past 40 years (**Table 1**).⁷ This focus has expanded from a singular emphasis on mortality to now include aspects of quality of life, health disparities and health equity, and social and physical environments. Yet, despite the expansion in these focus areas, the data and indicators that are often a part of public health data emphasize infectious disease, chronic conditions, mortality, and risk factor exposures, with less attention to other factors that influence health over generations (e.g., trauma), as well as measures of positive health and well-being and systemic inequities.

This reactive process has not only resulted in a large set of public health objectives and indicators, but it sometimes has confused those outside public health as to the boundaries and purpose of public health. Despite the fact that public health has been underfunded and declining in capacity over the past 50 years (see the following sections), the primary areas of measurement focus continue to grow, broadening a scope without commensurate support. The current Healthy People 2030 shows some promise towards this need for parsimony, but this set of measures is still too large to promote focused national health action.

^a Ten essential public health services: 1. Assess and monitor population health status, factors that influence health and community needs and assets. 2. Investigate, diagnose, and address health problems and hazards affecting the populations. 3. Communicate effectively to inform and educate people about health, factors that influence it, and how to improve it. 4. Strengthen, support, and mobilize communities and partnerships to improve health. 5. Create, champion, and implement policies, plans, and laws that impact health. 6. Utilize legal and regulatory actions designed to improve and protect the public's health. 7. Assure an effective system that enables equitable access to the individual services and care they need to be healthy. 8. Build and support a diverse and skilled public health workforce. 9. Improve and innovate public health functions through ongoing evaluation, research, and continuous quality improvement. 10. Build and maintain a strong organizational infrastructure for public health.

Table 1. Healthy People goals, priority areas, and objectives/targets over the past 40 years

Year	1990	2000	2010	2020	2030
Goals	-Decrease deaths throughout the lifespan -Increase independence of older adults	-Increase the span of healthy life -Reduce health disparities -Achieve access to preventive services for all	-Increase quality and years of healthy life -Eliminate health disparities	-Attain high-quality, longer lives free of preventable disease, disability, injury, and premature death -Achieve health equity, eliminate disparities, and improve the health of all groups -Create social and physical environments that promote good health for all -Promote quality of life, healthy development, and healthy behaviors across all life stages	-Attain healthy, thriving lives and well-being free of preventable disease, disability, injury, and premature death -Eliminate health disparities, achieve health equity, and attain health literacy to improve the health and well-being of all -Create social, physical, and economic environments that promote attaining the full potential for health and well-being for all -Promote healthy development, healthy behaviors, and well-being across all life stages -Engage leadership, key constituents, and the public across multiple sectors to take action and design policies that improve the health and well-being of all
Priority areas	15	22	28	42	23
# of objectives /targets	266	319	467	1,318	355

*Healthy People 2030 uses the term outcomes instead of goals.

Note: Healthy People 2030 also contains development and research objectives that, when added to the priority areas and core objectives listed in this table, would increase these numbers.

1.1b. Consideration: *Fragmentation of federal, state, and local funding and reporting requirements have contributed to the lack of crosscutting and aligned public health data.*

Federal agencies use population- or prevalence-based formula grants or competitive grants to award public health funds. How state and local funds are disbursed varies by state and is often based on health department activities and governance structures. County and city revenues contribute to public health activities but vary widely by geographic location. A common requirement of these funds is documentation of how funds were spent, either in terms of processes (e.g., number of cancer screenings) or outcomes (e.g., reduction in late-stage cancer diagnosis). In addition to government funds, resources from foundations or other donors often have other data requirements. This fragmentation, both in levels of funding and sometimes additive yet not converging data requirements, contributes to limited coordination, a lack of data parsimony, and poor data alignment.

Public health funding is often earmarked for program areas that align with traditional public health programs (e.g., maternal and child health, HIV/AIDS), which can create silos by specific health issues.⁸ Similar to the aforementioned prevalence-based formula grants, these silos have driven the content of public health data. Instead of focusing on the health of the population, public health data have focused on prevalence of specific issues, such as how many people are suffering from opioid addiction. These data have been useful in defining the scope of the problems and identifying hazards, but consequently, there is less focus on solutions or drivers of positive health and well-being.

Another challenge that leads to fragmentation is an overemphasis on transactional data (e.g., received an immunization, screening, provider visit); such data captures an interaction between individuals and the public health system and can include data from electronic health records, clinical and financial data systems, and enterprise resource planning systems. However, data on public health *action* or *implementation* are different from *transactional data* in that these implementation data not only track the transactions involved with the delivery of public health services, but also the quality of these services or the enforcement of a specific health-related policy. In theory, data on public health action or implementation are useful because they provide a time-based snapshot of public health services and policies that can be used to forecast the future need for these public health services or adjustments to public health policy and can be used to segment those forecasts by different population characteristics (e.g., race, gender, age, income, homelessness). However, in practice, much of the data on public health action or implementation do not contain detailed information on the characteristics of the population that are affected by the services or policies. Thus, the types of groups by which such data can be segmented is often limited, as described in Section 4 below.

1.1c. Consideration: *A wealth of data exists that could be leveraged to inform public health action, but public health does not consistently or readily have access to those data.*

Many companies are examining the data they collect and own and are seeking to repurpose those data for public good.^{9,10} For instance, person-level, transactional data exist on almost every aspect of our lives and are available to private industry. Technologies, such as GPS and accelerometry data from wearable devices, provide insight into bodily movements and location. Purchasing and travel activity provide insight on personal values, socioeconomic status, and healthy behaviors. Online habits, such as web searches or browsing histories, can illuminate what people want to know about health, as well as offer sensitive markers of health status.¹¹ Because of network effects or first mover effects (i.e., the benefits accrued from being the first to market), much of the data for a given technology type is monopolized by a single dominant company in that domain. Technology companies that seek to leverage their own products and holdings for social good are largely doing it on their own—that is, looking inward to their own data and tools and focusing on their own capabilities. One way in which some companies (e.g., *Twitter*) have tried to do this is by making subsets of their data openly available for researchers or government use, under specific data use or non-disclosure agreements. However, these data tend to be scrubbed of personal identifying information, and limited in scope or by region, which may be necessary for privacy and data security but can limit its usefulness for public health action. Currently, public health does not have a consistent, established relationship with technology companies to obtain or leverage the data or innovative analysis of unstructured and ‘big data’ that many of these companies are undertaking. However, a few notable exceptions provide promising models for such collaborations. Examples of ongoing efforts spearheaded by technology companies to use the person-level data these companies own include:

- *Google* researchers aggregated searches for cold and flu symptoms to identify whether they could provide early insight of disease outbreaks and trends. On aggregate, these searches were correlated with clinical reporting of flu and cold symptoms, but the methodology has received criticism.¹² This technique has since been used to improve the Centers for Disease Control and Prevention (CDC) syndromic surveillance, to support COVID-19 forecasting, and to predict other health and economic trends.
- Local public health agencies (including Chicago, New York, and Salt Lake City) have collaborated with companies, such as *Yelp* and *Twitter*, to pilot programs to identify food-borne illnesses based on reporting in tweets or restaurant reviews.^{13,14}
- *Facebook* has used its Data for Good platform to provide its data to international development stakeholders.¹⁵ This includes location data to build “disaster maps” to aid disaster response and using its algorithms to analyze content of Facebook posts to improve public health campaigns.
- *Uber Movement* makes public aggregated data of Uber trips in cities across the world to aid urban planning, such as highlighting areas that may need additional transportation infrastructure and mapping traffic collisions that could be reduced by implementing street calming policies.^{16,17}

But without coordination of these new public health data content efforts, this assortment of discrete technologies and company efforts means that any public health-focused efforts will be isolated and

disconnected from each other. Public health decisions can only be made on what is available, rather than something designed specifically to support public health on a consistent basis.

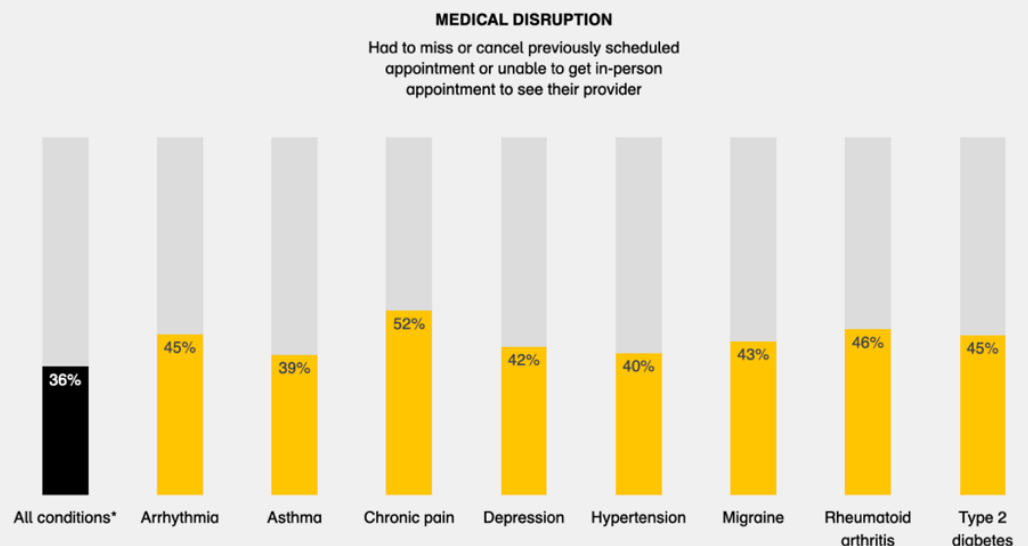
The COVID-19 pandemic also has spurred companies to create new efforts and explore how their existing data/tools could be used to address public health. Examples of COVID-19 specific technology and health innovations include:

- *Google's Community Mobility Reports*, which provide an example of how tech data can be used to provide data insight with greater speed.¹⁸ Google used location history data to show movement trends (e.g., visits to parks) in a variety of different regions. Other tech companies aggregated data from sources such as mobile devices, wearables, and cell towers to show how communities were interacting or engaging in social functions during quarantine. These examples demonstrate how these data could be used to observe how populations might respond to public health interventions.
- *Evidation Health* collects monthly data from a panel of 185,000 people across the United States, assessing their attitudes, behaviors, and health in response to the pandemic.¹⁹ **Figure 1** offers an example of data on medical disruption as a result of the pandemic.
- Some companies that develop wearable technologies started to use predictive analytics to detect COVID-19 or flu like symptoms.

Figure 1. Medical disruption in response to the COVID-19 pandemic

In the past two weeks, have you experienced any of the following with regards to your general or ongoing medical care (not related to coronavirus)?

PERCENTAGE OF RESPONDENTS WITH SPECIFIC MEDICAL CONDITIONS



*ACHIEVEMENT HEALTH PROFILE INCLUDES 39 MEDICAL CONDITIONS

N = 88,218 ACROSS ALL RESPONDENTS (NOT JUST LISTED CONDITIONS)
BASED ON RESPONSES FROM 3/31-4/8

© EVIDATION HEALTH, INC. 2020

evidation

From Evidation Health COVID-19 Pulse¹⁹

Having a parsimonious set of public health measures, as noted earlier, may make it easier to partner with technology companies and integrate or leverage unstructured and ‘big data’ sources and other technological innovations to promote public health.

2. Content of public health data

2.1 Key Action: Assess whether the content of public health data currently prioritized are what is needed to facilitate timely, proactive, and evidence-based decisions going forward.

Identifying a set of public health priorities could provide insight into needed adjustments to the type of public health data being collected or the way it is being analyzed or shared. While there is a vast array of public health data available, there are still gaps in fundamental areas that limit understanding and monitoring of positive health behaviors and well-being; upstream and root drivers of health—health inequities, structural inequities, and social determinants of health (see **Table 2** for definitions of these terms); resilience and preparedness; and the effectiveness of public health solutions.

For instance, public health data could place more emphasis on social context and other key factors (e.g., education, housing instability, food insecurity) that have a large impact in populations with vulnerabilities and place less emphasis on traditional health data that focus primarily on disease and disease burden. Data on lived experiences and community history (e.g., narratives), religion, banking (e.g., mortgage acceptance rates, redlining) and budgets (e.g., school budgets, county budgets), and racism and discrimination could provide critical insights into some of the root causes of health inequities.

Table 2. Definitions of terms used to capture upstream and root drivers of health

Key Term	Brief Definition
Health inequities	Systematic differences in the opportunities groups have to achieve optimal health, leading to unfair and avoidable differences in health outcomes. ^{20,21} The dimensions of social identity and location that organize or “structure” differential access to opportunities for health include race and ethnicity, gender, employment and socioeconomic status, disability and immigration status, geography, and more.
Structural inequities	Personal, interpersonal, institutional, and systemic drivers—such as, racism, sexism, classism, able-ism, xenophobia, and homophobia—that make those identities salient to the fair distribution of health opportunities and outcomes. Policies that foster inequities at all levels (from organization to community to county, state, and nation) are critical drivers of structural inequities.
Social, environmental, economic, and cultural determinants of health	The terrain on which structural inequities produce health inequities. These multiple determinants are the conditions in which people live, including access to good food, water, and housing; the quality of schools, workplaces, and neighborhoods; and the composition of social networks and nature of social relations.

From National Academy of Sciences²²

Data could capture infrastructure, including social infrastructure (e.g., social support, isolation, political representation, power imbalance) and physical infrastructure (e.g., transportation opportunities), to strengthen an evolving understanding of how these factors produce health (**Exhibit 2**). Given the dynamic nature of health, longitudinal data can assess changes over time as demographics shift due to climate change and other factors, for example, and as strategies to address inequity and upstream drivers of health are implemented.

Exhibit 2: Finding timely measures of physical, economic, and social environmental conditions affecting health needed for Public Health 3.0

New York University's Grossman School of Medicine conducted a rapid horizon scan on measures and analytic methods to incorporate timely information on social, economic, and environmental conditions affecting health into public health data systems in an effort to inform the development of Public Health 3.0 surveillance measures.²³ This scan showed that 'a gulf' exists between the technological and statistical innovations in academia and private industry and the day-to-day needs of public health practitioners. Few timely measures exist with sufficient spatial temporal granularity needed for public health departments to surveil acute stresses (e.g., COVID-19) or social determinants of health. Funding, partnerships, and data aggregation tools are needed if public health departments are to take advantage of advances in small area estimation, including tailored and adaptive spatial and visual data representation (e.g., drawing maps based on the prevalence of a certain public health phenomenon versus based on county boundaries), in ways that will not only protect privacy, but will be flexible and timely enough to meet evolving public health needs.

This section discusses considerations in transforming the content of public health data available to better capture positive health and well-being, data on the influence of the broader social and structural context, and data that sit at the intersectionality of public health with other sectors like behavioral health and disaster preparedness. Considering how to collect data in the most accessible and actionable format, and how to maintain parsimony across data sources (as mentioned earlier) will be important, as more data does not always translate to better information and insight.

2.1a. Consideration: *Despite some new advancements in integrating measures of subjective well-being, the public health data system does not consistently represent measures of positive health and well-being, inclusive of community well-being.*

There are several U.S. efforts underway that are working to include broader conceptualizations of well-being in surveys (**Table 3**). This includes the Gallup survey, which captures subjective well-being (e.g., optimism, hope, resilience) and happiness; Measure of America, which includes some sentinel measures of well-being and equity; and Well Being in the Nation, which includes measures of well-being and has advanced the consistent use of the Cantril ladder, a measure of life evaluation and expectation linked to other measures of mental health and well-being in a population.²⁴⁻²⁶ In addition to these established efforts, Healthy People 2030 has begun to more fully integrate subjective well-being and to push those measures in national surveys, such as the Behavior Risk Factor Surveillance System. CDC has integrated well-being concepts for some time, but currently those are mostly linked to Health-Related Quality of Life.²⁷

Areas such as lifelong learning and prosocial health behaviors, useful constructs for understanding drivers of health outcomes, are noticeably absent even from these efforts to advance well-being assessment. Within the constellation of these emerging efforts, only one public health data source captures the key constructs of hope, optimism, and life satisfaction. Though there has been advancement in the use of surveys to capture subjective well-being and understand the drivers of well-being (**Exhibit 3**), there are opportunities to use more information coming from the private sector to fill these gaps and consistently link information on sentiment and community needs to inform public health action. To date, these efforts are not standardized.²⁸

Table 3. Measure domains of thriving and flourishing in common data collection efforts that could strengthen a modern public health data system
 (“--“ means not present, X means present)

Dimensions	Well Being in the Nation ^b	NHIS ^c	NHANES ^d	Gallup ^e	Measure of America ^f	County Health Rankings ^g	BRFSS ^h
Subjective well-being	X	--	X	X	X	X	X
Cantril ladder	X	--	--	--	--	--	--
Life satisfaction	--	--	--	X	--	--	--
Thriving/flourishing	X	--	--	X	--	--	--
Hope	--	--	--	X	--	--	--
Optimism	--	--	--	X	--	--	--
Lifelong learning	--	--	--	X	--	--	--
Built and social environment (“place and planet”)	X	X	--	X	X	X	--
Community (e.g., social capital, loneliness, social networks)	X	X*	--	--	--	X	--
Civic engagement	X	--	--	X	X	--	--
Prosocial health behaviors	--	--	--	--	--	--	--
Health promotion behaviors	X	X	X	X	X	--	X
Economic opportunity	X	X	X	X	X	X	X

*added in 2020 questionnaire

^b <https://static1.squarespace.com/static/5ec9981e9cdcb22cafcc4750/t/5ed811b9a3592e107ce59f53/1591218618786/WIN-Report-060319-FINAL.pdf>

^c https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Survey_Questionnaires/NHIS/2020/EnglishQuest.pdf;
https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Survey_Questionnaires/NHIS/2020/EnglishQuest-ForQ3Q4.pdf

^d <https://wwwn.cdc.gov/nchs/data/nhanes/survey-contents-508.pdf>

^e <https://news.gallup.com/poll/trends.aspx>

^f http://measureofamerica.org/maps/?state^health^all_all^HDI^hdi

^g <https://www.countyhealthrankings.org/explore-health-rankings/measures-data-sources/2021-measures>

^h <https://www.cdc.gov/brfss/questionnaires/pdf-ques/2019-BRFSS-Questionnaire-508.pdf>

Exhibit 3: Understanding the drivers of well-being during a pandemic

The **National League of Cities** built a citywide coalition committed to using well-being data to advance health equity during the ongoing global pandemic.²⁹ Specifically, the coalition explored the drivers of well-being in Atlanta to understand and explain differences across populations using data from the **Gallup National Health and Well-Being Index**. A survey of Atlanta residents found differences in the drivers of well-being across the life span, with employment and work being important to younger populations (18-30 years old); personal finances to the middle-aged (31-59 years old); and health, access to health care, and sense of purpose being most important to older adults (60+ years old). Drivers of well-being also differed somewhat by ethnicity with work and employment, purpose, and safety and security mattering slightly more to Black populations relative to white populations, and personal finances, personal relationships, community, health status, and access to health care mattering slightly more to white populations than Black populations.

While this growing emphasis represents an important shift to focus on measures that can track thriving and flourishing, gaps remain in how consistently these measures are used across U.S. communities, how these measures are linked to resource allocation and other public health action, and the extent to which well-being also includes well-being measures of the environment and systems, often known as *community* and *civic well-being*. Work on community well-being has advanced both in the United States and globally.^{30,31} But some of this work has been criticized for not fully lifting up issues of equity, including but not limited to differential access to opportunity, historical and systemic barriers to well-being, and policies and practices that impede well-being. Work on community allostatic load as well as studies in the area of equity are important to consider when pursuing more comprehensive well-being measurement.³² **Table 4** outlines some areas of community well-being and civic well-being generally missing from current public health data.

Table 4. Measure domains of community and civic well-being that could strengthen a modern public health data system

Community well-being (measures of community health status, community conditions/assets)	Civic well-being (policies and governance)
Perception of community health	Civic engagement, well-being, and power
Thriving and flourishing	Equity tradeoffs/options of well-being investments
Health and well-being promoting amenities (<i>meaning features and conditions that actively promote positive health and well-being</i>)	Historical and systemic barriers to well-being
Valuation of community well-being and willingness to invest in well-being amenities	Well-being based policies
Intergenerational health (<i>or intergenerational health wealth</i>)	Governance models to advance well-being

2.1b. Consideration: *Despite a growing array of public health data, gaps still exist in the collection of upstream drivers and the social and structural context required to support public health action.*

Even with increasing amounts of health data being collected, public health data are limited in their ability to inform decision-making because data on many upstream factors that contribute to health are not regularly or consistently collected. A review of the Healthy People 2030 Leading Health Indicators by social determinants of health suggests continued gaps in current public health data (**Table 5**), particularly related to economic stability, neighborhood and the built environment, education, and community and social context. A broader review of national public health data sources for social determinants of health revealed similar gaps (see **Appendix A for expanded table**).

Table 5. Healthy People 2030 leading health indicators mapped to social determinants of health

Social Determinants of Health ³³		Healthy People 2030 Leading Health Indicators (“--” means not present)
Economic Stability	Employment	Employment among the working-age population
	Income	--
	Expenses	--
	Debt	--
	Medical Bills	--
	Support	--
	Hunger	Household food insecurity and hunger
Neighborhood and Built Environment	Housing	--
	Transportation	--
	Safety	Homicides Exposure to unhealthy air
	Parks	--
	Playgrounds	--
	Walkability	Adults who meet current minimum guidelines for aerobic physical activity and muscle-strengthening activity
	Access to healthy food options	Consumption of calories from added sugars by persons aged 2 years and over (2+ years)
	Zip code/geography	--
	Literacy	4th-grade students whose reading skills are at or above the proficient achievement level for their grade
Education	Language	--
	Early childhood education	--
	Vocational training	--
	Higher education	--
	Social integration	--
Community and Social Context	Support systems	--
	Community engagement	--
	Discrimination	--
	Stress and trauma	--

Health and Health Care	Health coverage	Persons with medical insurance (<65 years)
	Provider availability	--
	Provider linguistic and cultural competency	--
	Quality of care	--
Not aligned with social determinants	Receipt of treatment	Adolescents with major depressive episodes (MDEs) who receive treatment
	Receipt or use of diagnostic or preventative care	Children, adolescents, and adults who use the oral health care system (2+ years)
		Adults who receive a colorectal cancer screening based on the most recent guidelines
		Persons who are vaccinated annually against seasonal influenza
		Adults with hypertension whose blood pressure is under control
		Persons who know their HIV status (13+ years)
	Death and disease	Drug overdose deaths
		Suicides
		Infant deaths
		Maternal deaths
New cases of diagnosed diabetes in the population		
Substance abuse behaviors	Children and adolescents with obesity	
	Cigarette smoking in adults	
	Current use of any tobacco products among adolescents	
	Adults engaging in binge drinking of alcoholic beverages during the past 30 days	

2.1c. Consideration: *Disaster preparedness and resilience measures have evolved, but they still fall short with respect to equity orientation and agility.*

The National Health Security Preparedness Index (NHSPI) is one of the only nationally available comprehensive measures of disaster preparedness and resilience (including domains of health security surveillance, community planning and engagement coordination, incident and information management, health care delivery, and countermeasure management), but similar to other resilience measures, it focuses mainly on structural capacities such as people, institutions, policies, and programs that can “be easily observed and counted.”³⁴ While the NHSPI is a tremendous advancement in national- and state-level preparedness tracking, NHSPI measures currently lack the ability to capture concepts related to effectiveness, efficiency, or equity of disaster preparedness efforts (e.g., contextual equity), and are rarely linked together or with other relevant sources to capture preparedness or resilience. Further, the NHSPI includes many indicators, which challenges several of the data system principles noted in the preamble, namely the use of data for sensemaking and parsimony. Use of private sector sources to capture emerging health security threats is growing but certainly has not yet reached optimal use.

The CDC is exploring new ways to augment their syndromic surveillance system to allow mental health monitoring, particularly in the context of disaster response and recovery. The Council of State

and Territorial Epidemiologists (CSTE) is in the process of developing guidance for a series of mental health definitions that will be linked to queries that public health departments can use as part of ESSENSE on the National Syndromic Surveillance Platform.³⁵ Some states have begun to conduct behavioral health surveillance as part of their response to COVID-19.

For example, Washington state releases weekly behavioral health situation reports.³⁶ Using data from Washington's ESSENSE, the public health department in Washington state has been tracking weekly, since 2019, the relative count of emergency department (ED) visits for psychological distress, suicidal ideation, suspected suicide attempts, all drug-related visits, and alcohol-related visits. The situation report contains data for Washington state from the National Center for Health Statistics (NCHS) and the Census Bureau's Household Pulse Survey, which track symptoms of anxiety and depression among the U.S. population to learn about the impact of COVID-19.³⁷ Year-over-year percent changes in the sale of alcohol, tobacco, and cannabis, and percent change of NCIS handgun and firearm background checks—all factors correlated with mental health concerns—also are tracked on a monthly basis. Finally, the situation report contains an analysis of Tweets about COVID-19 geotagged in Washington state to track a 7-day moving average of deviation in expressions of positive sentiment, loneliness, and anxiety relative to a January 2020 baseline.

Other systems are being used for more sensitive disaster monitoring. For instance, throughout the pandemic, 2-1-1s worked with state and local health departments to share information, which was used to inform local action on testing and vaccine distribution (see **Exhibit 1**).

2.1d. Consideration: *New measures require ground-truthing and validation.*

As public health measures evolve and new measures are implemented at the national and local level, continuous evaluation of these measures is needed to ensure these measures reflect lived experiences and community needs. It will be important to capture an absolute or relative change in health status over time, particularly changes in health status that are sensitive to major policy changes that affect health and are reliable and comparable across settings.³⁸ In addition, using the method of ground-truthing can ensure that the data and findings from public health surveillance measures are checked or 'validated' using observations on the ground.³⁹

3. Volume and coordination of public health data

3.1 Key Action: Efficiently and ethically use the growing volume of public health data across sectors.

While there are some meaningful gaps in public health data, there is a significant amount of existing public health data that are not being optimally used. Despite understanding the critical need for public health agencies to improve the use and integration of structured and unstructured public health data, department and software silos often have impeded use and integration of public health data.

This section discusses the considerations for the efficient and ethical use of existing public health data, particularly to deal with the growing volume of unstructured data in the digital universe as well as integration and coordination of these data. This section describes the limitations in how existing public health data are applied, the challenges to using the exponentially growing unstructured data, and the challenges of bringing together siloed public health data systems given this volume.

3.1a. Consideration: *Many public health data systems are not able to capitalize on the existing volume of health-related data.*

There is a wide array of public health data collected through surveys, surveillance systems, and transactions with the public health system. As previously mentioned, within Healthy People 2030, there are 355 targets that require a diversity of public health data from the National Vital Statistics System to the National Health Interview Survey. Despite nearly 100 diseases and conditions electronically being tracked by state and local public health departments, there are still an array of data and laboratory reports that are being collected in paper formats. For example, in 2018, 20 percent of public health laboratory reports were still being received in paper format.⁴⁰ The CDC has prioritized transitioning more data to digital formats for more accurate and timely data. Additionally, the CDC is investing in Digital Bridge⁴¹ to improve the use of electronic health records for public health decision-making. Seven states are piloting Digital Bridge, extracting electronic health record data to monitor five conditions: chlamydia, gonorrhea, pertussis, Salmonella, and the Zika virus. However, there are still many states with a mix of electronic data that are not being shared between public health and health care to inform public health decision-making, let alone between public health and other key sectors that influence health.

3.1b. Consideration: *The sheer volume of unstructured data available in the digital universe also can be untenable.*

Unstructured data refers to non-standardized data from a variety of sources (e.g., social media, clouds, sensors) and can include text, images, audio, video, blogs, websites, and so forth.⁴²⁻⁴⁶ Unstructured data are growing exponentially, with over 2.5 quintillion bytes of unstructured data generated every day. Unstructured data are expected to continue to grow over time, with an estimated 65 percent annual growth rate, and to comprise approximately 95 percent of global data.⁴⁴ These unstructured data can provide critical insights into public health (**Exhibit 4**); but these data have not yet been effectively integrated into public health data systems. For example, clinical notes and other free text from electronic health records have been analyzed for insights into patient care and overprescribing and use of certain medications.⁴⁷ Text from social media posts has been analyzed to look for positive or negative sentiment providing insight into psychological well-being of a population, as well as self-reported symptoms of a disease to help estimate disease outbreak.^{48,49} There is emerging research about the use of images posted on social media to track drinking, smoking, and obesity related behaviors and the use of wearable sensors for remote patient monitoring and virtual health assessments.^{50,51}

The huge volume and complexity of existing and unstructured data make it a tedious task to extract useful information from different types of data, and its use is fraught with concerns about ethical tradeoffs. More data are not useful unless it can be used for proactive decisions.⁵² While past research has demonstrated how unstructured data can inform public health decision-making, these investigations have raised concerns about how to maintain privacy, especially if artificial intelligence is used to generate or distill information which could be considered personally identifiable medical information.⁵² Data ownership is another ethical challenge identified in using unstructured data and will require strong governance policies (described in more detail in the How? paper).

Exhibit 4: Improving public health outcomes using sentiment analysis to understand resident sentiment toward COVID-19 vaccination

The **Harvard Kennedy School's Ash Center** partnered with Bennett Midland and Zencity to use sentiment analysis to understand the “organic” resident discourse on vaccinations.⁵³ The information provided insights about the experiences associated with positive and negative perceptions of the vaccination process (e.g., capability to do things after receiving a vaccination, mistrust of the public health system). This project identified opportunities and challenges for the use of sentiment analysis in areas of public health. The project notes that access and use of sentiment analysis in public health data systems offer a way to capture and understand drivers of resident behavior, including those related to social determinants of health; new and timely data on the response of residents to interventions; an ability to quickly respond with targeted messaging when matched with data “know how”; and opportunities to test time-limited, high-priority projects in new ways.

3.1c. Consideration: *Siloed data and data systems have created barriers to using public health data to identify timely solutions and do not provide complementary data useful to decision-makers.*

The fact that race and ethnicity data were not collected regularly as part of initial lab testing for COVID-19 not only resulted in a delay in public health's examination of impacts by race and ethnicity, but also limited the ability to examine the impacts of state-level interventions on specific populations (e.g., were states' social isolation interventions affecting certain race, age, or gender populations disproportionately).⁵⁴ Without information about who is being affected by what level and quality of public health services and policies, public health departments are unable to assess the efficiency or effectiveness of public health services and policies. To support sense-making and decision-making, data on population characteristics, the transactions and quality of public health services, and the public health outcomes must be integrated or linked across what are often siloed data sources. This type of linkage, however, requires significant resources and expertise and can hinder public health departments from finding timely solutions to the nation's most pressing public health problems.

The public health data systems and digital infrastructure have also not been developed to support data-driven research of and decision-making based on population experience within and outside of the health care system.⁵⁵ This may be due, in part, to the historic separation of medicine and public health, which was fueled by the divergence of federal policies in the 1930s and 1940s^{56,57} separating public health and individual health care services. However, a modern public health data system will not only need to leverage the intersection between public health and health care, but also capitalize on the intersections with other sectors relevant to health and well-being (e.g., education, transportation, criminal justice) to truly understand the conditions that create health inequities (**Exhibit 5**).

Exhibit 5: Integrating state prison and jail data under the COVID Prison Project to help understand health equity among incarcerated populations during the global pandemic

There are currently no uniform standards for health data transparency among departments of corrections. To fill this gap in transparency and availability of data on the incarcerated population, the **University of North Carolina at Chapel Hill** established the **COVID Prison Project**.⁵⁸ This Project collected data from 53 prison systems (Federal Bureau of Prisons, Immigration and Customs Enforcement, Puerto Rico, and all 50 states) and over 50 of the largest jails through the use of web-scraping technology to automate daily data collection. This Project found that COVID-19 infected almost 400,000 people who are incarcerated in state or federal prison systems and almost 2,500 have died; however, despite this burden, vaccine rollout has been slow with only 17 states reporting data on vaccinations and, among those states, between 3-60 percent of prisoners vaccinated. When this Project tried to make comparisons of testing for COVID-19 across states, differential reporting made it impossible to accurately make these comparisons. Many deaths were misclassified, and race/ethnicity data was among the missing data elements, limiting the ability of the project to look at disproportionate impacts of COVID-19. While COVID-19 has, to some extent, increased data transparency in prisons—without additional momentum and interest, it is unclear whether these gains will be sustained.⁵⁹

Siloed data systems preclude the use of complementary data and the establishment of ‘data quilts’ (interconnections or mosaics of data that can be used together) with sufficient statistical power to produce reliable surveillance estimates (**Exhibit 6**).⁶⁰ The ‘holy grail’ of public health data infrastructure has been described as “unified representations of data that incorporate the full spectrum of the determinants of health using data from multiple sources versus advancing siloed data within a homogenous data system (pg. 1).”⁶¹ However, public health is only now trying to integrate data systems across health care and public health, which have been traditionally siloed (e.g., Digital Bridge), and, with the exception of isolated demonstration communities, has not yet started to focus on data infrastructure to capture population experiences outside of the health care system.⁶² Such data can provide important signal value, as in the case of school absences or law enforcement data that could signal emerging health needs in the community.

Exhibit 6: Accelerating public health data modernization efforts during a global pandemic

The **Task Force for Global Health, Inc.** and the **Public Health Informatics Institute** partnered with and listened to health department staff, Association of State and Territorial Health Officials (ASTHO), National Association of County and City Health Officials (NACCHO), CSTE, Association of Public Health Laboratories (APHL), and the CDC to determine the most relevant information related to the needs of local and state health departments and develop recommendations on a way forward to modernize public health data and infrastructure.⁶³ Despite the horrific impact of COVID-19 on U.S. communities, public health departments considered it a transformative opportunity for public health to build back better together. To do so, public health departments recommended a focus on equity, the development of a nationwide strategy for modernization (e.g., through a standard set of nationwide requirements), a common data architecture to eliminate data silos and ensure interoperability, and the need to improve the informatics capability of public health departments. Innovations from public and private sector were also highlighted as case examples showing how public health data can be used in more effective, efficient, and functional ways.⁶³

4. Precision, granularity, and timeliness of public health data

4.1 Key Action: Clarify the appropriate level of precision and granularity of public health data needed for public health practitioners, policymakers, and the public to make informed and timely decisions.

Public health recommendations are often made because of emerging needs, based on evolving, incomplete, and sometimes imprecise public health data. However, timely and precise (defined in this paper as exact and/or sharply defined) data are scarce, in large part because of the time and resources required. One challenge, however, is that if decisions change as more precise data become available, it can be difficult to communicate what has changed and why to the public and policymakers in ways they can understand. The ongoing global pandemic is an example of how evolving circumstances can lead to inconsistent and contradictory information that contribute to public mistrust of public health data and associated recommendations. This section discusses the considerations for establishing guidance and guidelines to determine and communicate the appropriate level of precision and granularity of public health data needed for public health decision-making, the time lag inherent in many data sources, and the role of technology in improving precision and granularity of public health data.

4.1a. Consideration: *There are resource, communication, and political issues associated with data granularity decisions.*

Standardized measures of race and ethnicity (e.g., Asian/Pacific Islander, Hispanic) combine extremely heterogeneous populations, which have very different lived experiences, cultures, languages, and lifestyles. From a practical perspective, one's country of origin and legal status also has real-world implications for accessing (or willingness to access) federal, state, and local services and supports. As a result, true health risks, outcomes, and trajectories of many populations are masked, severely limiting the development of tailored solutions to meet critical health needs (**Exhibit 7**). Similar challenges occur for other populations, such as gender minorities and LGBTQ+, individuals of multiple race and ethnic heritage, and for individuals in rural areas, where data are often combined across multiple geographic regions, posing challenges for localized response. Ensuring an equity-oriented data system requires careful consideration about what types of data are collected and at what level of granularity, and how to balance valid concerns around individual privacy, with the benefit of disaggregated data to inform public health and policy solutions. This point is particularly salient for some populations, which may be at risk if such data were made publicly available (e.g., undocumented residents).

Exhibit 7: *Strengthening health equity data gathering to better understand the genesis and impact of health disparities*

Texas Southern University and the **Deep South Center for Environmental Justice** conducted a retrospective evaluation of the Gulf Coast Historically Black Colleges and Universities-Community Based Organization Consortium's 2017-2020 work to determine its impact on individual communities, the Gulf Coast region at large, and to analyze project outcomes using a racial justice and equity lens and community-based participatory research approach to develop key priorities and recommendations for approaches to health equity data gathering and sharing.⁶⁴ Key priorities identified included using qualitative data as part of the gold standard for an equitable data system to help ensure public health data reflects lived experience and community needs; building cross-sector linkages to monitor co-occurring threats to health and human safety; and to ensure data, including data on laws and policies, is integrated and disseminated with a focus on accessibility and equity.

The lack of data granularity poses challenges for understanding what works, when, and for whom. For example, public health departments may be able to examine how many men or women received a specific public health service, but cannot look beyond a binary intersectionality (e.g., how many Black women living in poverty are receiving that same public health service). More detailed data on population characteristics are often collected via survey and not at the point of service. Combining data on public health action and implementation with disaggregated population-level data can be done through data integration across different data sources.

Currently, there is no consensus on what level of geographical and other data granularity is needed to support public health data decisions at the national versus local levels. For example, some have argued for greater granularity of ethnicity classifications to capture sub-group variations in health care, risk factors, and health behaviors, as noted above.⁶⁵ Precision public health emerged as “a new field driven by technological advances that enable more precise descriptions and analysis of individuals and population groups, with a view to improving the overall health of populations.”⁶⁶

While first coined in 2013 and 2014, the optimal population strata to guide precision public health is still being determined.⁶⁷ It is important to acknowledge, however, that the identification of a specific and universal level of data granularity may risk further fragmenting communities facing difficult decisions around the level of disaggregation pursued in either data collection or analysis due to dwindling state and local governmental resources. There could be an unintentional overriding of the goals of a community in which greater levels of disaggregation would be beneficial to local decision-making.

Executive Order on Equitable Data

In a recent executive order, President Biden established an Interagency Working Group on Equitable Data, which is tasked with identifying “inadequacies in existing Federal data collection programs, policies, and infrastructure across agencies, and strategies for addressing any deficiencies identified; and support[ing] agencies in implementing actions, consistent with applicable law and privacy interests, that expand and refine the data available to the Federal Government to measure equity and capture the diversity of the American people.”

4.1b. Consideration: *A lot of public health data suffers from significant time lags in favor of precision, but the timeliness and corresponding precision of data needed by decision-makers varies.*

The types of public health data have adapted to changing public health data and informatics needs over time, moving from counts and trends to causal inferences to geospatial inferences over time.⁶⁸ Public health practitioners often rely on federal data systems to create those inferences, but those data must be collected, cleaned, and de-identified, creating significant time lags often measured in years. This requires significant resources and can hinder public health departments in making timely decisions because they are constantly relying on survey data that may be lagging by a year or more. Making key decisions with outdated information can negatively impact public health, particularly if populations are dealing with a public health emergency.⁶⁹ For example, the lack of timely, federal data made it more difficult to use to inform a pandemic containment strategy and shed new light on the limitations of our current public health data system for supporting real-time decision-making. While countries like South Korea, Singapore, and New Zealand were disseminating public health data in near real-time, the U.S. data lagged and were often lacking critical information, such as race and ethnicity.⁷⁰ An August 2020 analysis of COVID-19-related cases and deaths found the *data quality* differences between South Korea and the United States correlated with differential response and death rates.⁷⁰ The authors cited that some of the differential response could be attributed to South Korea’s coordinated network of public health centers across 250 districts, which had the data infrastructure to readily feed intelligence to the Korea Centers for Disease Control and Prevention for timelier and more targeted pandemic action. In the absence of available U.S. federal data, the

COVID-19 Tracking Project emerged as a public-private partnership to provide transparent and timely data on COVID-19 cases, deaths, and testing.⁷¹

There are circumstances when more and less precise data are appropriate, and it also can come down to a decision of data quantity versus timeliness. Decision-making related to emerging public health issues, such as COVID-19, often are fraught with a great deal of uncertainty, and conditions are continually evolving. Rapid health surveillance can potentially “pass” with less precision. Washington state public health has been forecasting mental health impacts of COVID-19 to help inform a series of proactive public health interventions to stem the tide of anticipated mental health problems, as described in Section 2. Many of these proactive measures (e.g., an education campaign to teach people how to cope with stress and where to find mental health resources) can be implemented with little to no harm, and only have the opportunity to help, so the risk for acting with less precise data is low. Polling is a common practice that can also produce targeted information for decision-making in near real-time but may lack precision. The Kaiser Family Foundation conducted a Tracking Poll in mid-July to assess the impact of COVID-19 on adult mental health. The Tracking Poll found that 53 percent of adults reported a negative impact on their mental health due to worry and stress over the pandemic, which helped to raise awareness among policymakers and the public of the rising mental health burden due to COVID-19.⁷²

These are good examples where less precision was warranted, but before policymakers and other funders dedicate more comprehensive resources to public health interventions and plans, more precise data to understand the potential return on investment often is needed. Research studies often require higher precision data to minimize error in studies and to ensure accurate results before drawing population conclusions.

Public health officials must make a determination about whether public health data are ‘fit for purpose’ to inform decisions as conditions evolve; however, these nuanced decisions can be challenging to convey. Communication guidance and guidelines are needed for how public health officials should handle explaining the precision of the data by articulating the strengths and weaknesses as well as best use cases of different data sources for pending public health needs.⁷³

4.1c. Consideration: *Technology solutions may offer templates in how best to balance precision and granularity with real-time response.*

Existing public health data challenges related to the speed and precision trade-off may be overcome by learning from and partnering with technology companies, which have resources and data capabilities that, in many cases, far exceed the typical public health sector. These companies have data that are already continually collected and analyzed. Using massive amounts of data in combination with sophisticated analytic techniques, companies are able to create accurate profiles of individuals or population segments based on numerous characteristics. These capabilities can be used to rapidly identify behavior changes, to answer emerging questions, or to spot shifts in sentiment. Social media posts or web searches provide real-time, granular information that can

reveal mental states, such as depression or anxiety, or whether someone has searched for information about how to hurt themselves or others.^{11,74}

Technology offers a convenience factor that could broaden the scope of collection. Many products that collect data are already in our households, which could help lessen a reliance on traditional data collection procedures that are more laborious and take longer to implement. One impact of this could be to increase the representation in public health data of an otherwise-hidden population. People who might be difficult to reach or who might not come in contact with health systems might not show up in current surveillance systems. However, the vast majority of people go online at some point, and many more people have cell phones, including those experiencing homelessness or other populations with vulnerabilities.^{75,76} Attention should still be paid to individuals that may not be online due to the lack of broadband in some regions, limited access to free public internet, inconsistent access to the internet due to economic instability, and varying degrees of digital literacy. It will be critical to acknowledge and address this digital divide so as to not create further inequity among populations that bear significant public health burdens but that are not reflected in online data.⁷⁷

Conclusion

This paper reviewed several challenges for using data for proactive and timely decision-making. Public health data often are reactive to emerging areas of focus, rather than proactively driving decisions, and that reactivity has contributed to a large number of siloed systems and indicators. The systems that do exist suffer from significant time lags and gaps in meaningful data to capture positive health and well-being, preparedness and resilience, and upstream, social, and structural factors at both individual and community levels. Despite the existence of these gaps, a large volume of existing public health data is not being utilized effectively. Unstructured data are growing exponentially, yet public health has not systematically partnered with technology companies to leverage these data and, therefore, innovative efforts to use the data for public health decision-making are disconnected.

Despite these challenges, there are opportunities for improving the consistency, parsimony, and relevance of data in the modern public health data system. A parsimonious set of core public health measures, particularly at the national level, would facilitate timely, proactive, and evidence-based decisions. The modern public health data system could be strengthened by including data on the root causes and upstream drivers of public health outcomes, as well as from other sectors whose data may signal emerging health needs. Better data system governance needs to be in place to promote the efficient and ethical use of the growing volume of public health data and to clarify the appropriate level of precision and granularity of public health data needed for public health practitioners, policymakers, and the public to make informed decisions. If public health data are to have an impact, reviewing the content of public health data to determine how it can be used for proactive public health action and how it can be used with a more deliberate equity lens is essential.

Appendix A. Gaps in Social Determinants of Health Data from Varied National Public Health Data Sources

("--" means not present, X means present)

Social Determinants of Health		Healthy People 2030 Leading Health Indicators	Behavioral Risk Factor Surveillance System (BRFSS)	National Health and Nutrition Examination Survey (NHANES)	National Health Interview Survey (NHIS)	Well-Being in the Nation	National Equity Atlas	County Health Rankings and Roadmaps
Economic Stability	Employment	X	X	--	X	X	X	--
	Income	--	X	X	X	X	X	X
	Expenses	--	--	--	--	--	--	--
	Debt	--	--	--	--	--	--	--
	Medical Bills	--	--	--	--	--	--	--
	Support	--	--	--	--	--	--	--
	Hunger	X	X	X	X	--		X
	Access to a car	--	--	--	--	--	X	--
	Rent/housing burden	--	--	--	--	--	X	X
	Access to internet	--	--	--	--	--	--	X
Neighborhood and Built Environment	Housing	--	X	X	X	--	X	--
	Transportation	--	--	--	--	--	X	--
	Safety	X	--	X	X	X	--	--
	Parks	--	--	--	X	--	--	--
	Playgrounds	--	--	--	X	--	--	--
	Walkability	X	--	--	X	--	--	--
	Access to healthy food options	X	X	X	--	--	--	X

Social Determinants of Health		Healthy People 2030 Leading Health Indicators	Behavioral Risk Factor Surveillance System (BRFSS)	National Health and Nutrition Examination Survey (NHANES)	National Health Interview Survey (NHIS)	Well-Being in the Nation	National Equity Atlas	County Health Rankings and Roadmaps
	Zip code/geography	--	X	--	--	--	--	--
Education	Literacy	X	--	--	--	--	--	--
	Language	--	--	--	--	--	--	--
	Early childhood education	--	X	--	X	X	X	--
	Vocational training	--	X	--	X	X	X	--
	Higher education	--	X	--	X	X	X	--
Community and Social Context	Social integration	--	--	--	--	--	--	--
	Support systems	--	--	--	--	X	--	--
	Community engagement	--	--	--	--	--	--	--
	Discrimination	--	--	--	--	--	--	--
	Stress and trauma	--	X	--	X	X	--	X
Health and Health Care	Health coverage	X	X	X			--	X
	Provider availability	--	X	--	--	X	--	--
	Primary care providers	--	X	--	--	X	--	--
	Health care access barriers	--	X	--	--	--	--	--
	Provider linguistic and cultural competency	--	--	--	--	--	--	--
	Quality of care	--	--	--	--	--	--	--

Social Determinants of Health		Healthy People 2030 Leading Health Indicators	Behavioral Risk Factor Surveillance System (BRFSS)	National Health and Nutrition Examination Survey (NHANES)	National Health Interview Survey (NHIS)	Well-Being in the Nation	National Equity Atlas	County Health Rankings and Roadmaps
	Self-reported health and well-being status	--	X	X	X	X	--	X
	Chronic health conditions	--	X	X	X	--	--	X
	Current medications and/or treatments	--	X	--	--	--	--	--
	Exercise	--	X	X	X	--	--	--
	Receipt of treatment	X	X	--		X	--	--
Not aligned with social determinants	Receipt or use of diagnostic or preventative care	X	X	--	X	X	--	X
	Death and disease	X	--	--	--	X	--	X
	Substance abuse behaviors	X	X	X	X	X	--	X

Acknowledgements

Each paper in this series of white papers has been reviewed by Mahshid Abir of the RAND Corporation and the University of Michigan and Karen Smith of Public Health Strategies. We appreciate these reviewers for their insights and contributions. We also thank Burness for their editorial and communication support for each paper.

References

1. Hanlon P, Carlisle S, Hannah M, Reilly D, Lyon A. Making the case for a 'fifth wave' in public health. *Public Health*. 2011;125(1):30-36.
2. Leveraging Community Information Exchanges (CIEs) for Equitable and Inclusive Data: CIE Data Equity Framework [PowerPoint]: Community Information Exchange; 2021.
3. Leveraging Community Information Exchanges for Equitable and Inclusive Data: CIE Community Profiles: Community Information Exchange; 2021.
4. Grounds K, Johnson B, Christenson C. *Leveraging Community Information Exchanges (CIEs) for Equitable and Inclusive Data: CIE Case Study*. San Diego, Calif.: 211 San Diego; 2021.
5. Assessing the Value, Best Practices, and Potential Uses of 211 Data to Inform Public Health During the Pandemic and Beyond [press release]. Alexandria, VA: United Way Worldwide; 2021.
6. CDC. 10 Essential Public Health Services. Centers for Disease Control and Prevention. <https://www.cdc.gov/publichealthgateway/publichealthservices/essentialhealthservices.html>. Published 2021. Accessed January 4, 2021.
7. 2020 DHP. *Appendix 7. Building on Past Challenges of the Healthy People Initiative*. Online: U.S. Department of Health & Human Services; 2020.
8. Meit M, Knudson A, Dickman I, Brown A, Hernandez N, Kronstadt J. *An Examination of Public Health Financing in the United States*. Washington, D. C.: NORC at the University of Chicago; 2013.
9. Robbins R. Tech companies are crucial players in the coronavirus response. Are they contributing what's most needed? *STAT News*. Health Tech Web site. <https://www.statnews.com/2020/03/25/tech-companies-coronavirus-response/>. Published 2020. Accessed 2020.
10. Romm T, Dwoskin E, Timberg C. U.S. government, tech industry discussing ways to use smartphone location data to combat coronavirus. *The Washington Post* 2020; Tech Policy.
11. Stephens-Davidowitz S. Dr. Google Will See You Now. *The New York Times* 2013.
12. Lazer D, Kennedy R, King G, Vespignani A. The Parable of Google Flu: Traps in Big Data Analysis. *Science*. 2014;343(6176):1203-1205.
13. Lehman S. Twitter helps Chicago find sources of food poisoning. *Healthcare & Pharma*. 2014. <https://www.reuters.com/article/us-chicago-twitter-food-poisoning/twitter-helps-chicago-find-sources-of-food-poisoning-idUSKBN0GQ25820140826>.
14. Shaprio M, Vilven D. *Using Yelp Restaurant Reviews to Identify Foodborne Illness*. Online: Salt Lake County Health Department; 2018.
15. Cheney C. What the Facebook scandal means for 'data for good'. *Inside Development*. 2018. <https://www.devex.com/news/what-the-facebook-scandal-means-for-data-for-good-92425>.
16. Introducing Uber Movement [press release]. Online: Uber Newsroom 2017.
17. Announcing Uber Movement Street Speeds [press release]. Online: Uber Newsroom 2019.
18. COVID-19 Community Mobility Reports. Google. <https://www.google.com/covid19/mobility/>. Published 2020. Accessed 2021.
19. Evidation. COVID-19 Pulse: Delivering regular insights on the pandemic from a 150,000+ person connected cohort. Evidation Health, Inc. <https://evidation.com/news/covid-19-pulse-first-data-evidation/>. Published 2020. Accessed March 11, 2021.

20. Braveman P. Health disparities and health equity: concepts and measurement. *Annual Review of Public Health*. 2006;27:167-194.
21. Organization WH. Health inequities and their causes. World Health Organization. <https://www.who.int/news-room/facts-in-pictures/detail/health-inequities-and-their-causes>. Published 2018. Accessed April 8, 2021.
22. The Root Causes of Health Inequity. In: National Academies of Sciences E, and Medicine, Division HaM, Practice BoPHaPH, et al., eds. *Communities in Action: Pathways to Health Equity*. Washington, D. C.: National Academies Press; 2017.
23. Thorpe LE, Chunara R, Roberts T, et al. *Public Health Surveillance 3.0: A 2021 Horizon Scan of Emerging Timely Measures of Physical, Economic and Social Environmental Conditions Affecting Health [White Paper]*. New York, N. Y.: NYU Grossman School of Medicine; 2021.
24. Well Being in the Nation (WIN) Measures. Well Being in the Nation Network. <https://www.winmeasures.org/statistics/winmeasures>. Published undated. Accessed January 4, 2021.
25. Clifton J. The State of World Happiness in 2019. In. *Opinion*. Vol 2021. Online: Gallup; 2019.
26. Lewis K, Burd-Sharps S. The State of America's Well-Being. In. *Measure of America of the Social Science Research Council*. Vol 2020. Online: Measure of America of the Social Science Research Council; 2013.
27. CDC. Health-Related Quality of Life (HRQOL). Centers for Disease Control and Prevention. <https://www.cdc.gov/hrqol/wellbeing.htm>. Published 2018. Accessed 2021.
28. Luhman M. Using Big Data to study subjective well-being. *Current Opinion in Behavioral Sciences*. 2017;18:28-33.
29. Cities NLo. *The State of Well-being in Atlanta A Baseline Analysis*. Online: National League of Cities; 2021.
30. Community wellbeing indicators. In. *what works wellbeing*. Vol 2021. Online: Scribd; 2017.
31. VanderWeele TJ. Measures of Community Well-Being: a Template. *International Journal of Com*. 2019;2:253-275.
32. Chandra A, Cahill M, Yeung D, Ross R. *Measuring Community Stress*. Online: Robert Wood Johnson Foundation; 2018.
33. Organization WH. About Social Determinants of Health. World Health Organization. http://www.who.int/social_determinants/sdh_definition/en/. Accessed April 25, 2020.
34. National Health Security Preparedness Index, Access the new 2020 Index release data, available as of June 25, 2020. Prepared. <https://nhspi.org/>. Published 2021. Accessed January 4, 2021.
35. CSTE. Using the power of epidemiology to fight the spread of COVID-19. CSTE. <https://www.cste.org/>. Published undated. Accessed April 10, 2021.
36. Health WSDo. *Week of March 29, 2021 COVID-19 Behavioral Health Impact Situation Report*. Online: Washington State Department of Health; 2021.
37. CDC. Household Pulse Survey. Centers for Disease Control and Prevention. National Center for Health Statistics Web site. <https://www.cdc.gov/nchs/covid19/pulse/mental-health.htm>. Published 2021. Accessed April 8, 2021.
38. Thacker SB, Stroup DF, Carande-Kulis V, Marks JS, Roy K, Gerberding JL. Measuring the public's health. *Public Health Reports*. 2006;121(1):14-22.
39. Sadd J, Morello-Frosch R, Pastor M, Matsuoka M, Pritchard M, Carter V. The truth, the whole truth, and nothing but the ground-truth: Methods to advance environmental justice and researcher-community partnerships. *Health Education & Behavior*. 2014;41(3):281-290.
40. CDC. Partnerships Push Progress. Centers for Disease Control and Prevention. Public Health Surveillance and Data Lab Reporting Web site. <https://www.cdc.gov/surveillance/initiatives/lab-reporting.html>. Published 2018. Accessed April 8, 2021.
41. Bridge D. Digital Bridge: About. Digital Bridge. <https://digitalbridge.us/about/>. Published 2021. Accessed February 15, 2021.
42. Lomotey RK, Deters R. Topics and terms mining in unstructured data stores. 2013 IEEE 16th international conference on computational science and engineering; 2013.
43. Lomotey RK, Deters R. RSEnter: Terms Mining Tool from Unstructured Data Sources. *International Journal of Business Process Integration and Management*. 2013;6(4):298-311.
44. Lomotey RK, Jamal S, Deters R. SOPHRA: a mobile web services hosting infrastructure in mHealth. Paper presented at: First international conference on mobile services; 2012; New York, N. Y.
45. Scheffer T, Decomain C, Wrobel S. Mining the Web with active hidden Markov models. Paper presented at: International conference on data mining2001; New York, N. Y.
46. Wang Y, Kung LA, Byrd TA. Big data analytics: understanding its capabilities and potential benefits for healthcare organizations. *Technol Forecast Soc Change*. 2018;126:3-13.
47. Luo Y, Thompson WK, Herr TM, et al. Natural Language Processing for EHR-Based Pharmacovigilance: A Structured Review. *Drug Safety*. 2017;40(11):1075-1089.
48. Fung ICH, Tse ZTH, W. FK. The use of social media in public health surveillance. *Western Pacific Surveillance and Response Journal*. 2015;6(2):3.
49. Wongkoblap A, Vadillo MA, Curcin V. Researching mental health disorders in the era of social media: systematic review. *Journal of Medical Internet Research*. 2017;19(6):e228.

50. Garimella VRK, Alfayad A, Weber I. Social media image analysis for public health. Paper presented at: 2016 CHI Conference on Human Factors in Computing Systems 2016.
51. Seshadri DR, Davies EV, Harlow ER, et al. Wearable Sensors for COVID-19: A Call to Action to Harness Our Digital Infrastructure for Remote Patient Monitoring and Virtual Assessments. *Frontiers in Digital Health*. 2020;2(8).
52. Adnan K, Akbar R. An analytical study of information extraction from unstructured and multidimensional big data. *Journal of Big Data*. 2019;6:91.
53. Zencity, School HK, BennettMidland. *Interim Report Improving Public Health Outcomes Using Sentiment Analysis: COVID-19 vaccine resident sentiment*. Zencity et al.; 2021.
54. Meraji SM. Code Switch [Internet]: NPR; 2020. Podcast: 18 minutes. Available from: <https://www.npr.org/2020/04/21/840609912/the-news-beyond-the-covid-numbers>
55. Massoudi BL, Chester KG. Public health, population health, and epidemiology informatics: recent research and trends in the United States. *Yearbook of Medical Informatics*. 2017;26(1):241.
56. Sledge D. Divergent Paths. In: *Health Divided: Public Health and Individual Medicine in the Making of the Modern American State*. Kansas: University Press of Kansas; 2017:157-178.
57. Sledge D. Health Divided. In: *Health Divided: Public Health and Individual Medicine in the Making of the Modern American State*. Kansas: University Press of Kansas; 2017:179-194.
58. Hill UoCC. *COVID-19 Prison Project [White Paper]*. Princeton, N. J.: Robert Wood Johnson Foundation; 2021.
59. *The COVID-19 Prison Project*. Princeton, N. J.: Robert Wood Johnson Foundation; 2021.
60. Anandarajan M, Hill C. Data quilting: Art and science of analyzing disparate data. *Cogent Business & Management*. 2019;6(1):1629095.
61. Rosella LC, Agrawal A, Awadalla P, et al. *Building a data science platform for better health*. New Digital Research Infrastructure Organization; 2020.
62. How Sharing Data Digitally Benefits Health [press release]. Centers for Disease Control and Prevention 2018.
63. Singletary V, Kerr J, McPhillips-Tangum C. *Building Back Better: Transforming U.S. public health data and infrastructure to protect health and achieve health equity [White Paper]*. Public Health Informatics Institute; 2021.
64. Bullard R, Wright B. *Assessing the Value, Best Practices, and Potential Uses of 211 Data to Inform Public Health During the Pandemic and Beyond [White Paper]*. Houston, TX: Texas Southern University Foundation; 2021.
65. Villarroel N, Davidson E, Pereyra-Zamora P, Krasnik A, Bhopal BS. Heterogeneity/granularity in ethnicity classifications project: the need for refining assessment of health status. *European Journal of Public Health*. 2019;29(2):260-266.
66. G. B, Bauskis A, Pachter N, et al. 3-Dimensional facial analysis—facing precision public health. *Frontiers in Public Health*. 2017;3:31.
67. Coljin C, Jones N, Johnston IG, Yaliraki S, Barahona M. Toward Precision Healthcare: Context and Mathematical Challenges. *Front Physiol*. 2017;8:136.
68. Wang YC, DeSalvo K. Timely, granular, and actionable: informatics in the public health 3.0 era. *American Journal of Public Health*. 2018;108(7):930-934.
69. Chan JL, Purohit H. Challenges to transforming unconventional social media data into actionable knowledge for public health systems during disasters. *Disaster Medicine and Public Health Preparedness*. 2020;14(3):352-359.
70. Maxmen A. Why the United States is having a coronavirus data crisis. *Nature*. 2020;585:13-14.
71. Project TCT. Why it Matters. The Atlantic. <https://covidtracking.com/about/why-it-matters>. Published 2021. Accessed January 4, 2021.
72. Hamel L, Kearney A, Kirzinger A, Lopes L, Muñana C, Brodie M. *KFF Health Tracking Poll - July 2020*. Online: KFF; 2020.
73. Wang H, Cleary PD, Little J, Auffray C. Communicating in a public health crisis. *The Lancet Digital Health*. 2020;2(10):e503.
74. Stephens-Davidowitz S. Fifty States of Anxiety. *New York Times* 2016; Sunday Opinion.
75. Jahani E, Sundsøy P, Bjelland J, Bengtsson L, de Montjoye YA. Improving official statistics in emerging markets using machine learning and mobile phone data. *EPJ Data Science*. 2017;6(1):3.
76. Leinen S. *Mobile Lifestyles: The Homeless and Their Cell Phones* [Doctoral Dissertation], Regent University; 2017.
77. Shi L, Stevens GD. *Vulnerable Populations in the United States*. Manhattan, N. Y.: Wiley; 2021.

Transforming Public Health Data Systems

How? The design of the
modern public health data
system

This white paper is part of a series of papers for use by the National Commission to Transform Public Health Data Systems, an independent commission convened by the Robert Wood Johnson Foundation (RWJF). These papers were used to inform Commission discussions and recommendations. The papers represent insights from literature review, a diverse set of stakeholder perspectives, and inputs from RWJF grantee projects related to the topic of the Commission. The white papers were developed in support of the Commission by the RAND Corporation, a nonprofit research organization, within its division, RAND Social and Economic Well-Being. Corresponding authors are [Anita Chandra](#) and [Laurie Martin](#).

**National Commission to
Transform Public Health
Data Systems**

May 2021

Table of Contents

Introduction	2
The Operation of a Modern Public Health Data System	3
1. Governance and data stewardship	4
2. Data sharing	15
3. Standards of access and interoperability	24
4. Privacy and security	36
Conclusion	41
Acknowledgements	41
References	41

Introduction

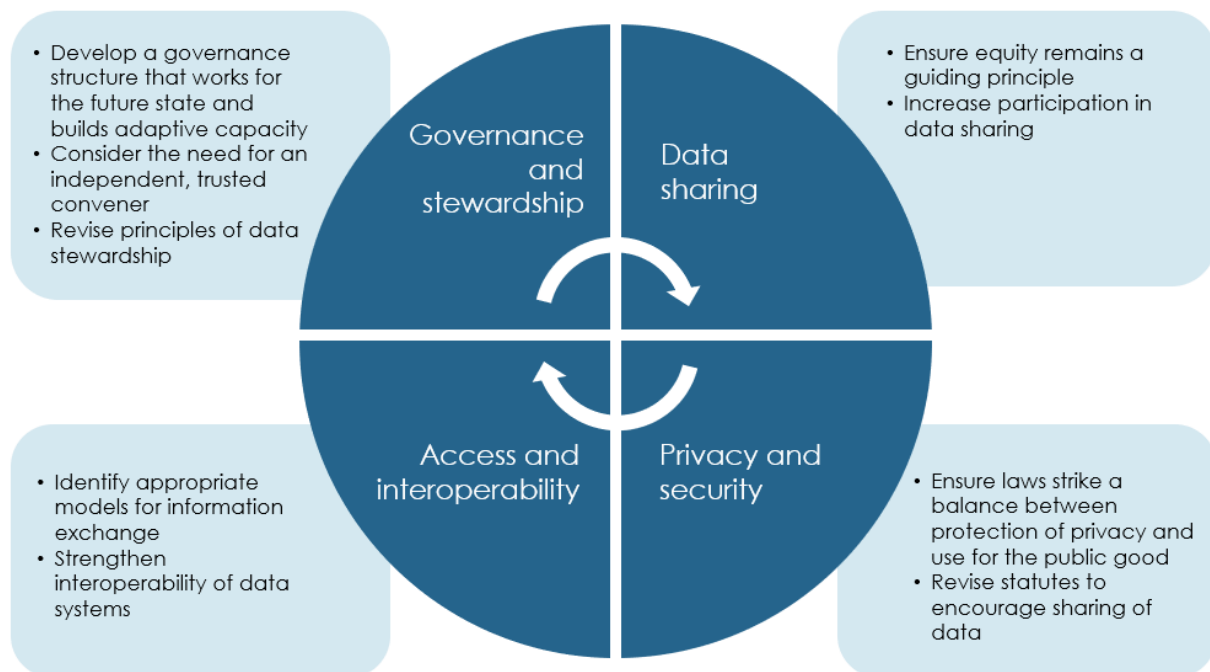
The first two white papers for the National Commission provoked considerations around **why** an equity-oriented public health data system is necessary, and **what** data and measures must be included and elevated to achieve this vision. This third paper in the series focuses on critical features of system design to motivate considerations around **how** the current system could evolve into a modern, equity-oriented public health data system that is not only agile, responsive, and adaptive to emerging conditions, but also leverages diverse data to generate usable findings and integrated knowledge, contextualized through the lens of equity, that are timely and readily available for decision-makers.

Implementation of such a public health data system, however, brings practical challenges related to entrenched interests, legacy systems, antiquated data collection processes, cost, lack of standardization and interoperability, privacy and legal protections, lack of trust, and public health workforce capabilities.¹ As difficult as it has been to witness the failures of our current data system to support a swift, equitable, and coordinated response to a worldwide pandemic, and to watch the erosion of the public's trust in both data and data stewards, this crisis continues to reveal a window of opportunity to reimagine what an equitable public health data system could and should be, to leverage current and critical data modernization efforts, and to think carefully about how to achieve and sustain that vision. The tensions and tradeoffs are clear in the discussion of **how**, because of the push and pull that can happen in these data modernization discussions—while the urgency to get something done can be practically valuable, these efforts stop at incremental or minimalist changes that can be implemented quickly. On the other hand, the nation can think in transformative ways, but that can feel untenable or ambiguous at times, particularly if the United States must reimagine core tenets like purpose and consider widespread systems change.

The Operation of a Modern Public Health Data System

This paper summarizes core issues, innovations, and considerations from the health system and other systems that may provide insight as to a viable path forward, while still integrating and pushing the boundaries of system-minded and transformation-minded thinking, as defined in the preamble. This paper is organized around four major issues: 1) governance and data stewardship, 2) data sharing, 3) access and interoperability, and 4) privacy and security (**Figure 1**). As noted in the prior white papers, this paper is also informed by stakeholder input, literature review, and grantee insights. The examples and exhibits included in this paper are meant to be exemplar, rather than exhaustive.

Figure 1. Equity-oriented data system: Core features and considerations



1. Governance and data stewardship

Governance refers to the structures and processes by which people in societies make decisions and share power, creating the conditions for ordered rules and collective action. Governance over such a large and variable system has inherent challenges given the diversity of actors and interests, which are not always aligned. However, parallels may be drawn, and insights gleaned from the governance of two other systems that share similar characteristics: governance over natural resources and the governance of space. In each of these cases, there is an underlying premise that a governance model is needed to guide collective action around a shared asset (e.g., natural resources, space, data) where stakeholders have diverse interests, “ownership” of the asset is less clear, and the asset can be considered a collective good.²⁻⁴ Whether and how to develop an overarching governance system requires consideration of different types of governance models and how to catalyze convening and buy-in around the optimal governance structure. Related considerations include who could serve in a convening role for an equity-focused public health data system and tenets of data stewardship required to keep pace with rapidly emerging technologies for data collection, analysis, and use of data.

1.1 Key Action: Develop a governance structure that is equitable, leverages the diversity of data across sectors, supports timely decision-making at multiple levels, and builds in adaptive capacity.

Given the rapid pace of innovation in information technology and changes in what, how, when, and why data are collected, and by whom, an important consideration will be not just what governance structure works best to guide the current public health data system, but how to ensure that any governance structure adopted is well suited to iteration and able to address future health data considerations.⁵ Given the data system principle around adaptability to emergent threats (preamble, Principle 2) as well as recognition that modern also means agile, a more flexible governance structure is key. Related to this approach to governance are decisions around binding versus non-binding governance structures, centralized versus decentralized models, and the potential for adaptive governance.

A strong governance structure could prevent the use of public health data from becoming what is known in economic science as a tragedy of the commons, where “individual users have open access to a resource (in this case, data) unhampered by shared social structures or formal rules that govern access and use, act independently according to their own self-interest, and contrary to the common good of all users, cause depletion of the resource through their uncoordinated action.”⁶ Within a public health data system, depletion could occur if individuals or other data owners stop sharing data due to the way in which it is being handled, used, and protected. At the same time, there are concerns that preemption, a “legal doctrine that allows a higher level of government to limit or eliminate the power of a lower level of government to regulate a specific issue,” and particularly state preemption in the time of COVID-19, has harmed local public health efforts and worsened health and economic inequities, elevating a need to think carefully about the relationship between state and local governments and governance models.⁷

1.1a. Consideration: *Non-binding governance holds innate appeal, but relies heavily on standardization, interoperability, and buy-in of diverse stakeholders.*

An important consideration is the extent to which public health data governance should be binding or non-binding. This distinction refers to whether there are repercussions for not following data governance standards and regulations. Some aspects of data governance may be binding, such as with data security or privacy, where there are consequences for mishandling of data. Non-binding governance may exist in the form of guidelines or tenets, which may be useful particularly when consensus on details cannot be achieved and there is an interest in not being overly prescriptive or adding layers of complexity to data sharing activities.³

Although there is innate appeal in non-binding governance, development of a shared language, definitions, thresholds, and risk levels, among others, becomes critical. Efforts towards standardization and interoperability, therefore, become important considerations in a data governance structure. Taking an example from space governance, for example, it was noted that government and commercial space operators have differing levels of “acceptable risk” for space collision, which also may or may not align with the protection of the space environment.³ In this case, it was recommended that “acceptable collision risk levels...be set by international guidelines or national regulations [to] ensure the safety and stability of our orbital environment.”³

If a convening body is conceptualized as the backbone of any data governance structure, then a set of rules, perhaps along the lines of a Data Constitution, may help to build muscles, giving the public health data system the strength needed to leverage data for the public good.³ In addition, statutes or modifications to existing data sharing regulations or privacy acts could support more timely and proactive analytics for the public good. Although these efforts would require diverse input, political will, and stakeholder buy-in, groundwork has been laid by numerous entities who have identified key tenets for data use that could be leveraged (see **Table 1** below). However, equity is an important consideration as well; despite higher transaction costs, diversity in knowledge and perspective can add value to governance and decision-making, reducing the risk of poor system outcomes.⁸⁻¹⁰

1.1b. Consideration: *Centralized models allow for greater control, but are less efficient, while decentralized models facilitate local response, but lack standardization.*

The debate around centralized versus decentralized governance structures is particularly relevant for the public health data system. Decentralized models push responsibility for data system governance to local levels, where models of data sharing and guidelines for data stewardship and governance are locally determined. Local governance affords an opportunity to build trust in the data system and provide a more agile and customized response to local needs.¹¹ However, the lack of standardization in governance across decentralized data systems poses barriers for more widespread data sharing with the state and federal government and for pooling data across local jurisdictions to assess emerging health needs. A centralized data governance structure allows for greater control over standards, but criticisms of this model are that it is less efficient (particularly at a national level), may be slower to react to problems, and may be more open to corruption or political influence. This

tradeoff between centralization to ensure standardization and decentralization to allow for local freedom will be an important consideration for the public health data system.¹¹

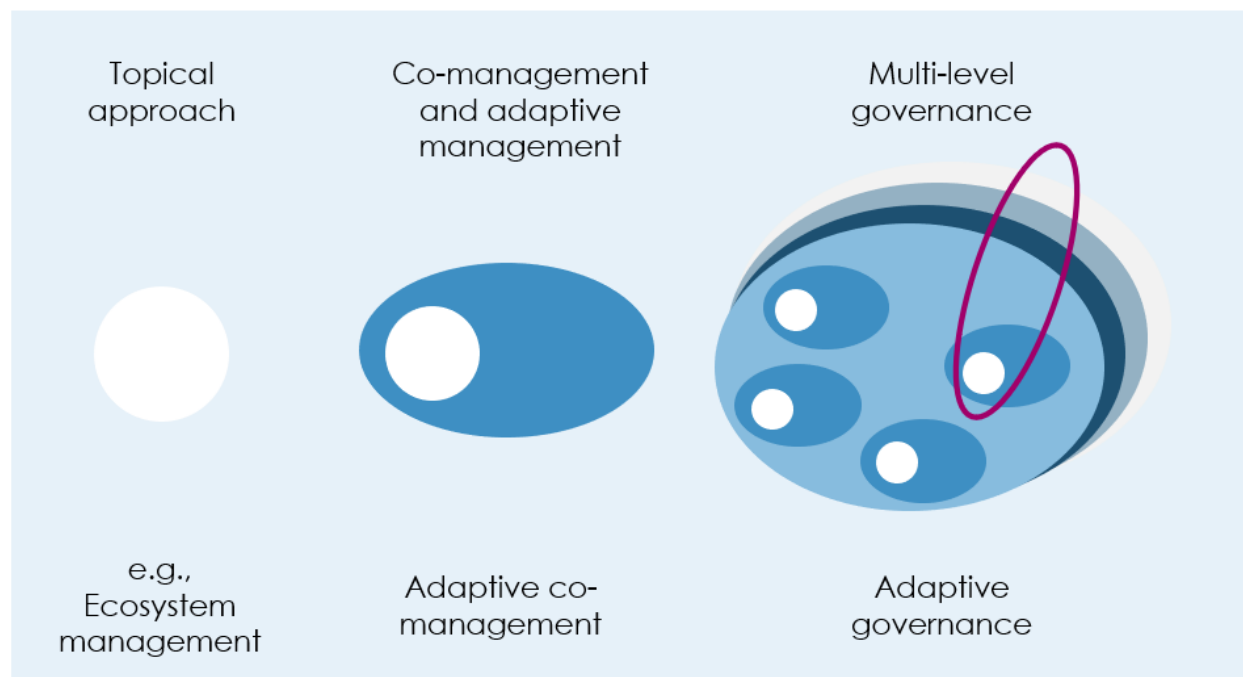
1.1c. Consideration: *The 'nesting principle' for robust governance of large-scale common-pool resources may be applicable to public health data but requires substantial coordination and collaboration.*

An analysis of sustainable models of common-pool^a environmental resources indicated that the governance systems were often “organized in multiple layers of nested enterprises” (see **Figure 2** below and preamble Figure 5).⁴ Although developed within the field of environmental management, this ‘nesting principle’ for governance of large-scale common-pool resources may be relevant for the public health data system. One potential advantage of nested governance is that “smaller organizations become part of a more inclusive system without giving up their essential autonomy,” which can contribute to the robustness of the system.¹² For example, a nested system allows for more tailored (decentralized) governance at a local level (lower-level) while simultaneously working with higher governance levels that may be capable of addressing problems that exceed the capacity of lower-levels to address on their own.¹² However, one important consideration of this model for the public health data system is determining what those levels of governance look like: are they organized by geographic region? Population? Topic (privacy, standardization)?

Another important consideration is deciding which tasks will be handled at which levels of governance. This requires substantial collaboration and coordination and revolves around the notion of subsidiary, meaning that “any particular task should be decentralized to the lowest level of governance with the capacity to conduct it satisfactorily.”^{2,13} This assignment of tasks is more difficult than it sounds, however, as it requires appraising capacity of various levels of governance, and we know that capacity varies greatly across the field of public health even within similar levels (e.g., local public health departments). A final consideration is how to manage interactions among the various levels of governance to ensure that any governance structure does not undermine work noted above to strengthen interoperability, data sharing, and equity.¹³

^a A common-pool resource is a resource that benefits a group of people, usually by consumption. Common-pool resources (e.g., forests, fisheries) are susceptible to overuse and provide diminished benefits to everyone if each individual pursues his or her own self-interest.

Figure 2. Nesting of governance within complex adaptive systems



Adapted from: Rechkemmer, A. (2012) Climate Change, Development and Health: The Role of Adaptive Governance¹⁴

1.1d. Consideration: *Adaptive governance affords flexibility in governance over time but relies on a learning-based collaboration across stakeholders that could prove challenging.*

Adaptive governance refers to an “evolution of the rules and norms that promote the satisfaction of underlying human needs and preferences given changes in understanding, objectives, and the social, economic, and environmental context.”⁹ In the context of a public health data system, this theory of governance may afford flexibility as the needs and preferences for health data evolve, new types of data become available (e.g., from private industry), and technological advancements become widespread. For example, in an adaptive model, a public health data system may start in a decentralized model and be governed locally but also has the capability to adapt quickly to a centralized model of governance during public health emergencies, when collaboration and more centralized approaches are helpful for coordinated response. Adaptive governance provides space for the detection of opportunities and problems as they arise and the modifications to be made to address emerging needs.³ Because adaptive governance reflects a learning-based collaboration across multiple stakeholders, its success may be dependent on the willingness of stakeholders to take on and give up responsibility as the system evolves and understand opportunity contexts that may signal the need for adaptation. While this may be easier to achieve in the context of preserving a local natural resource, where adaptive governance has been used successfully, this may be more challenging for public health and with stakeholders who have a diverse and vested interest in health data.¹⁵ However, aspects of adaptive governance hold promise for informing a potential governance structure of a public health data system.

1.2 Key Action: Consider the need for an independent, trusted convener, who or which can bridge sectors to facilitate development of an equity-oriented data system.

In examples of governance over natural resources, the move towards adaptive governance was often catalyzed by a crisis that mobilized individuals to think differently, build trust, and develop a system vision that combined seemingly opposing viewpoints (e.g., conservation and development).¹⁵ It could be argued that the COVID-19 pandemic has served as a similar catalytic awakening for the public health data system. This crisis has mobilized action at the local, state, and federal level and coalitions and collaboratives already focused on interoperability and data sharing have a renewed sense of urgency. A next step could be the establishment of a convener, who or which can harness the collective action and power of thought occurring across these groups to develop a governance structure and path forward that facilitates a vision of an equity-oriented public health data system.

1.2a. Consideration: *A convener or convening body could support the evolution of the public health data system, but it is not clear who (or what group) should serve in this role, and not all stakeholders believe it is necessary.*

In 2007, the Agency for Healthcare Research and Quality (AHRQ) released a request for information (RFI) on the concept of a national health data stewardship entity (NHDSE)¹⁶ with the proposed characteristics of an NHDSE summarized in **Exhibit 1**. Although framed with a particular focus on health care data, the proposed mission, precepts, and characteristics of the NHDSE offered by AHRQ for comment have widespread applicability to a broader health data system. A summary report of the comments received, however, suggested that there was little consensus in understanding around the concepts of data stewardship and governance, and AHRQ received both support and opposition for the idea.¹⁷

As summarized by Rosenbaum, “proponents viewed a broadly conceived steward as offering an essential oversight mechanism for health data issues, organizing the various data collection, aggregation and sharing systems, assuring privacy, empowering consumers, and fostering collaboration among stakeholders. Opponents objected to a stewardship entity as an unnecessary competitor; others raised concerns about the absence of a clear legal authority on which to act.”¹⁷ Although AHRQ did not move forward with the development of an entity at that time, future efforts to identify or develop a convener could benefit from a deeper understanding of the feedback received and challenges faced, as well as an assessment of whether overarching objections still hold given the evolution in technology, data security, and understanding of the value of data sharing for both individual health management and the public good that has occurred over the past 15 years.

Exhibit 1: Proposed characteristics of NHDSE offered by AHRQ in 2007

1. Objective
2. Independent
3. Knowledgeable
4. Responsive
5. Trustworthy
6. Adaptable
7. Transparent
8. Timely
9. Collaborative
10. Sustainable

No clear individual or collaborative body has emerged as an obvious choice for a health data system convener. It has been suggested that perhaps the Surgeon General might serve in this role, setting expectations for health progress as a country and prioritizing a smaller set of public health measures for a more unifying and inspiring national health vision, which would then trickle to cabinet-level agencies. Although promising for visibility, it may be important for the convener or body to be independent and accountable to the public with full moral and decision-making authority, and not subject to political isolation or pressure. Others have suggested that the field of public health, or the Centers for Disease Control and Prevention (CDC) specifically, should be the convener, although the current pandemic has recently eroded trust in these institutions, and, again, there may be concerns around political pressure and financial ties to federal and state government that may impact objectivity.¹⁸ In the space governance example, industry associations play vital roles in developing industry consensus, sharing best practices and expected norms of behavior, and promoting industry objectives and aspirations.³

An important clarification in the framing of the convener role will be to distinguish it from management, which refers to the process of decision-making and control over resources within a given institutional setting.⁹ A convener, however, connects to management across diverse sectors and utilizes indirect control in the form of governance to align decisions and behaviors.⁹ An important characteristic of the convener, therefore, is the ability to work across and meaningfully engage diverse sectors. In addition to public health and health care data, data from social services, education, justice, the environment, and private industry, among others, hold immense potential for public good and the development of equitable, proactive solutions to strengthening the health and well-being of our nation. There may be an opportunity to consider whether a cross-system convener or governing body may be a viable solution.

1.3 Key Action: Revise core tenets of data stewardship and data governance to keep pace with rapidly emerging technologies for data collection, analysis, and reuse.

Success of the broader public health data system to advance health equity and improve population health and well-being is heavily dependent on data stewardship. Stewardship can be thought of both as an institutional commitment to and a collection of methods for data management that address the acquisition, storage, and aggregation of data for scientific and societal benefit, while protecting against privacy and security breaches and misuse.¹⁷ Data stewards have a fiduciary (or trust) level of responsibility toward the data.¹⁷ Data governance, described above, is the process (e.g., broad policies, approaches) by which stewardship responsibilities are carried out. Data stewardship and governance, however, have not kept pace with rapidly emerging technologies that have an impact on data collection, access, and use. Issues related to data ownership, access, and trust remain unclear; there is a need for methodologic advances to facilitate use of data for public benefit; and there is an opportunity to reassess the necessary competencies of the public health workforce with respect to data stewardship and reuse of data for public good.

In recent years, cross-sector efforts and strategic plans have offered core tenets and frameworks for advancing data modernization and stewardship. These core tenets, summarized in **Table 1**, cover issues of equity, governance, infrastructure, innovation, interoperability, partnerships, and data

security. Although each was written through a specific lens and, in some cases, as a guiding document for a specific organization, collectively they provide important insight on key data stewardship tenets for a modern, equity-oriented health data system. Included in the table are core tenets drawn from the *Centers for Disease Control's (CDC) Data Modernization Initiative*; the *Office of the National Coordinator for Health Information Technology (ONC) 2020-2025 Federal Health IT Strategic Plan*; the *Council for State and Territorial Epidemiologists' (CSTE) report, "Driving Public Health in the Fastlane"*, *The National Committee on Vital and Health Statistics*, the *World Economic Forum*, and the *American Medical Informatics Association*.¹⁹⁻²³

Table 1. Core tenets of recent data use frameworks

Categories	Core tenets	CDC 2020 ^b	ONC 2020 ^c	CSTE 2019 ^d	NCVHS 2014 ^e	WEC 2013 ^f	AMIA 2012 ^g
Equity	Equity		X				
	Steward responsibility	X	X		X	X	X
	Representativeness						
Governance	Individual Rights				X		X
	Patient-centered		X				
Infrastructure	Information Technology	X		X			
	Data Quality			X	X		X
	Workforce education/innovation	X		X			X
Innovation	Promote innovation	X	X	X			X
Interoperability	Regulatory standards						
	Data standards			X			
Partnerships	Public-private partnerships	X	X	X			X
	Public-community partnerships	X	X	X			X
	Intergovernmental collaboration	X					X
Security	Accountability/Enforcement		X		X	X	X
	Transparency						X
	Privacy/Protection	X	X	X	X	X	X
	Public trust						X
	Access	X	X				X

^b CDC. (2020). Data Modernization Initiative. <https://www.cdc.gov/surveillance/surveillance-data-strategies/dmi-investments.html>.

^c HHS Office of National Coordinator for HIT. (2020). "2020-2025 Federal Health IT Strategic Plan." <https://www.healthit.gov/topic/2020-2025-federal-health-it-strategic-plan>.

^d Council of State and Public Health Epidemiologists. (2019). "Driving Public Health in the Fastlane".

https://cdn.ymaws.com/www.cste.org/resource/resmgr/pdfs/pdfs2/Driving_PH_Print.pdf.

^e National Committee on Vital and Health Statistics. (2014). "Health Data Stewardship: What, Why, Who, How". <https://ncvhs.hhs.gov/wp-content/uploads/2014/05/090930lt.pdf>.

^f World Economic Forum. (2013). "Unlocking the Value of Personal Data."

http://www3.weforum.org/docs/WEF_IT_UnlockingValuePersonalData_CollectionUsage_Report_2013.pdf.

^g Hripcsak, G., Bloomrosen, M., FlatleyBrennan, P., Chute, C. G., Cimino, J., Detmer, D. E., ... & Keenan, G. M. (2014). Health data use, stewardship, and governance: ongoing gaps and challenges: a report from AMIA's 2012 Health Policy Meeting. *Journal of the American Medical Informatics Association*, 21(2), 204-211.

1.3a. Consideration: *Lack of clarity around data ownership can bring cascading challenges related to who can access and share data, with whom, and for what purpose.*

The issue of data ownership is perhaps one of the biggest challenges to address in the context of data use for public good. There is a case to be made that data are owned by the entity that paid for or authorized its collection (e.g., government), such as with state or national surveys. Another position may be that the entity who collected the data (e.g., health care providers, researchers, private industry) own the health data they collected, given the resources and time spent collecting it and the competitive advantage such data may provide within their industry. A third position is that data are owned by the individual providing it, particularly given that in many cases the data contain personal information about that individual.

In the context of a public health data system, the issue of ownership becomes particularly important, given that data ownership is likely to vary depending on the type of data, method of data collection, whether consent has been provided, and for what purpose. As data get pooled and shared more readily, the question of ownership becomes even more important because it points to who can grant permission to access the data and for what purpose and who should profit (financially or otherwise) from the use, sharing, or even selling of data. From an equity perspective, the question of data ownership points to who has the “right” to tell the story stemming from the data, and whether there should be special provisions for where and how that story is told, particularly for special populations.

Exhibit 2: Legal authorities and barriers to COVID-19 data disaggregation

The **Network for Public Health Law (NPHL)** is undertaking a study to understand the authority of health departments to collect race and ethnicity data, where there are legal barriers to their reporting it, and obligations to share data in response to requests made by the public pursuant to records access laws.²⁴ Surveying the laws of 10 diverse states, NPHL found that health departments are authorized to collect data about race and ethnicity, and although the information reported to health departments is confidential, de-identified data may be shared and published. While all 50 states have laws giving the public the right to access government data, records access laws differ slightly in what they exempt from disclosure, and many require the state to balance privacy interest of the people whose information they possess against the public’s interest in disclosure. Early conclusions from this work suggest that the individuals about whom COVID-19 data relates have legitimate privacy interests that health departments must respect. If an individual can be re-identified from disaggregated data, its release might constitute an unwarranted invasion of privacy, particularly in a state with a constitutional or statutory right of privacy.

The lack of clarity around data ownership raises important considerations related to data stewardship, access, and use cases for the data. The first relates to whether there should be a review process of not only the use case, but also of the findings to help ensure that the data are being used for public good. This may facilitate the rebuilding of trust among populations who have been harmed by unethical research practices and provide an opportunity to engage individuals proactively and meaningfully in the knowledge generation process who have been previously excluded. Another consideration pertains to Institutional Review Board (IRB) oversight. Typically, the governing IRB is located at the institution where the data are being housed, analyzed, and reported. However, in new models of data access, where the data stay behind firewalls and the queries go to the data, it could be that the responsibility for review and approval of data use is better left to the data steward, rather than an external IRB. Although shifting this responsibility to the data steward would add burden to that role, it may also facilitate equitable access to data for users who are not affiliated with an institution that has an IRB.²⁵

Another consideration is the need to guard against the use of a public health data system for financial gain. If the system is successfully built on the premise of data reuse for public good, knowledge that the data were used for financial gain or competitive advantage could quickly undermine the trust of the public and participating sectors. Equity in data access will also be an important consideration, to ensure that the system is not structured in such a way that gives individuals or entities with more resources access to more information. Given that many data sources require purchase of a data license or are cost prohibitive for less-resourced public health departments, it will be important to develop strategies (e.g., free to non-commercial users, or sliding fee scale) that ensure equitable access to the data.

1.3b. Consideration: *Despite promising developments, public health is not yet consistently using methodologies to efficiently leverage the diversity and volume of data available.*

The integration of data from a wide range of sources, the sheer volume of health-related data being generated including from social media, sensor technology, and new sources of audio and video data, and increased computing power and technological innovation hold great promise for the development of proactive, data-driven solutions to improve health, equity, and well-being.²⁶ With these changes, however, comes a need for new methodologies to analyze the data efficiently, cost-effectively, and accurately. For example, methodologies for creating robust national estimates from EHR data have yet to be developed.²⁷ There is also an opportunity to leverage and refine innovations like artificial intelligence, machine learning, natural language processing, and other methodologies for predictive analytics and the generation of actionable solutions.

To ensure that data governance and protection of privacy keep up with the pace of IT innovation, methodologic advancements with respect to the de-identification of data that lessen the likelihood of re-identification could also be explored.²⁸ Methodologic approaches to allow for the disaggregation and analysis of data by geography or population characteristics could also have far-reaching implications for advancing health equity. Currently, data are either omitted or aggregated with other data if the number of individuals within a “cell” are too small, for the purpose of protecting the privacy of individuals. Oversampling of specific populations or geographies is one strategy for overcoming this challenge, although this approach is resource intensive and may not be sustainable.

Approaches that utilize data from multiple years or introduce noise in the data have also been leveraged to examine populations of interest with smaller numbers, but each has its corresponding limitations.

Exhibit 3: Data platform viability

The **Health Care Cost Institute** is assessing the feasibility of developing a permanent platform for making private data available to researchers for noncommercial use. Through their work, they are developing a conceptual framework and learning from both data suppliers and data users about potential facets of such a platform including: perceptions of infrastructure, access and governance, legal frameworks, data linkages, privacy and de-identification of data, and researcher collaboration and standards of data use.²⁹

In 2012, the Obama administration invested \$200 million in the National Big Data Research and Development Initiative to harness emerging structured and unstructured data sources.³⁰ In 2016, a strategic plan for research and development of ‘big data’ was developed by the Office of the President, National Science and Technology Council to take stock of what changes would be needed for the federal government to keep pace with the growing array of disparate, distributed, dynamic, and real-time data sets.³¹ This plan included activities to advance new analytic techniques; explore the trustworthiness of data; enhance the cyber infrastructure; explore data sharing, privacy, security, and ethics of using ‘big data’; and to determine needed workforce capabilities. Building on this work, the 2020 Federal Data Strategy lays out a series of principles, practices, and actions to improve the federal government’s approach to stewarding data and using data for public good.³² In 2020, the CDC received \$500 million through the Coronavirus Aid, Relief, and Economic Security (CARES) Act to strengthen public health data and surveillance infrastructure of the United States through “modern, interoperable, and real-time public health data and surveillance systems that will protect the American public.”³³

Figure 3 describes the specific limitations of current public health data systems and opportunities for modernization that the CDC is seeking to align their investments to address. These opportunities include using data for predictive modeling, rapid analytics with real-time data, and developing interoperable systems that translate to action. Consistent across these federal plans is the identified need for modern information technology, a skilled workforce, and data governance policies to improve accountability and transparency when using new structured and unstructured types of data.

Figure 3. Realities and opportunities of the Centers for Disease Control and Prevention's data modernization initiative³⁴



1.3c. Consideration: The current public health workforce may not have the skills and competencies necessary to operate in a modern information age.

There is growing concern that the public health workforce, in general, may not have the skills and competencies necessary to efficiently operate in a modern information age. Individuals skilled in informatics are choosing private industry over public health given the opportunity for better pay and advancement. In some cases, there is no job classification specific to informatics, making it difficult for public health departments to attract the right candidates.

The public health workforce could benefit from expanding analytic competencies through a multi-pronged approach. First, schools of public health could review course offerings and strengthen requirements around informatics and analysis of complex data for all students (not simply those

pursuing a more quantitative track). Second, there may be an opportunity for a narrative shift that highlights important challenges and unique opportunities in public health that require top talent. Although some individuals may still opt for a higher salary, others may be enticed by the opportunity to be trailblazers in an emerging area. Third, workforce exchanges, cross-sector models of job-sharing (e.g., where an individual splits time across employment at a public health department and academic institution or private company), and models where certain data analytic tasks are completed by a centralized regional or national group, rather than conducted at the local level, for example, could help build competencies and fill this gap.

These points raise a larger consideration around whether there should be formalized competencies for data stewards more broadly, and whether public health professionals are equipped for this role. Data stewards of the future, for example, will need to have not only data governance expertise, but also expertise in data management and informatics, and hold many of the same characteristics outlined above for a system convener such as independence, objectivity, and responsiveness. There may be an opportunity to consider in more detail what competencies the data stewards of the future will need, how to ensure we are collectively developing the data stewards of the future through education and training programs, and whether there is an opportunity to standardize the position through a certification program or equivalent.

2. Data sharing

As noted in the What? paper, the idea that an individual's health, social service, education, and other data can be used not only for the purpose of improving the health of that individual, but to improve the health of the population, is increasingly acknowledged but is not currently being optimally collected or utilized for decision-making. Those not in the field of public health often struggle to understand what makes population health different from other conceptualizations of health, why non-clinical data are important to understanding the health of populations, and how individual-level data can inform the health of populations. Without a clear understanding or recognition of the value of such data for public good, the interest and willingness of other sectors to participate in data sharing efforts may be dampened.

In general, reuse of existing data for the purposes of public good (e.g., strategic planning, surveillance, research) is considered advantageous in that it saves resources, reduces duplication of effort, minimizes burden on individuals, and can result in a larger, more complete set of data from which to draw conclusions.³⁵ Sector-spanning terminology like “data aggregation” may be useful for engaging sectors less familiar with public health in that it helps to clarify how the data may be used to shape action, may alleviate concerns around privacy, and elevates the idea that our public health data system should utilize information about all individuals, not just those seeking medical care. This section discusses two considerations related to data sharing: 1) how to ensure equity remains a core tenet in data sharing efforts, and 2) how to increase participation in data sharing efforts.

2.1 Key Action: Ensure that equity remains a guiding principle in data sharing efforts.

The events of 2020, including calls for social justice and the COVID-19 pandemic that disproportionately affected people of color, presents an opportunity for careful and perhaps uncomfortable self-reflection on the reality that a stronger, more integrated public health data system has not been achieved, despite numerous calls and initiatives to build cross-sector collaboration, facilitate community-partnered work, and overcome technical challenges. As noted in the preamble and the Why? paper, there is growing recognition that this lack of progress is due, at least in part, to deep system inequities and core values related to self-preservation that make those in power reluctant to share resources, data, and power. For example, current incentive structures in the health care system are tied to individuals seeking medical care rather than healthy communities, and making data (e.g., on race or ethnicity) more readily available to a wide range of stakeholders could shed light on inequitable practices and result in unwanted negative attention.

Instead of working around the edges, what elements of system design will generate a truly transformative shift in our public health data system with respect to equity? Three considerations are discussed in this section: 1) how to uniformly collect data that includes special populations and facilitates data disaggregation, 2) how to shift the paradigm of engagement from community participation to co-creation, and 3) how to ensure equitable access to public health data. Additional equity considerations are discussed in relevant sections throughout this paper.

2.1a. Consideration: *Challenges to an equity-oriented public health data system include an inability to disaggregate data and the frequent absence of underrepresented populations from data collection activities.*

The United States does not consistently or uniformly collect data critical to understanding the health, well-being, or lived experiences of millions of individuals living in the United States, severely limiting our ability to develop data-driven solutions for minorities or populations that have vulnerabilities.^{36,37} As noted in the Why? paper, even in the midst of a worldwide pandemic, as of August 2020, basic demographic characteristics like race and ethnicity were not collected or were unknown for 51 percent of the cases, masking severe disparities in case and fatality rates early on and delaying the development and implementation of data-driven solutions.³⁸ As noted in the What? Paper, the lack of granularity in currently used data and measures poses significant challenges to advancing population health and well-being.

Exhibit 4: Identifying systems-level barriers to data disaggregation

The **Croal Services Group, LLC** conducted a series of in-depth interviews on the logistical barriers to collecting and reporting race and ethnicity public health surveillance data.³⁹ The team identified numerous barriers shaped by the sociopolitical nature of race and what they note as the historic and contemporary failures of leaders and systems to prioritize data equity. Leadership may be the most important factor because data equity and the collection of complete and disaggregated race and ethnicity data can be (de)prioritized at all levels. Improvements in race and ethnicity surveillance data were exemplified by leaders making it a priority during the COVID-19 pandemic, as demonstrated in their case study on North Carolina, which achieved a 98 percent completion rate for race and ethnicity data.⁴⁰ Barriers included a lack of universally applied federal standards (e.g., from Office of Management and Budget) and data categories that are inadequate for data equity given lack of disaggregation. The chronically underfunded and under-resourced public health system is also a barrier, given that initiatives to improve and expand data equity may compete with other priorities for limited resources. The authors note that data equity requires deep investments in engaging with communities to create meaningful categories; training and data systems modernization; and changes to race and ethnicity data collection standards, policies, and practices.

Despite advancements in survey methodologies, public health surveillance efforts, and sharing of health data through electronic health records (EHR), critical equity questions remain: Who is *not* included in those data? And who have we missed? An important consideration of developing an equity-oriented public health data system that produces actionable information to address inequities will be assessing who is and who is not represented in current health data and developing solutions to ensure accurate and complete representation. This could include strengthening representation of individuals who do not regularly see a provider (or see a provider that does not utilize an EHR or actively share data), which may be particularly valuable for detecting the early spread of infectious diseases or other health-related trends.⁴¹ It also suggests a need to strengthen inclusion of individuals with cognitive, developmental, or other disabilities; children; individuals in group home settings; those in prison or jail; homeless persons; and undocumented residents, among others. Data collection from these and other special populations poses additional challenges due to special protections, consent and assent procedures, and necessary collaboration with other systems (e.g., justice), which may explain why they tend to be overlooked.

New protocols and methods of data collection could be developed to ensure that the public health data system includes data on all individuals, not just those for whom data are easier to collect. At the same time, more advanced data collection methods and methodologies must also be critically examined as they could potentially exacerbate inequities in data representation if, for example, existing data or data collection methods designed for one population are applied to another without consideration of whether such approaches are fully appropriate or representative of new spaces,

cultures, or populations.⁴²⁻⁴⁵ Another potential challenge may arise in the development of methodologies, including algorithms, where developers establish benchmarks for acceptability. An accuracy level of 90 percent may be considered strong, but results in data that is more precise (useful) for some groups and less precise for others.

2.1b. Consideration: *Although community engagement efforts are expanding, practitioners have not sufficiently moved from episodic community participation to full participation and co-development.*

Most health data systems were developed without the input of individuals whose data are being collected, and largely for the purposes of financial accountability, rather than improving the population's health. Development of an equitable health data system could benefit greatly from meaningfully and authentically engaging diverse stakeholders (including individuals) in the design process. That tenet described in the Why? paper, "Nothing about us without us," remains resonant.

Exhibit 5: Incorporating data from residents

Streetwyze⁴⁶ is developing a new technology platform for community residents so they can gather and incorporate their own data into publicly available, place-based data to construct more complete pictures of how conditions of place impact health. These data can be used to demonstrate how key challenges to achieving health equity are experienced and to encourage broader participation, especially by those who represent vulnerable populations, in the co-creation of solutions to the most significant challenges.

Initially, data collection efforts and corresponding solutions were developed with minimal to no community input. Recognition that these solutions often fell short of expectations led researchers and practitioners to seek input from community members via focus groups and interviews in the design process, although criticisms of this approach are that it obtains limited input and lacks meaningful, sustained engagement with the community. In an effort to develop data collection efforts and health solutions that were more in line with community needs, promoted meaningful engagement, and built community capacity, the field shifted again towards the utilization of community-based participatory research (CBPR) methods.⁴⁷ However, adherence to CBPR principles and methods remain quite variable, and it has been argued that this approach does not go far enough to dismantle the power imbalance between professionals (e.g., researcher, practitioner) and community residents, particularly when it comes to interpreting, contextualizing, and translating data into information and insight.

An important consideration is whether there are opportunities to further shift this paradigm towards co-development that ensures the community has power to drive local solutions. True transformation to an equity-oriented data system necessitates tough conversations around issues of power, expertise, and trust. For example, whose goals or needs are reflected in design decisions specific to the public health data system? Whose expertise is truly valued (e.g., those with lived experience or

those with advanced degrees) in decisions around what, when, why, and from whom data are collected? Who becomes more or less “visible” in the system, and what might the consequences be (good or bad) for vulnerable populations? If approached through the lens of co-creation, a stronger, more equitable data system could emerge.

Exhibit 6: Grassroots action infrastructure for public health

Community Science is examining the role of national grassroots action infrastructure (grassroots organizations and intermediaries) for public health and social justice. One of their case studies is examining the work of **Citizens Committee of New York City (CitizensNYC)**, an intermediary organization whose mission has been to help New Yorkers come together and improve the quality of life for residents in their neighborhoods, and to serve as a community hub that promotes philanthropic justice by being a “connector rather than a gatekeeper.” CitizensNYC fielded a survey through an extensive network of grassroots community groups during the pandemic and received over 800 responses from majority Black and Hispanic residents, which provided an urgent list of community needs. In addition to using these results to shape their grantmaking efforts, an interactive mapping tool was also made widely available to support local donor decision-making, and they are striving to sustain active community engagement through new platforms that have allowed CitizensNYC to hear directly from grantees about what is happening in their communities.⁴⁸

Exhibit 7: Shaping an equitable data ecosystem through community partnership

Although early in the process, the **Rising Equitable Community Data Ecosystems (RECoDE) Learning Council** is encouraging genuine participation by communities through a series of interviews and workshops to support the creation of more equitable, accessible community data ecosystems, which will require shifting power and resources to communities, fostering a sense of ownership, and ensuring that data systems have feedback loops that enable communities to use and benefit from the data they share.⁴⁹

2.1c. Consideration: *Public health data are difficult to access for many stakeholders.*

While data sharing is an important aspect of a public health data system, it is important to reflect on who has timely access to those data. Representatives from non-health sectors have noted that they are often asked to share their data, with little opportunity to leverage the combined data for their own work. Even within the public health system, COVID-19 brought similar challenges as local health departments fed data into state data systems but were not able to access or analyze those collective data to understand state and local trends for local decision-making.⁵⁰ Access to some types of data,

such as exact dates or geographic information, are accessible only through restricted-use files that users access after a detailed application process, purchase of a data license, or travel to a research data center.^{51,52} These expenses may be difficult for smaller organizations to afford, further perpetuating inequities in data access. Such one-way data flows, and structural barriers to data access, pose important challenges for the development of local data-driven solutions and are important considerations for a new public health data system.

Who has access to public health data is not only an important ethical question, but one that also has practical implications for the field of public health. Making public health data more widely available would allow a range of stakeholders (not just public health departments) to leverage the data for local decision-making and provide an opportunity for innovation and new models of working together that may address resource and personnel constraints. It is not clear, however, whether public health data should be viewed as “open access,” where anyone can access and analyze the data, regardless of their analytic capabilities. Questions include:

- Should there be certain requirements to gain access to data?
- Should these requirements differ for public health, researchers, community groups, journalists, or individuals?
- What should those requirements be and how should we ensure those requirements do not create inequitable barriers to accessing data?
- How do we balance equitable access with the potential for misinformation or the potential for harm that could come to groups represented in the data?

Answers to these questions are also relevant to considerations around governance and stewardship, discussed above.

2.2 Key Action: Increase participation in data sharing efforts.

Although there are many promising examples of how communities have come together to share data across sectors,^{53,54} these examples are more often the exception, rather than the rule. Increasing participation in data sharing efforts will require careful consideration of two challenges that have hindered widespread data sharing: 1) variable infrastructure and limited resources, and 2) limited emphasis on the value of data for common good.

2.2a. Consideration: *Although the federal government has implemented policies to promote data sharing of certain clinical and research data, widespread data sharing with public health is hindered by variable infrastructure and limited resources.*

As noted in the What? paper, our currently fragmented public health data system is an eclectic mix of legacy systems, disease-specific systems, and unique measures that will require considerable

investment to align and integrate.⁵ As a result, the seemingly simple request of asking organizations to collect and share data on a more granular level (e.g., using more refined race or ethnicity categories), or to share data they have already collected (e.g., social service utilization, or health data obtained from wearable technology), for broader public good can require significant time and resources. Organizations may need to make changes to their data collection procedures, consider implications for informed consent, ensure privacy in the new use case, and ensure processes for secure data transfer and tracking. It is not clear whose responsibility it is to cover these costs. In fact, the significant investment that organizations make in collecting and maintaining their own data can sometimes be used as a rationale for *not* sharing data more broadly, particularly if they do not perceive any benefit in return.

This tension is common in the research community as well, although some federal funders such as the National Institute of Health (NIH) and the National Science Foundation (NSF) have policies in place to encourage data sharing.⁵⁵ The NIH, since 2003, has required that projects receiving over \$500,000 must make final research data “as widely and as freely available as possible while safeguarding the privacy of participants” through the timely dissemination of data in data archives such as the ICPSR, data enclaves, or mixed mode sharing.⁵⁶ Since 2011, NSF has required all proposals to include a data management plan that details a way for principal investigators to share data with other researchers within a reasonable timeframe and at a no more than incremental cost.⁵⁷

In recent years, federal agencies like the Centers for Medicare & Medicaid Services (CMS) have begun to leverage their position as a large funder of health care services to incentivize consistent data collection, interoperability, and data sharing. For example, the Meaningful Use program is an incentive program designed, in part, to promote collection of standardized information and to share relevant data with public health agencies (e.g., data on immunizations, reportable conditions, and other data for surveillance and public health registries).^{58,59} Medicare’s Merit Based Incentive Program (MIPS) is furthering the coordination of EHR interoperability and data sharing.⁶⁰ Although efforts like these provide strong incentives, participation and implementation is far from universal and is heavily focused on the sharing of data collected within the clinical encounter.

Despite federal investment in health care infrastructure, public health has received far fewer federal resources to maintain and modernize the public health infrastructure. For example, from 2012 to 2018, the state public health workforce decreased by about 10 percent, which equates to a loss of about 10 full-time equivalent workers per state.⁶¹ Particularly relevant is a reduction in the public information (-33 percent) and public health informatics (-29 percent) workforce at the state level between 2010 and 2013.⁶² General federal and state funding for public health has also declined over the past decade.⁶³ This has resulted in wide variability in the public health infrastructure on top of what was already a heterogeneous group of local public health departments that vary in size, geographic reach, populations served, location, economic circumstances, governance structure, and governing organization to which they are accountable.⁶⁴ Although recent investments and incentive programs in the health care system have been designed or adjusted, in part, to support data sharing more readily with public health, the public health system may not be able to receive, process, or use

the data being sent. Inequities in the public health infrastructure have critical cascading effects for health equity overall.

2.2b. Consideration: *Limited emphasis on the value of data for common good, including the value of public health data specifically, has hindered data sharing.*

Exhibit 8: Barriers to sharing data across criminal justice and health settings to advance equity

The **University of Arkansas** examined the collection, usage, and sharing of public health data to better surveil and respond to health threats, improve population health outcomes, and reduce health disparities among criminal justice involved populations.⁶⁵ Their work has identified a number of barriers to sharing data across criminal justice and health settings to advance equity. For example, many organizations collect data on important pieces of the overall narrative of incarceration and health but getting buy-in from stakeholders and a commitment to prioritize data systems is a challenge. Many national surveillance systems do not survey institutionalized populations and do not ask about previous experiences of interactions with the criminal justice system. In addition, jail- and prison-based medical records and health data are often housed by private, state-contracted providers, and making inroads with these groups can be challenging. Data quality is also a concern as medical services that are provided by the state or non-profit providers often include data of lower quality and accessing health data through the sheriff's office or a state department of corrections remains a challenge. While electronic health records are more common in larger correctional facilities, many smaller jails still rely on paper records. It is also difficult to link the health of the individual before, during, and after incarceration. Health care use prior to arrest and incarceration is only available if collected upon entry to the facility and is typically self-reported. While it is possible to link post-release health data with diagnoses and health needs during incarceration, such linkage is an arduous task and must be conducted manually every time it is needed, regardless of whether such a linkage has been completed in the past by others. Collectively, these insights are informing the conceptualization of a public health and criminal justice data sharing system.

The issue of data privacy is commonly cited as a rationale for not sharing data. Individuals and organizations are quick to envision scenarios where personal data are inadvertently released with devastating consequences. While data privacy must be taken seriously (see Privacy below for more information on this issue), at the same time, data have tremendous potential for social good. It is not unreasonable to question whether our current rules and regulations related to data privacy have negatively impacted the narrative around the benefits of data sharing for public benefit, and to consider how we might balance these competing concerns. There may also be an opportunity to be clearer in our narrative about how sharing data for public good can also have a direct benefit for individuals. In recent years, for example, we have seen an increasing amount of media coverage

highlighting various ways that climate change (a similarly nebulous topic that will require collective action) impacts the health of individuals, and this may serve as a model for public health.^{66,67}

Public health as a field has been plagued by a general lack of understanding about what it is, what it does, and what it includes, and this awareness links to interest in public health data. Most Americans are not familiar with public health and do not readily share that they see public health action in their daily lives. In contrast, individuals are able to comprehend the value of the health care system, transportation, or education system, where the perceived benefits to residents and the community are more obvious, and individuals can visibly see those systems in action. This suggests a need to not only develop messaging around what public health is, its value to society, and the critical importance of data in achieving its mission, but also to consider how to make the work that public health does, as well as the public health infrastructure in place to support health and well-being, more visible to residents and policymakers. This is particularly important in light of the need to combat misinformation and to elevate credible health communicators in our health narrative.^{68,69} Increasing the visibility of public health could help other sectors see more easily how their data could serve as important inputs to population health and health-related decision-making.

Most of the public health data collected could, when properly handled, be used for the common good and shared for public benefit. The amount of health and other data that are currently being collected and shared via our online activities (that we knowingly or unknowingly consent to) is staggering and is currently being used largely for commercial gain. That more of these data are not being used also for public good is a missed opportunity, particularly during a crisis that has spurred many of these companies to examine the public health potential of their assets and resources. As referenced in the What? paper, more work is needed to engage private industry and other sectors that own data about the value of sharing that data consistently for the purposes of strengthening community health and well-being. Messaging to advance the sharing of data across sectors, however, will likely require more than an altruistic argument.

Although some businesses may be willing to share data as a social good (assuming the costs and risks of doing so are low), others may require a more clear and compelling narrative that ties to other outcomes they may be more interested in (e.g., public relations, healthier workforce, improved bottom line), helping them to see “what’s in it for them.” Another possibility is seeking out companies whose missions are naturally aligned with public health goals. For example, Nike or other athletic brands could benefit from an increase in physical activity, as would public health. At the same time, businesses that transparently share data may catalyze a recognition among their customer base and competitors around the value of sharing data for public benefit, which may translate to increased buy-in around data sharing over time.

Changing the national narrative for both individuals and data owners, however, will require a multi-pronged effort, starting with national leadership. There may be an opportunity to elevate the importance of the public’s health through regular reporting on a sentinel indicator or set of indicators by top executive leadership and state and local leadership, which could signal political support for data to improve the health of the nation. This returns to the principle of parsimony, noted in the preamble and the What? paper, calling for a more coherent set of indicators that could be easily used by business leaders, amplifying the value of public health and public health data. There may also be opportunities to partner with well-established business or sector leadership, which may serve

as catalytic change agents. By being one of the first to share data, they can serve as a “proof of concept” for other businesses or industries. Engaging a diverse set of journalists, social media outlets, and news organizations across the political spectrum, from different geographies, and of different lived experiences to develop communication strategies and tailored messaging may also help change the narrative around the value of sharing data for the public good.⁷⁰

3. Standards of access and interoperability

Section 2 above focused on core considerations around increasing participation in data sharing efforts and ensuring that such data sharing practices are equitable. Those considerations can be viewed as early steps in a theory of change that utilizes incentives (financial or otherwise) to encourage data sharing and address current system deficits, and changes narratives and mindsets around how people understand health, public health, and the public health data system (section 2.2). Section 3 of this paper builds on this theory of change to elevate considerations around data and system standards that could promote interoperability and strengthen data sharing.

3.1 Key Action: Identify appropriate models for information exchange.

A wide range of models are currently in place to facilitate data sharing, and while the details of each has been uniquely negotiated by the participating organizations, they fall within several basic models or designs for information exchange, described below. As noted in the preamble, this translation of data into information in ways that are useable is a key part of useable knowledge development. Given that each has its own strengths and weaknesses, it will be important to prioritize the use cases (e.g., tracking chronic disease burden, real-time outbreak surveillance, research) for public health data to help determine which models or mix of models will optimally inform the advancement of the public’s health. Related to this is a consideration of the data user, and whether the needs of state and local health departments could or should be prioritized over the needs of federal agencies or researchers to facilitate a more timely, localized response to emerging health challenges. Such a prioritization, however, could perpetuate inequities in data access. Another important consideration is assessing what it will take to migrate legacy data systems to a point where they can actively and efficiently participate in any model of data sharing with minimal human effort and securing the funding and other supports needed to sustain and maintain these new systems, including ongoing testing and refinement. It will be important, however, to ensure that we are not further burdening organizations with increased complexity and cost in the pursuit of efficiency regardless of the model(s) selected.⁷¹ **Figure 4** shows the continuum of information exchange models:

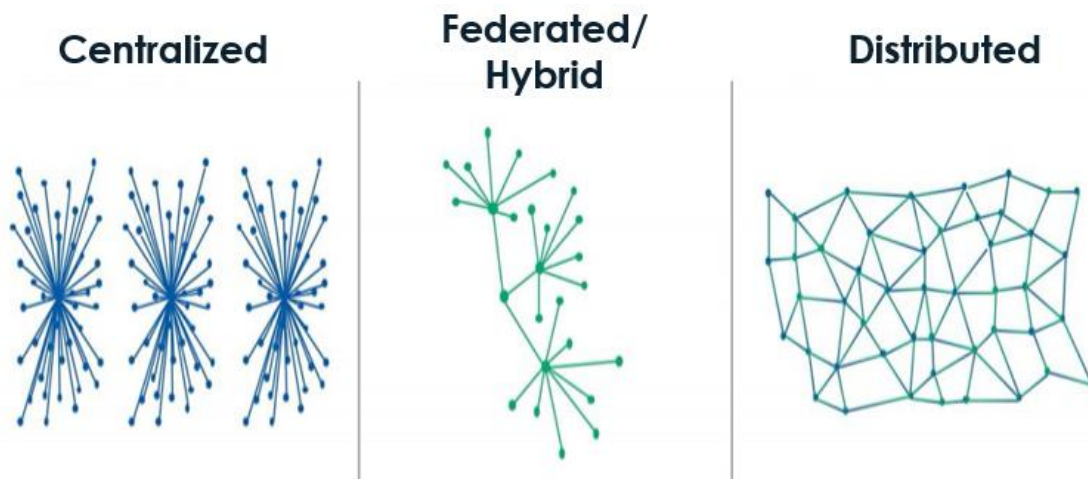
- **Centralized:** Data are sent to a centralized database or repository by a network of users. The data can then be accessed by participants in accordance with the policies and procedures governing that repository.
- **Federated:** Data are stored in separate data repositories or nodes, which are interconnected in a federation.⁷¹ The nodes have a central management framework and set of agreed upon

tenets and standards that creates a uniform environment so that the member networks can share data.^{71,72}

- **Hybrid:** Data are stored in separate data repositories or nodes, but a central repository contains a limited set of standardized data on every individual in the network along with a record locator service, which enables tracking of data across nodes.^{73,74}
- **Distributed:** Data are stored in a distributed network, where users maintain control of their data but provide access to other users by enabling user privileges.

In the following sections, we describe key considerations about each approach.

Figure 4. Models of information exchange



3.1a. Consideration: *Centralized models are relatively simple to deploy, but data access can be more burdensome, and data are at higher risk from a privacy and security standpoint.*

In a centralized system, data are sent to a centralized database or repository by a network of users. The data can then be accessed by participants in accordance with the policies and procedures governing that repository.^{74,75} One example of a centralized model is the North American Association of Central Cancer Registries (NAACCR).⁷⁶ In addition to establishing a set of data standards and facilitating the tracking of cancer patients across participating states, NAACCR is developing the Virtual Pooled Registry Cancer Linkage System (VPR-CLS) to efficiently connect researchers performing minimal risk linkage studies with population-based cancer registries.⁷⁶ Although centralized models of data sharing have advantages in that they are relatively simple to deploy, affordable to maintain, and provide central control of the data, this centralized control can also be a limitation if the party responsible for overseeing the database or repository is slow to adapt to emerging technologies or address issues that could improve the quality, completeness, or value of the data. Another potential challenge is finding an entity that is sufficiently trusted or well-regarded to serve in this role. Other challenges may involve the potential for duplicate records and burdensome processes for data access that may affect productivity or a timely response.^{73,77}

Centralized data models are also at higher risk from a privacy and security standpoint, as a breach in data protections places the entire database at risk.

3.1b. Consideration: *Federated data systems reduce privacy and technical risk inherent with data transfer, but it can be challenging to reach agreement from federation members and beneficiaries on important aspects of data sharing.*

Unlike centralized systems, federated data systems do not have a single owner. Rather, data live in separate data repositories or nodes (e.g., health provider, public health department), which are interconnected in a federation.⁷¹ The nodes have a central management framework and set of shared tenets and standards around security, auditing, authentication, and access rights that creates a uniform environment so that the member networks can share data.^{71,72} These shared tenets and standards are decided on by members of the federation and therefore vary across federated systems. Application programming interfaces (APIs) provide a secure means of accessing local data of the federated system.⁷³ The federated data model reduces privacy and technical risk inherent with data transfer, as the data are not shared. Rather, the query is sent to each data node, processed locally, and then aggregated and returned to the individual making the request.⁷¹ Despite its benefits, this model of data sharing adds complexity to the process, as one of the biggest challenges can be reaching agreement from federation members and their beneficiaries on shared goals and tenets, types of data included, metrics of success, and approved data use.⁷¹

3.1c. Consideration: *Hybrid models combine the functionality of both centralized and federated models but may be better suited for clinical and research applications.*

A common model of health information exchanges is a hybrid model, which combines the functionality of both the centralized and federated infrastructure models. In this model, a central repository contains a limited set of standardized data on every individual in the network along with a record locator service, which enables tracking of individual movement between providers or services.^{73,74} In addition to obvious clinical applications to track patient care utilization, it can also be useful in research for developing patient cohorts of populations of interest. Once the cohort has been identified, the central repository provides information on where the source data can be requested for each potential participant within the federation.^{73,74} This model provides more flexibility for clinical analytics and medical research than either the centralized or federated models, but it is not clear whether it is as useful for informing local decision-making around issues of population health.

3.1d. Consideration: *Distributed network models are similar to federated models, but they do not require a priori agreements on all aspects of data use and can be more costly and require trust among members of the network.*

Similar to a federated model, a distributed model does not have a centralized location where data are stored. Rather, data are stored in a distributed network, where users maintain control of their data but provide access to other users by enabling user privileges. This enables data to be integrated from disparate sources.^{27,75,78} A key difference between a distributed model and a federated model

is that the distributed model does not require an agreed-upon set of shared standards as with the federated model.⁷⁵

When used in conjunction with software platforms, such as PopMedNet, that facilitate distributed querying and analytic tools that protect patient privacy, health departments and researchers can query multiple data sources simultaneously.⁸⁰ Examples of this model include the Patient-Centered Outcomes Research Network (PCORnet), used commonly for research, and MDPHnet, a distributed health data network in Massachusetts used often for public health surveillance.²⁷ In the latter example, public health officials use PopMedNet to query the network, which returns de-identified aggregate counts, while network software maintains a data repository behind the firewall of each practice to ensure the privacy of individuals.¹³ Distributed network models may be particularly well suited for models of information exchange that give individuals more control over their own data. OAuth, for example, is an open authentication protocol commonly used by technology companies, like Amazon and Google, that allows a data owner to grant permission for a third party to access their information on other websites or in other applications, without sharing passwords.^{81,82} CommonHealth is another example.⁸³ Although the distributed model helps to overcome some of the privacy concerns of other models and is not vulnerable if a single data source goes off-line, it can be costly to develop and maintain and requires significant effort to build trust among members of the network.

Exhibit 9: Putting individuals in charge of their own data

CommonHealth, developed by the nonprofit The Commons Project in collaboration with a range of private and public partners, is an app that helps put power over health data in the hands of individuals, who then share with entities they trust.⁷⁹ CommonHealth interposes a trust framework (intended to be similar to Consumer Reports) to help individuals decide which apps are trustworthy. The app works with a third-party ecosystem to ensure those entities to be trustworthy. CommonHealth aims to serve everyone on an equal playing field so that for-profit stakeholders can compete within it.

3.1e. Consideration: *Blockchain is an emerging technology that holds promise for overcoming challenges of privacy and trust in data, but it can be expensive to maintain and has not been widely tested in the health sector.*

Blockchain is not noted in Figure 4 but may be useful to consider given the pace of innovation. Blockchain is a peer-to-peer distributed ledger technology that is gaining interest as an effective way to improve the transparency of data exchange and trust in data-sharing systems.^{84,85} It uses a distributed network, where each member stores an identical copy of the blockchain that is updated in real time. As members in the network add transactions, entries are validated, encrypted, and digitally signed, after which the data entries become immutable and cannot be changed or removed. These processes help to increase confidence among users that the data are secure, accurate, and have not been tampered with, which helps to build trust.^{77,84,86,87} Blockchain technology also allows

the user to share data in compliance with privacy laws like the Health Insurance Portability and Accountability Act (HIPAA), while retaining full control of that information.^{77,85}

While blockchain technology has been in use in other sectors like banking, its application for sharing public health data is relatively new and deserves further study. The public health field may benefit from the implementation of pilot programs to test potential use cases, like the pilot program for an EHR blockchain conducted by the CDC's Office of Technology, National Center for Health Statistics, and IBM.⁸⁵ Potential limitations of this model include a concern around bandwidth, storage capabilities, and maintenance costs for a growing blockchain that may drain resources, particularly for less-resourced organizations.⁸⁷ Issues of data ownership and regulatory issues such as who has the right to grant data access and for what purpose are also unclear.⁸⁷ Data standardization within and across blockchains is also an area in need of attention, if the goal is to use such data for surveillance or research activities.⁸⁷

3.2 Key Action: Strengthen interoperability of data systems.

The lack of data standards and the sheer number of unique information systems, most of which are not interoperable, pose major technical challenges for efficient and effective use of public health data, particularly given the data volume and variety issues explored in the What? paper. In addition to the technical challenges of interoperability, standardization around documentation and exchange of data specific to social risk factors, including social determinants of health, will be an important consideration in the development of an equity-focused health data system. This section highlights four considerations for strengthening the interoperability of data systems: 1) the need to leverage data integration techniques and establish a minimum viable product; 2) the potential role of modular approaches; 3) how to standardize measures, particularly of social determinants of health and other factors like racism, that impact health but are not uniformly defined; and 4) the role of policies, incentives, and collaborations to catalyze interoperability between public health, health care, and other sectors.

Exhibit 10: The promise and potential of integrated data systems for improving health equity

Project Evident of the Tides Center is conducting an assessment of data transformation to improve health equity, including how to do it and who needs to be involved.⁸⁸ They describe integrated data systems as having the following characteristics: data that are linked at the individual level, data that are systematic and consistent, and data that are focused at the intersection of data from different partners or domains. They note three key factors that have combined to accelerate the use of integrated data, increase analytical capacity, and create a culture of evidence-building and use in the social sector. These include policies that reinforce the critical role of data; increased funding from government and philanthropy; and advances in technology, data standards, and statistical methods. At the same time, the team identified several barriers to generating evidence. First, some data points are not collected because vendors and government actors do not adequately incorporate community and practitioner input when deciding “what matters.” Second, integrating data across systems from different sectors is challenging due to outdated technology or system designs and lack of data standardization, making it difficult to generate evidence based on integrated data. Third, data holders from different sectors have few incentives to share their own data with others.

3.2a. Consideration: *Methods for data integration continue to advance but are not being leveraged enough across sectors to support sense-making and decision-making.*

Data integration involves combining data from multiple sources (including, but not restricted to, surveillance systems), and data integration techniques can be useful whether they are complex or simple. Dashboards are one simple approach public health practitioners have used to integrate and distill data into more consumable formats, which can be used to promote data-driven decision-making or to support education and learning through data visualization.⁸⁹ Public health dashboards do not follow any standardized guidance, however, and thus while some elements may appear frequently across dashboards, many are only somewhat similar or not similar enough themselves to be integrated in a meaningful way.

Data integration methods in **Exhibit 11** provide a structure for producing ‘data quilts’ that leverage overlap in geographic coverage timing, and/or substantive outcomes of existing data sources, to produce improved surveillance estimates (e.g., more precise, greater periodicity, and/or better ability to measure small subpopulations like vulnerable populations or small geographic areas). **Table 2** shows a variety of methods for combining secondary data that range across the use of multiple secondary data sources (e.g., spatial interpolation of county health rates using census tract population counts); secondary and primary data (e.g., constrained regression estimation of census tract health rates using geocoded survey microdata and county health rates); to multiple primary data sources (e.g., combined-survey regression estimation of tract health rates using multiple

surveys, or survey and census microdata). Successful data integration requires stakeholder input to identify the priorities and preferences for surveillance.

Exhibit 11: Data integration methods

- Methods using secondary data (Macro): meta-analysis; data smoothing; spatial interpolation
- Methods using primary & secondary data (Macro+Micro): small-area estimation; constrained regression
- Methods using primary data (Micro): combined-survey regression with multiple imputation

To be able to access and integrate data for public health purposes, metadata on available data sets are needed. For example, the CDC has catalogued their data sets and the meta data and searchable system are available to CDC researchers on the CDC intranet.⁹⁰ Public health more broadly needs a similar set of metadata to be able to efficiently and effectively use the wide array of other data sets in key sectors, such as education, transportation, housing, and economics, which influence health, well-being, and equity outcomes. Unique approaches are being utilized to develop metadata across sectors for more holistic research. For example, *Data Lumos* was developed by the Inter-university Consortium for Political and Social Research (ICPSR) and is a "crowd-sourced repository for valuable government data" on the subjects of employment, labor, citizenship, and housing.⁹¹ *Data Lumos* contains over 5 million variables from over 15,000 studies and almost 100,000 publications.

Public health does not necessarily need to gather and store these data, but instead should focus on making these data more accessible so data can be integrated with other sources. However, given the complexity and difficulty in using the current volume of data, there may be lessons public health can learn from the technology industry's use of minimum viable product (MVP). The purpose of a MVP is to allow for an agile product development that allows for users to augment an initial simple set of features, often opting for the least complicated yet useful set of features to allow for users to achieve the vision of the product.⁹² Should public health identify a minimum data set or MVP equivalent? Adopting a 'less is more' approach for public health data would require some consensus about what the bare minimum or core basic set of public health data is that should be available to all public health departments to ensure they can work towards identified public health priorities. As previously mentioned, these data could reflect health's intersection with multiple sectors—education, housing, economics, etc.—and could be enhanced with local data. Clear guidance for how to construct the minimum data set and standardized definitions of measures would need to be developed to support implementation of such an approach in the future.

3.2b. Consideration: *Modular approaches and identification of common elements that facilitate interoperability have been leveraged in other industries, but not as widely in health data systems.*

One critical consideration for interoperability is how to structure a public health data system that allows for local flexibility, while also ensuring that data collected at a local level can be easily aggregated with data collected elsewhere. Modularity is a “general systems concept that describes the degree to which a system’s components can be separated and recombined and refers to...the degree to which the rules of the system architecture enable or prohibit the mixing or matching of components.”⁹³ Modularity can be contrasted with the consolidation or integration of systems, which is another approach to ensuring interoperability, but one that provides less flexibility to meet emerging or unique needs. While the health care system (and, by default, health care data), in general, is moving towards consolidation and integration, it is reasonable to ask whether modularity, which has been leveraged successfully in other industries, may be a helpful construct when thinking about the public health data system.⁹³

Using LEGO as an analogy, there are hundreds, if not thousands, of different types of LEGO bricks, but each type of brick is uniform in its size, shape, and color. The number of configurations achievable from even a fraction of available LEGO bricks is limitless. Extending this analogy to the public health data system, the standardization of measures would not only facilitate interoperability, but would also allow stakeholders to select and use measures most relevant for their work, greatly increasing the flexibility of the public health data system. This notion of “recombinability” in a modular system, or all the ways in which component parts (measures) can be put together, becomes more valuable as the demands of the (public health data) system become more heterogeneous, increasing system fitness.⁹³ Modular approaches have also been used in many other industries, including the shipping industry, where standardization of shipping containers and free use of patents to build machines that lift containers off of ships and place them on trucks not only removed bottlenecks, but also lowered costs significantly and allowed the industry to grow quickly.⁹⁴

3.2c. Consideration: *Important determinants of health remain poorly conceptualized, resulting in a lack of standardized measures that hinders interoperability, though ongoing efforts to address this challenge hold promise.*

Data on social determinants of health (SDOH) such as food security, housing stability and quality, transportation, and racism can help to shape the development of programs and policies that address upstream drivers of health and health inequity. Yet, despite being important concepts, as noted in the What? paper, they remain poorly conceptualized and often miss key aspects of inequity processes and outcomes. This results in a lack of standardized definitions and measures across data collection efforts. Several efforts are underway to standardize the collection of SDOH data that could provide important advancements towards addressing health equity (**Exhibit 12**). A key consideration, however, is ensuring that these and related efforts are integrated and align, to the extent possible. If multiple guidelines, definitions, and coding practices stem from these related efforts, and particularly if sectors elect to follow different recommendations, the challenge of interoperability and standardization for data sharing will remain. It may be worth considering whether a parallel to the American National Standards Institute (ANSI) or the National Institute of Standards

and Technology (NIST) could be developed for the public health data system, building on existing collaborations and efforts currently underway.^{95,96} These institutes promote and facilitate the establishment of consensus standards, measures, and conformity assessment systems and safeguard their integrity to build trust. However, such standardization may still be plagued by an incomplete set of domains and may miss deeper historical and systemic inequity issues, called out in more detail in the What? paper.

Exhibit 12: Efforts to standardize the collection of SDOH data

Objective 1c of the **ONC Federal Health IT Strategic Plan** is to integrate health and human services information.²³ One strategy offered as part of Objective 1c is to “advance standardization and interoperability of SDOH and social services data across federal programs through the use of standard health IT terminologies, definitions, and methods of data collection and exchange.” Another is to “capture and integrate SDOH data into EHRs and clinical decision support to inform care delivery (including referrals and integration of medical and social care) and to address health disparities in a manner that is ethical and consistent with routine patient care.” In February 2021, ONC released a special emphasis notice of interest under the Leading Edge Acceleration Projects (LEAP) in applications to advance health IT standards and tools to improve SDOH data exchange.

The **Gravity Project** is a national collaborative to advance interoperable SDOH data.⁹⁷ This is done through a collaborative process where stakeholders develop and test consensus-based standards to facilitate SDOH capture and data exchange across health and social services systems and settings. Through the development and testing of use cases, the Gravity Project “seeks to identify data elements and associated value sets to represent SDOH information documented in EHRs across four clinical activities: screening, diagnosis, goal setting, and intervention activities.” The Gravity Project is also collaborating with organizations to address coding gaps and developing and testing HL7 FHIR Implementation Guides.

The **3-D Commission** is a partnership between the Rockefeller Foundation and Boston University that is seeking to improve health-related decision-making by creating a demand for public and private investment in the social determinants of health and creating a common language between fields of data science, health determinants, and policy and practice decision-making.⁹⁸ The commission’s report is expected to be released in late spring 2021.

3.2d. Consideration: *Policies, incentives, and collaborations help catalyze interoperability, but they remain fragmented, which can create further inequities in the public health data system.*

EHRs are now used by more than 85 percent of ambulatory providers and by the vast majority of hospitals, making their data a valuable source of data for public health use.⁹⁹⁻¹⁰¹ In recent years, the federal government has leveraged incentive programs to promote interoperability and the collection of a standardized set of data through CMS' Meaningful Use program and, more recently, MIPS. These incentive programs have tied provider payments to standards of data capture and information exchange.²³ In 2020, CMS finalized a rule to advance interoperability by promoting the flow of electronic health information (EHI) and providing patients with access to their health information.^{23,102} The CURES Act is another statute designed to strengthen interoperability. In 2020, ONC released the CURES Act Final Rule “designed to drive interoperability of EHI by supporting the use of ... Fast Healthcare Interoperability Resources (FHIR) standards for application programming interfaces (APIs).”^{23,103} Use of FHIR has broad federal support and fosters data sharing between a wide range of potential users including patients, providers, and other health care entities.

Exhibit 13: *Creating a model for data-driven policymaking for Native Hawaiian and Pacific Islander populations in the time of a pandemic to help promote equity*

The **Native Hawaiian and Pacific Islander (NHPI) Data Policy Lab** conducted a series of key informant interviews and focus groups to elevate issues related to the availability of disaggregated NHPI COVID-19 data during a pandemic.¹⁰⁴ Several key insights emerged. First, apply community-based participatory principles to governance structure, and incorporate community stakeholders into organization formation and decision-making processes. Second, strong data privacy policies are necessary to prevent the release of personally identifiable information to law enforcement for immigration purposes and to prevent commercial exploitation. However, in health emergencies such as pandemics, health systems should have policies in place that can lower thresholds to a previously agreed upon level that is triggered by state or federal declarations of public health emergencies. Third, utilization of machine learning and artificial intelligence must be coupled with additional oversight to ensure appropriate data sets have been selected and algorithms are audited for bias. Fourth, incentivize data collection and any oversampling necessary to allow for reporting of population subgroups such as Chamorus, Fijians, Marshallese, Native Hawaiians, Samoans, and Tongans, among others. The suggested subgroups are those that make up the majority of NHPIs in the country. These subgroups also capture the different types of experiences and identities that impact access for NHPIs through one or more of these designations: Indigenous, Citizens, Birthright citizens of U.S. Territories, Nationals, Legal permanent residents, Undocumented, and Compact of Free Association Migrants. Fifth, enforce data collection and reporting of OMB 15 categories at a minimum. And sixth, implement data processing transparency policies that describe how steps are taken to process and impute data, imputation techniques, and the effect of imputation for small populations.

Although these efforts are important contributions towards interoperability, particularly for clinical data stored in EHRs, several considerations remain. First, while EHRs are widespread, there remain pockets of providers who do not utilize them or who do not participate in CMS incentive programs. This leaves open the question of how to incentivize interoperability among non-participating health systems, particularly if the patient population of that health system is underrepresented in current data systems. Second, although policies and strong incentive programs are promising change agents, their reach is limited to the health care system and the providers who participate, in part because EHR data systems were not designed to be shared outside specific entities. As a result, more work is needed to ensure interoperability of EHR data with relevant data from other services and sectors, including private entities, and to overcome the historic divide between the public health and health care sectors. Third, improving interoperability could benefit from ongoing technological and methodologic advancement in how EHR systems interact and data are shared, with particular attention to interoperability across sectors and with public health specifically. Revisiting earlier considerations around equity and resource allocations for public health infrastructure, such innovations could focus on novel approaches to efficiently “wrapping” legacy systems with newer technologies and capabilities (without creating additional complexity), as opposed to replacing existing systems outright, which may be cost prohibitive for many public health departments.¹⁰⁵ At the same time, federal and state governments should also adopt technology emerging from interoperability standards.

A number of forums and collaboratives have coalesced around issues of interoperability and data sharing across diverse sectors (**Exhibit 14**). Collectively, these forums offer readily available expertise, experience, and lessons learned that could be leveraged to further tackle existing and emerging challenges related to interoperability and standardization. Fostering stronger connection, collaboration, and alignment between these groups has the potential to expand impact and result in an even stronger set of recommendations, guidelines, and best practices for interoperability across a diversity of settings and sectors. As noted in the What? paper, technology companies (e.g., digital health companies) could also be an important partner in thinking through interoperability considerations because many have been expressly digital from the beginning. Companies are not usually stuck in digital capability traps (e.g., challenges with migration towards digital systems) and have interoperability baked into their business model.

Exhibit 14: Learning forums and collaboratives focused on interoperability and data sharing

All In: Data for Community Health is a learning network of communities, representing over 150 collaborations across the country, testing new ways to improve community health through multi-sector partnerships working to share data. National and regional programs coordinate technical assistance and facilitate peer-to-peer learning activities across partner networks to build the evidence base to advance practice, identify gaps, highlight investment needs, and inform policy.¹⁰⁶

Data Across Sectors for Health (DASH) is a RWJF-funded effort that supports community collaborations in their efforts to address locally determined problems or goals; enhance communities' ability to plan; make decisions; implement health improvement activities through sharing of data and information; and identify methods, models, and lessons that can be applied locally and shared with other communities who wish to improve their ability to share data and information across sectors. DASH has launched a funding opportunity called Community Impact Contracts-Strategic, Timely, Actionable, Replicable, Targeted (CIC-START), which aims to help local collaborations catalyze their efforts to share and use multi-sector data to improve community health. Awardees also receive support to participate in All In: Data for Community Health.^{107,108}

Digital Bridge is a forum comprised of individuals from health care, public health, and information technology. Goals of Digital Bridge include easing the burden and costs for all stakeholder groups through a unified approach to information exchange; advancing greater standards-based information exchange across public health and health care; and laying the foundation for greater bidirectional exchange of data so that clinicians can be more informed about population health, environmental risks, and outbreaks. Past projects included the design of a nationally scalable, multi-jurisdictional approach to electronic case reporting (eCR) and has recently developed the ExeCC (Expanding eCR's Capacity and Capability) use case workgroup, focused on generic enhancements to the existing eCR infrastructure that would support additional reporting beyond nationally notifiable conditions.¹⁰⁹⁻¹¹¹

GO FAIR implementation networks are also creating specific materials and tools as elements of the Internet of FAIR Data and Services (IFDS). The FAIR guiding principles for scientific data management and stewardship were published in 2016 and include guidelines to improve the findability (e.g., through accurate, standardized metadata), accessibility (e.g., balanced against privacy, resources), interoperability (e.g., architectures, formats), and reuse of digital assets (e.g., generation of new knowledge).¹¹²

Stewards of Change Institute recently launched the National Interoperability Collaborative, a community of networks designed to promote greater information sharing, interoperability, and collaboration in human services, public health, health IT, education, public safety, emergency response, and other relevant domains related to and including SDOH and well-being. Their goal is to enable more and better partnerships, collaborations, relationships, and linkages at the local, state, and national levels.¹¹³

4. Privacy and security

With respect to data sharing for reuse, privacy is often cited as a barrier. Organizations not interested in sharing data point to privacy laws as justifications for not sharing data with public health entities. However, others have speculated that some organizations use privacy laws as an excuse to not share data, when their real rationale for not sharing data is more connected with notions of competitive advantage or self-preservation.¹¹⁴ This section focuses on considerations to advance aspects of the legal and regulatory environment that continue to protect privacy and promote data security while reflecting the changing needs of an evolving public health data system. These include ensuring laws strike a balance between privacy and benefits of using data for public good, modifications to informed consent, and revisions to language in HIPAA's Privacy Rule.

4.1 Key Action: Ensure laws strike a balance between the protection of privacy and benefits of using data for public good.

Privacy protections, particularly of personally identifiable information, are important for ensuring that individuals are not harmed by the inadvertent release of information. At the same time, overly restrictive privacy protections can be problematic for public health or other beneficial reuse cases, because data are either not released or not released in a way that would allow for maximum public benefit. How data are collected, stored, accessed, and analyzed have also evolved in recent years, and privacy laws have not kept pace with rapid advancements in health IT and data management, and vary considerably by state.

There is precedent for revisiting privacy laws to achieve a better balance between the protection of privacy and the benefits of data sharing. In 2020, legislation was passed related to 42 CFR part 2, which covers substance use treatment information.¹¹⁵ Substance use is the only condition that has its own privacy law, given potential ties to illegal behavior, but the lack of available data for reuse has made it difficult to address emerging public health concerns, such as the opioid crisis.¹¹⁵ Furthermore, because 42 CFR part 2 inhibited the sharing of information between providers, it was more difficult for individuals to receive comprehensive and wrap around services. Recent legislation was passed to bring this law closer in alignment to HIPAA, allowing the individual to give a one-time consent to share information about their receipt of treatment for substance use with covered entities and providers. The legislation also allows the data to be used with prescription drug monitoring programs, opening the door to state-level data systems.

Two potential challenges will need to be addressed in current privacy laws to achieve a balance between privacy and benefits of data reuse. The first is that state privacy laws, which supersede federal regulations, vary considerably with respect to privacy and data reuse. The second relates to emerging laws governing what Nicholson and colleagues have termed “shadow health records,” which are “collections of health data outside the health systems that provide detailed pictures of individual health.”¹¹⁶

4.1a. Consideration: *State laws vary considerably with respect to privacy and reuse, making coordination with public health difficult.*

Federal regulations may be superseded by more stringent state and local privacy laws.¹¹⁷ Knowing whether and how data can be shared with public health or reused for public benefit can be difficult given the stacked nature of federal, state, and local privacy laws. Particularly challenging is that state privacy laws vary considerably in the degree of privacy protections afforded. For example, some states' laws lack specificity about who can access data, do not address secondary disclosures of information, and do not address or take into consideration the use case of the data being obtained.^{118,119} On the other end of the spectrum, some states' privacy laws are written in a way that limits the ability of public health to exchange identifiable information across jurisdictions.^{118,119} While the lack of specificity around data access and reuse poses a potential privacy risk in some states, other states have developed privacy protections so stringent that the law is obstructing public health activities.

The Model State Public Health Privacy Act of 1999 was developed to help states provide “strong privacy safeguards for public health data while preserving the ability of state and local public health departments to act for the common good.”¹¹⁸ The extent to which states leveraged this model act in the development of their own privacy protections is not clear, and the model act is specific to the protection and use of a more narrowly conceptualized set of data than are available today. Regardless, it establishes a framework and language for achieving this critical balance between privacy and reuse that may be helpful in the current day as state and local jurisdictions revisit statutes governing the protection of health data.

4.1b. Consideration: *While emerging laws governing “shadow health records” provide privacy protections, they could negatively impact the reuse of such data for public health.*

As noted above, “shadow health records” describe collections of health data that have been de-identified and are no longer protected by HIPAA but provide enough information that through ‘big data’ analytics and methodologic advances can be combined with other unprotected data to provide fairly comprehensive pictures of individual health.¹¹⁶ This practice essentially legally skirts existing privacy laws, in that HIPAA and other privacy regulations do not account for this type of technological advancement.¹¹⁶ The challenge is that these shadow health records can be used for commercial gain, or worse—malintent—but at the same time, could provide a promising opportunity for research, innovation, and public health practice that is not possible under current privacy laws.¹¹⁶

In response, Europe has implemented the General Data Protection Regulation (GDPR) that went into effect in 2018, and California recently passed the Consumer Privacy Act (CCPA) that went into effect in 2020.^{120,121} Both laws attempt to extend privacy protections to shadow health records held by companies. GDPR, for example, requires companies to obtain consent, to collect and use only as much data as necessary, and prohibits processing of health data. GDPR does include an exception for “public interest in the area of public health” although the scope of this exception remains unclear.¹¹⁶ CCPA created notice and access requirements for businesses, and consumers can opt out of the sale of their information or request that certain information be deleted. CCPA has

exceptions for research in the public interest, but it is not clear whether this applies to organizations who would commercially benefit from the research.¹¹⁶ These laws have the potential to give individuals more control of the use of their data, but requirements like informed consent could have unintended consequences for equity. Given the resources required to collect consent from large numbers of individuals, industry, and in particular larger industry, may be able to obtain documentation more easily than academic, non-profit, or governmental sectors, resulting in unequitable access to health data.¹¹⁶ Given the potential value of such records for public health, emerging privacy laws at the federal, state, and local levels should aim to balance these concerns, clarify exceptions, and ensure that such data can be equitably accessed.

4.1c. Consideration: *Consent processes specify what may be shared, with whom, and under what circumstances, but this level of specificity hinders data reuse.*

Consent to treat (for the provision of medical care) and informed consent (for participation in research studies) provide information about benefits and risks of receiving care or participating in research, describe potential costs to the individual, and outline whether, when, and with whom information collected through the encounter will be shared. Because consent forms are often written for an explicit purpose, this level of specificity can hamper the reuse of health data by public health, even with existing exceptions to existing privacy rules.

It has been suggested that current consent processes could be modified to allow patients to “opt-in” to broader use of their data. If done well, this could promote transparency, give consumers greater control over their data, and alleviate some of the reluctance that providers and other organizations have around sharing data with public health and others for the purpose of public benefit. One challenge is that current consent processes do not discuss the potential value of reusing the data for public benefit. Beyond the legal and privacy considerations, successful implementation and adoption of a more generalized consent will also require narrative shifts around the value of such data for public benefit.

Use of plain language in the development of the consent form will be an important consideration to ensure that individuals can make an informed decision. Critiques of current consent procedures note that the process is more focused on legal protection of the organization seeking consent than ensuring the full understanding of the person from whom consent is being sought. Legal jargon, limited opportunity to meaningfully review and seek clarification before signing, and availability of consent forms in a limited number of languages contribute to this challenge.

One example of this type of consent comes from the technology field, where companies ask individuals to directly authorize providing technology-related data for public benefit through existing protocols like OAuth. OAuth, as noted earlier, is a protocol that allows users to authorize applications to interact with one another for purposes such as logging in to one application with credentials from another application.¹²² These technologies could provide a potential model from which to build a more general consent for data reuse process.

4.2 Key Action: Revise language in the Privacy Rule and other statutes to encourage sharing of identifiable data with public health.

Although many privacy laws allow the sharing of data with public health entities through public health exceptions, these exceptions are often ill-defined and differentially interpreted, resulting in decisions to err on the side of not sharing data for fear of violating those laws. Clarification around these exceptions, and consideration of whether public health agencies should become a covered entity to facilitate the flow of health-related data for public health practice, may be valuable. And while privacy rules differentiate between activities such as public health and research, neither is well defined and there is little recognition of the potential for overlap. In January 2021, the Department of Health and Human Services (HHS) released proposed modifications to the HIPAA Privacy Rule to support, and remove barriers to, coordinated care and individual engagement.¹²³ The Office of Civil Rights developed many of the proposals related to privacy standards and considered the extent to which each proposed modification would impact privacy protections relative to the benefits of making PHI more available for coordination of care or case management. The modifications did not propose changes to the disclosure of such data to public health authorities.

4.2a. Consideration: *The Privacy Rule of HIPAA includes an exception that allows disclosure to public health authorities authorized by law, but “authorized by law” is not defined in the Privacy Rule and is differentially interpreted.*

The Privacy Rule of HIPAA specifies that a covered entity may not disclose personal health information without individual written authorization, or consent (**Table 2**). However, the Privacy Rule also allows for a series of exceptions, one of which is sharing data with “public health authorities authorized by law to collect or receive such information for public health purposes.”¹²⁴ Because the Privacy Rule does not require that data holders share data with public health, but merely allows the sharing of data, organizations have been forced to grapple with what “authorized by law” means.¹¹⁹ The net result is that data holders have resisted sharing data and have placed the burden on public health authorities to demonstrate their legal right to access the data. Some federal agencies require statutory authority to be explicitly specified before they will release restricted or nonpublic data.¹²⁵ Demonstrating this legal need requires resources and legal expertise that many local public health departments do not have.

It could be beneficial to revise the types of activities that can be classified as being done for “public health purposes.” Current statutory language classifies this as data for the purpose of preventing or controlling disease, injury, or disability. As the scope of public health evolves to proactively improve health and well-being more broadly and address inequity, an important consideration will be how to align the public health exception with the future needs of public health practitioners that extends beyond disease surveillance. It is also worth considering whether public health entities could become covered entities to facilitate the sharing of health data. Although public health departments that fall under the “health care clearinghouse” designation can be considered covered entities, this designation does not apply to all public health departments.³³ Because larger, better resourced public health departments may meet this designation, this has the potential to widen gaps and inequities in data access between those communities that are more and less resourced.

Table 2. HIPAA law and public health data disclosure

Covered entities	What is covered	When can public health data be disclosed
<ul style="list-style-type: none"> • Health care providers • Health plans • Health care clearinghouses • Business associates of covered entities 	Protected Health Information (PHI): Contains individual identifiers including: <ul style="list-style-type: none"> • Names • Social security numbers • Photos • Addresses • Any other reasonable basis to identify an individual 	Covered entities are permitted but not required to disclose PHI without the individual's authorization to public health authorities for public health needs including: <ul style="list-style-type: none"> • Disease reporting • Vital events reporting • Public health surveillance HIPAA exemptions may be superseded by more stringent state and local laws
	Limited data sets: Can retain identifiers only in the form of <ul style="list-style-type: none"> • Location data • Dates • Ages 	After a data use agreement between the covered entity and the data recipient is established. Covered entities are not required to provide limited data sets.
	NOT de-identified data: aggregate data or data stripped of all identifiers	Not covered by HIPAA

4.2b. Consideration: *Privacy laws differentiate between activities such as research and public health practice, but neither is well defined and there is little acknowledgement of the potential for overlap.*

Another area of confusion within privacy statutes is that there are often separate exceptions for research and public health. However, regulations do not provide clear enough definitions or guidance on how to distinguish these activities. Particularly challenging is that both researchers and public health practitioners gain access and use data for health-related purposes, with the ultimate goal of improving health. This confusion was noted at the federal level in a recent HHS report that noted “Although ‘research’ and ‘public health’ are defined in statute... There is a disconnect with how ‘research’ and ‘public health’ are interpreted from one agency to the next [in the HIPAA Privacy Rule]. This causes frustration and dissimilar outcomes across the department” (p. 17).¹²⁵

In 2004, CSTE provided a clarification, which could be potentially helpful, where public health practice was defined as “the collection and analysis of identifiable health data by a public health authority for the purpose of protecting the health of a particular community” (p. 16).¹²⁶ Research was defined as “the collection and analysis of identifiable health data for the purpose of generating knowledge that will benefit those beyond the participating community who bear the risks of participation” (p.15).¹²⁶ This distinction could be potentially beneficial for communicating to individuals about the value of sharing health data. It is not known, for example, whether individuals

would be more willing to consent to the reuse of their health data if they knew the data were being used to promote the health of their community, where the benefits could be more directly felt. However, the word ‘community’ in this definition would need to be clarified to specify whether it refers to a geographically bound region, or groups of people bound together by shared characteristics, health factors, or lived experiences. As the definition of community expands to groups of individuals with shared characteristics, the lines between research and public health again begin to blur.

Conclusion

This white paper laid out practical challenges to achieving a modern, equity-oriented public health data system related to entrenched interests, legacy systems, antiquated data collection processes, cost, lack of standardization and interoperability, privacy and legal protections, lack of trust, and public health workforce capabilities.⁴¹ Such challenges have resulted in many shying away from true systems transformation, and instead opting to work around the edges. However, scrapping our current data system and starting over is both unreasonable and impractical. This paper offers a number of key actions and considerations that could serve as important levers in an evolution of the public health data system related to governance and data stewardship, data sharing, access and interoperability, and privacy and security that could inform meaningful transformation. Important work, both within and outside the public health system, is currently underway to tackle many of these challenges, but significant work remains, and difficult decisions and trade-offs lie ahead. It is also becoming increasingly clear that public health cannot do this alone—the success of our nation’s public health data system will rely on the active involvement of diverse sectors and systems to help figure out how to design a modern, equity-oriented public health system. In the final paper of this series, *Who and What Next?*, we describe potential roles for each sector in achieving this vision.

Acknowledgements

Each paper in this series of white papers has been reviewed by Mahshid Abir of the RAND Corporation and the University of Michigan and Karen Smith of Public Health Strategies. We appreciate these reviewers for their insights and contributions. We also thank Burness for their editorial and communication support for each paper.

References

1. Van Panhuis WG, P. P, Emerson C, et al. A systematic review of barriers to data sharing in public health. *BMC Public Health*. 2014;14:1144.
2. Marshall GR. Nesting, Subsidiarity, and Community-based environmental Governance beyond the Local Scale. *International Journal of the Commons*. 2007;2(1):75-97.
3. Oltrogge DL, Christensen IA. *Space Governance in the New Space Era*. NASA;2019.
4. Ostrom E. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press; 1990.
5. DeSalvo K, Hughes B, Bassett M, et al. *Public Health COVID-19 Impact Assessment: Lessons Learned and Compelling Needs*. Discussion Paper, Washington, DC: National Academy of Medicine; 2021.
6. Tragedy of the commons. Wikipedia. https://en.wikipedia.org/wiki/Tragedy_of_the_commons. Published 2021. Accessed January 4, 2021.
7. Haddow K, Carr D, Winig BD, Adler S. *Preemption, Public Health, and Equity in the Time of COVID-19*. COVID-19 Policy Playbook; 2021.
8. Aslin H, BROWN V. *Towards whole of community engagement: a practical toolkit*. 1 ed. Canberra: Murray-Darling Basin Commission; 2004.
9. Hatfield-Dodds S, Nelson R, Cook D. Adaptive governance: An introduction, and implications for public policy. Paper presented at: Annual conference of the Australian Agricultural and Resource Economics Society 2007; Queenstown, NZ.
10. Lebel L, Anderies JM, Campbell B, et al. Governance and the Capacity to Manage Resilience in Regional Social-Ecological Systems. *Ecology and Society*. 2006;11(1):19.
11. Denford J. *Idea to retire: Decentralized IT governance*. Online: Brookings; 2016.
12. Marshall GR. *Economics for Collaborative Environmental Management: Renegotiating the Commons*. London, U. K.: Earthscan Publications; 2005.
13. Young OR. Institutional interplay: The environmental consequences of cross-scale interactions. In: Ostrom E, Dietz T, Dolsak N, S. SP, Stonich S, Weber EU, eds. *The Drama of the Commons*. Washington, D. C.: National Academy Press; 2002:263-291.
14. Rechkemmer A. Climate Change, Development, and Health: The Role of Adaptive Governance. GRF One Health Summit; 2012; Davos, Switzerland.
15. Schultz L, Folke C, Österblom H, Olsson P. Adaptive governance, ecosystem management, and natural capital. *PNAS*. 2015;112(24):7369-7374.
16. Quality AfHRA. *National Health Data Stewardship*. HHS;2007.
17. Rosenbaum S. Data Governance and Stewardship: Designing Data Stewardship Entities and Advancing Data Access. *Health Services Research*. October 2010;45(5):1442-1455.
18. Pollard MS, Davis LM. *Decline in Trust in the Centers for Disease Control and Prevention During the COVID-19 Pandemic*. Santa Monica, CA: RAND Corporation;2021. RRA308-12.
19. CDC. Data Modernization Initiative: An Urgent Need to Modernize. <https://www.cdc.gov/surveillance/surveillance-data-strategies/dmi-investments.html>. Published 2020. Accessed 2020.
20. Epidemiologists CoSaPH. Driving Public Health in the Fastlane. https://cdn.ymaws.com/www.cste.org/resource/resmgr/pdfs/pdfs2/Driving_PH_Print.pdf. Published 2019. Accessed 2020.
21. Forum WE. *Unlocking the Value of Personal Data: From Collection to Usage*. Cologny, Geneva Canton, Switzerland; 2013.
22. Kanaan SB, Carr JM. *Health Data Stewardship: What, Why, Who, How: An NCVHS Primer*. National Committee on Vital and Health Statistics; 2009.
23. Technology TOotNCfHI. *2020-2025 Federal Health IT Strategic Plan*. Office of the Secretary, United States Department of Health and Human Services; 2020.
24. Law NfPH. *COVID Data Disaggregation: Legal Authorities and Barriers [White Paper]*. Princeton, N. J.: Robert Wood Johnson Foundation; 2021.
25. Suver C, Thorogood A, Doerr M, Wilbanks J, Knoppers B. Bringing Code to Data: Do Not Forget Governance. *Journal of Medical Internet Research*. 2020;22(7).
26. Dash S, Shakyawar SK, Sharma M, Kaushik S. Big data in healthcare: management, analysis and future prospects. *Journal of Big Data*. 2019;6:54.
27. Yoon P, Pollock D, Foldy S. National Public Health Informatics, United States. In: Magnuson J, Dixon B, eds. *Public Health Informatics and Information Systems*. Springer, Cham.: Health Informatics; 2020.

28. Hripcsak G, Bloomrosen M, Flatley Brennan P, et al. Health data use, stewardship, and governance: ongoing gaps and challenges" a report from AMIA's 2012 Health Policy Meeting. *Journal of American Medical Informatics Association*. 2014;21:204-211.
29. *Data Platform Viability Project Project Update*. Washington, D. C.: Health Care Cost Institute; 2021.
30. Kalil T, Zhao F. Unleashing the Power of Big Data. 2013. <https://obamawhitehouse.archives.gov/blog/2013/04/18/unleashing-power-big-data>.
31. Beninson L, Brown Q, Burrows E, et al. *The Federal Big Data Research and Development Strategic Plan*. National Coordination Office for Networking and Information Technology Research and Development; 2016.
32. 2020 Action Plan. Federal Data Strategy Web site. <https://strategy.data.gov/action-plan/>. Published 2020. Accessed January 2021.
33. CDC. Data Modernization Initiative: An Urgent Need to Modernize. Centers for Disease Control and Prevention. Public Health Surveillance and Data Web site. <https://www.cdc.gov/surveillance/surveillance-data-strategies/dmi-investments.html>. Published 2020. Accessed January, 2021.
34. Timely. Accurate. Accessible. The New World of Public Health Data [press release]. Online: Centers for Disease Control and Prevention undated.
35. SOCI, NIC. *Interoperability Insights: Demonstrating the Need and the Benefits of Connecting Health and Human Services*. Online: Stewards of Change Institute; 2020.
36. Krahn GL, Walker DK, Correa-De-Araujo R. Persons With Disabilities as an Unrecognized Health Disparity Population. *American Journal of Public Health*. 2015;105:S198-S206.
37. Trevathan E. Foreword. In: Drum EE, Krahn GL, Bersani Jr. H, eds. *Disability and Public Health*. American Public Health Association and American Association on Intellectual and Developmental Disabilities. Vol vii-viii. Washington, D.C.: American Public Health Association; 2009.
38. Krieger N, Testa C, Hanage WP, Chen JT. US Racial and ethnic data for COVID-19 cases: still missing in action. *The Lancet*. 2020;396:e81.
39. Wong NT, Glenn K, Whitley J, Wong A, Redwood Y. *Identifying Systems-level Barriers to Complying with and Expanding National Race & Ethnicity Data Disaggregation Standards [White Paper]*. Princeton, N. J.: Robert Wood Johnson Foundation; 2021.
40. Kaiser Family Foundation is tracking race and ethnicity data on state COVID-19 surveillance outcomes, including vaccinations. Kaiser Family Foundation. <https://www.kff.org/state-category/covid-19/covid-19-disparities/>. Published 2021. Accessed April 5, 2021.
41. Stoto MA. What did the 2009 H1N1 pandemic teach us about influenza surveillance systems? *Future Virology*. 2013;8(9).
42. Boyd RW, Lindo EG, Weeks LD, McLemore MR. On Racism: A New Standard For Publishing On Racial Health Inequities. In. *Health Affairs*. Vol 2021. Online: Health Affairs; 2020.
43. Hicks M. Fixing Tech's Built-In Bias. *American Scientist*. 2018;106(5):314.
44. Ito J. Supposedly 'Fair' Algorithms Can Perpetuate Discrimination. 2019. <https://www.wired.com/story/ideas-joi-ito-insurance-algorithms/>.
45. Kahn J. A.I. and tackling the risk of "digital redlining". In. *Fortune*: Fortune Magazine; 2020.
46. Discover The Power Of Street Wisdom. Streetwyze. <https://www.streetwyze.com/>. Published 2021. Accessed April 5, 2021.
47. Israel BA, Schulz AJ, Edith A, Becker AB. REVIEW OF COMMUNITY-BASED RESEARCH: Assessing Partnership Approaches to Improve Public Health. *Annual Review of Public Health*. 1998;19:173-202.
48. SAMHSA *White Paper Promoting Collaboration Between the Adverse Childhood Experiences and Suicide Prevention Fields to Identify, Assess, and Increase Resilience Among Youth with Strengthening the National Grassroots Action Infrastructure for Public Health and Social Justice: Moving from Mutual Aid and Protest to Power and Health*. Gaithersburg, MD: Community Science; 2021.
49. data.org. Rising Equitable Community Data Ecosystems (RECoDE) Project Announces the Formation of New Learning Council. data.org. <https://www.data.org/equitable-community-data-learning-council/>. Published 2021. Accessed April 5, 2021.
50. Vestal C. Lack of Public Data Hampers COVID-19 Fight. *Stateline*. 2020. <https://www.pewtrusts.org/en/research-and-analysis/blogs/stateline/2020/08/03/lack-of-public-data-hampers-covid-19-fight>.
51. CDC. Restricted-Use Vital Statistics Data. Centers for Disease Control and Prevention. National Center for Health Statistics Web site. <https://www.cdc.gov/nchs/nvss/nvss-restricted-data.htm>. Published 2020. Accessed April 8, 2021.

52. HRS. Accessing CMS Data. Institute for Social Research. <https://hrs.isr.umich.edu/data-products/restricted-data/cms-data>. Published 2020. Accessed April 8, 2021.
53. Camden Coalition of Healthcare Providers. Connecting Data. Camden Coalition of Healthcare Providers. Published 2021. Accessed April 20, 2021.
54. Allegheny County Analytics. Allegheny County Data Warehouse. Allegheny County Department of Human Services. <https://www.alleghenycountyanalytics.us/index.php/dhs-data-warehouse/>. Published 2021. Accessed December 17, 2020.
55. University N. Data Management: Federal Funding Agency Requirements. Northwestern University. Federal Funding Agency Requirements Website. <https://libguides.northwestern.edu/datamanagement/federalfundingagency>. Published 2011. Accessed January 4, 2021.
56. NIH. NIH Data Sharing Policy and Implementation Guidance. National Institutes of Health. Grants & Funding Website. https://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm. Published 2020. Accessed January 4, 2021.
57. NSF. Dissemination and Sharing of Research Results - NSF Data Management Plan Requirements. National Science Foundation. <https://www.nsf.gov/bfa/dias/policy/dmp.jsp>. Published undated. Accessed January 4, 2021.
58. CMS. *Medicare Promoting Interoperability Program Stage 3 Eligible Hospitals, Critical Access Hospitals, and Dual-Eligible Hospitals Attesting to CMS Objectives and Measures for 2018*. Centers for Medicare & Medicaid Services; 2018.
59. EHR. *Eligible Professional Meaningful Use Table of Contents Core and Menu Set Objectives*. Online: Centers for Medicare & Medicaid Services; 2014.
60. CMS. Participation Options Overview. Centers for Medicare & Medicaid Services. Quality Payment Program Web site. <https://qpp.cms.gov/mips/overview>. Published undated. Accessed December 4, 2020.
61. Carlin M, Ensign K, Person CJ, Kittle A, Meadows K. State of the Public Health Workforce: Trends and Challenges Leading Up to the COVID-19 Pandemic. *Journal of Public Health Management and Practice*. 2021;27(1):92-93.
62. Beck AJ, Boulton ML. Trends and Characteristics of the State and Local Public Health Workforce, 2010–2013. *American Journal of Public Health*. 2015;105:S303-S310.
63. Officials AoSaTH. ASTHO Profile of State and Territorial Public Health, Volume 5. ASTHO. <https://astho.org/profile>. Published 2020. Accessed October 14, 2020.
64. Medicine Io, Century CoEPHPfts. 6, Public Health Agencies: Their Roles in Educating Public Health Professionals. In: Gebbie K, Rosenstock L, Hernandez LM, eds. *How Will Keep the Public Healthy? Educating Public Health Professionals for the 21st Century*. Washington, D.C.: National Academies Press; 2003.
65. Zaller N. *Analyzing barriers to sharing data across criminal justice settings and with health settings to advance health outcomes and equity [White Paper]*. Princeton, N. J.: Robert Wood Johnson Foundation; 2021.
66. Hollis A. Climate Change Isn't Just a Global Threat—It's a Public Health Emergency. In. *Time*. Vol 2050: The Fight for Earth: Time Magazine; 2019.
67. Moore S. Climate Change and COVID-19. *News Medical Life Sciences*. 2021. <https://www.news-medical.net/health/Climate-Change-and-COVID-19.aspx>.
68. National Academy of Medicine Launches Initiative Supported by YouTube on Evaluating the Authoritativeness of Online Providers of Health Information [press release]. Washington, D.C.: The National Academies of Sciences Engineering Medicine 2021.
69. Medicine NAo. Principles for Defining & Verifying the Authority of Online Providers of Health Information. National Academy of Medicine. <https://nam.edu/programs/principles-for-defining-and-verifying-the-authority-of-online-providers-of-health-information/>. Published 2021. Accessed April 8, 2021.
70. Nelson C, Chandra A, Miller C. Can measures change the world. *Stanford Social Innovation Review*. 2018;16:1.
71. Forum WE. *Federated Data Systems: Balancing Innovation and Trust in the Use of Sensitive Data*. World Economic Forum; July 2019.
72. VMware. What is a federated network? VMware, Inc. <https://www.vmware.com/topics/glossary/content/federated-network>. Published 2021. Accessed 2021.
73. Bresnick J. *How Health Information Exchange Models Impact Data Analytics*. Health IT Analytics: xtelligent HEALTHCARE MEDIA; 2015.
74. McCarthy DB, Propp K, Cohen A, Sabharwal R, Schachter AA, Rein AL. Learning from Health Information Exchange Technical Architecture and Implementation in Seven Beacon Communities. *EGEMS (Wash DC)*. 2014;2(1):1060.
75. Tournon M. Centralized vs Decentralized vs Distributed Systems. In. *berty*. Vol 2021: berty; 2019.

76. Resources and Projects. North American Association of Central Cancer Registries. <https://www.naaccr.org/>. Published 2018. Accessed January 4, 2021.
77. Khurshid A. Applying Blockchain Technology to Address the Crisis of Trust During the COVID-19 Pandemic. *JMIR Med Inform*. 2020;8(9):e20477.
78. Platt R, Brown JS, Robb M, et al. The FDA Sentinel Initiative - An Evolving National Resource. *New England Journal of Medicine*. 2018;379(22):2091-2093.
79. Introducing CommonHealth. The Commons Project Foundation. <https://www.commonhealth.org/>. Published 2021. Accessed April 5, 2021.
80. Davies M, Erickson K, Wyner Z, Malenfant JM, Rosen R, Brown J. Software-Enabled Distributed Network Governance: The PopMedNet Experience. *eGEMs (Generating Evidence & Methods to improve patient outcomes)*. 2016;4(2):5.
81. OAuth. Wikipedia. <https://en.wikipedia.org/wiki/OAuth>. Published undated. Accessed January 4, 2021.
82. An open protocol to allow secure authorization in a simple and standard method from web, mobile and desktop applications. OAuth. <https://oauth.net/>. Published undated. Accessed January 4, 2021.
83. Introducing CommonHealth. CommonHealth. <https://www.commonhealth.org/>. Published 2020. Accessed January 4, 2021.
84. Chattu VK, Nanda A, Chattu SK, Kadri SM, Knight AW. The Emerging Role of Blockchain Technology Applications in Routine Disease Surveillance Systems to Strengthen Global Health Security. *Big Data and Cognitive Computing*. 2019;3(2):25.
85. Goldstein P. CDC Tests Blockchain Tech to Keep Track of Health Data. In. *FedTech*. Online2018.
86. Linn L. Blockchain for Health Data and Its Potential Use in Health IT and Health Care Related Research. Paper presented at: OSEHRA Innovation Webinar 2018.
87. Ramos MV, Jain R, Del Valle P, Torio C. *Beyond the Data: Is Blockchain the Answer to Improving Global Health?* Online: PharmExec.com; 2020.
88. Vuong B, Bauer T, Martinez A, Charlier C. *Transforming Data to Improve Health Equity: How to Do It and Who Needs to Be Involved* Princeton, N. J.: Robert Wood Johnson Foundation; 2021.
89. Sarikaya A, Correll M, Bartram L, Tory M, Fisher D. What do we talk about when we talk about dashboards?. *IEEE transactions on visualization and computer graphics*. 2018;25(1):682-692.
90. Matters MD, Lekachivili A, Savel T, Zheng ZJ. Developing metadata to organize public health datasets. *AMIA Annu Symp Proc*. 2005;2005:1047.
91. Michigan TRotUo. Data Lumos. The Regents of the University of Michigan. <https://www.datalumos.org/datalumos/>. Published 2021. Accessed January 4, 2021.
92. Schuh G, Doelle C, Schloesser S. Agile Prototyping for technical systems –Towards an adaption of the Minimum Viable Product principle. Paper presented at: DS 91: Proceedings of NordDesign 2018; 14th-17th August 2018., 2018; Linköping, Sweden.
93. Schilling MA. Toward a General Modular Systems Theory and Its Application to Interfirm Product Modularity. *The Academy of Management Review*. 2000;25(2):312-334.
94. Levinson M. *The Box: How the Shipping Container Made the World Smaller and the World Economy Bigger - Second Edition with a new chapter by the author*. 2nd ed. Princeton, N. J.: Princeton University Press; 2016.
95. ANSI. The American National Standards Institute oversees standards and conformity assessment activities in the United States. American National Standards Institute. <https://www.ansi.org/>. Published 2021. Accessed 2021.
96. Technology NloSa. Measure, Innovate. Lead. U.S. Department of Commerce. <https://www.nist.gov/>. Published Undated. Accessed 2021.
97. International H. Gravity Project. Health Level Seven International. <http://www.hl7.org/gravity/#:~:text=The%20Gravity%20Project%20seeks%20to%20identify%20coded%20data,four%20clinical%20activities%3A%20screening%2C%20diagnosis%2C%20planning%2C%20and%20interventions>. Accessed April 5, 2021.
98. University ToB. About; Why the 3-D Commission. Rockefeller Foundation-Boston University 3-D Commission. <https://3dcommission.health/about#:~:text=The%203-D%20Commission%20aims%20to%20bridge%20these%20gaps,science%2C%20health%20determinants%2C%20and%20policy%20and%20practice%20decision-making>. Published 2020. Accessed April 5, 2021.
99. Aliabadi A, Sheikhtaheri A, Ansari H. Electronic health record –based disease surveillance systems: A systematic literature review on challenges and solutions. *Journal of American Medical Informatics Association*. 2020.

100. Hatef E, Weiner JP, Kharrazi H. A public health perspective on using electronic health records to address social determinants of health: The potential for a national system of local community health records in the United States. *International Journal of Medical Informatics*. April 2019;124:86-89.
101. Miri A, O'Neill DP. Accelerating Data Infrastructure For COVID-19 Surveillance and Management. In. *Health Affairs Blog*. Vol 2020. Health Affairs; April 14, 2020.
102. CMS. Policies and Technology for Interoperability and Burden Reduction. Centers for Medicare & Medicaid Services. <https://www.cms.gov/Regulations-and-Guidance/Guidance/Interoperability/index>. Published 2021. Accessed April 4, 2021.
103. Technology TOotNCfHI. ONC's Cures Act Final Rule supports seamless and secure access, exchange, and use of electronic health information. HealthIT. <https://www.healthit.gov/curesrule/>. Published Undated. Accessed 2020.
104. NHPI. *Native Hawaiian and Pacific Islander (NHPI) Data Policy Lab Transforming Public Health Data Systems [White Paper]*. Princeton, N. J.: Robert Wood Johnson Foundation; 2021.
105. Glaser J. It's Time for a New Kind of Electronic Health Record. In. *Harvard Business Review*. Online: Harvard Business Publishing; 2020.
106. All In Home Page. All In. <https://www.allindata.org/>. Published undated. Accessed April 4, 2021.
107. Dash Connect Home Page. Data Across Sectors for Health. <https://dashconnect.org/>. Published 2021. Accessed April 4, 2021.
108. O'Neil S, Hoe E, Ward E, Goyal R, Staatz C. *Data Across sectors for health initiative: promoting a culture of health through cross-sector data networks*. Princeton, N. J.: Mathematica;2020.
109. Digital Bridge About Page. Digital Bridge. http://digitalbridge.us/about/?_sm_au_=iVVmrn5njT2nQfWV. Published 2021. Accessed April 4, 2021.
110. Past Projects. Digital Bridge. <https://digitalbridge.us/past-projects/>. Published 2021. Accessed April 4, 2021.
111. Current Projects. Digital Bridge. <https://digitalbridge.us/current-projects/>. Published 2021. Accessed April 4, 2021.
112. Implementation Networks. GO FAIR. <https://www.go-fair.org/implementation-networks/#:~:text=A%20GO%20FAIR%20Implementation%20Network%20%28IN%29%20is%20a,and%20self-governed%20consortia%20working%20across%20disciplines%20and%20countries>. Published undated. Accessed April 4, 2021.
113. Stewards of Change Institute Home Page. Stewards of Change Institute. <https://stewardsofchange.org/>. Published 2018. Accessed April 4, 2021.
114. Nissenbaum H. Privacy as Contextual Integrity. *Washington Law Review*. 2004;79(1).
115. Fact Sheet: SAMHSA 42 CFR Part 2 Revised Rule [press release]. Online: HHS Press Office 2020.
116. Price II WN, Spector-Bagdady K, Minssen T, Kaminski ME. Shadow Health Records Meet New Data Privacy Laws. *Science*. 2019;363(6426):448-450.
117. Wartenberg D, Thompson WD. Privacy Versus Public Health: The Impact of Current Confidentiality Rules. *American Journal of Public Health*. 2010;100:407-412.
118. Gostin LO, Hodge J, Valdiserri RO. Informational privacy and the public's health: The model state public health privacy act. *American Journal of Public Health*. 2001;91(9):1388-1392.
119. Hodge J, J. G., Kaufman T, Jaques C. *Legal Issues Concerning Identifiable Health Data Sharing Between State/Local Public Health Authorities and Tribal Epidemiology Centres in Selected U.S. Jurisdictions*. Online: CSTE; 2009.
120. California Consumer Privacy Act (CCPA). Stae of California Department of Justice. <https://oag.ca.gov/privacy/ccpa>. Published 2021. Accessed 2021.
121. Wolford B. What is GDPR, the EU's new data protection law? GDPR.eu. <https://gdpr.eu/what-is-gdpr/>. Published 2021. Accessed 2021.
122. Sobers R. *What is OAuth? Definition and How it Works*. Online: Varonis; 2018.
123. Services HaH. Proposed Modifications to the HIPAA Privacy Rule To Support, and Remove Barriers to, Coordinated Care and Individual Engagement. In: Services HaH, ed2021.
124. CDC. Provisions Relevant to Public Health Practice. Centers for Disease Control and Prevention. National Healthcare Safety Network Website. <https://www.cdc.gov/nhsn/hipaa/index.html>. Published 2015. Accessed April 8, 2021.
125. The Data Initiative Team. *The State of Data Sharing at the U.S. Department of Health and Human Services*. Office of the Chief Technology Officer U.S. Department of Health and Human Services; 2018.
126. Hodge J, J. G., Gostin LO. *Public Health Practice vs. Research*. Atlanta, GA: Council of State and Territorial Epidemiologists; 2004.

Transforming Public Health Data Systems

Who and what next?
Opportunities for a new kind
of modernization

This white paper is part of a series of papers for use by the National Commission to Transform Public Health Data Systems, an independent commission convened by the Robert Wood Johnson Foundation (RWJF). These papers were used to inform Commission discussions and recommendations. The papers represent insights from literature review, a diverse set of stakeholder perspectives, and inputs from RWJF grantee projects related to the topic of the Commission. The white papers were developed in support of the Commission by the RAND Corporation, a nonprofit research organization, within its division, RAND Social and Economic Well-Being. Corresponding authors are [Anita Chandra](#) and [Laurie Martin](#).

**National Commission to
Transform Public Health
Data Systems**

May 2021

Table of Contents

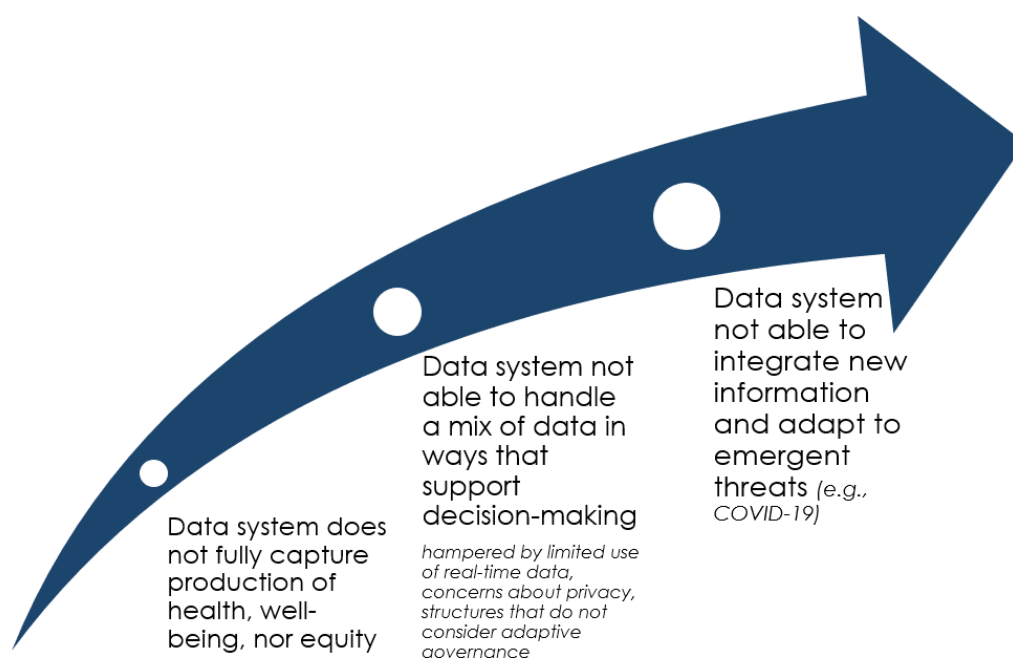
Introduction	2
Revisiting Key Actions	3
A Call to Action.....	5
Federal government	6
State government.....	9
Local government.....	10
Health care systems	11
Philanthropy.....	11
Technology companies.....	12
Nonprofit/Community-Based Organizations.....	13
Professional associations (e.g., NACCHO, ASTHO, NASEM)	14
Schools of public health and policy	15
How Far Away Are We?	16
Considerations from Systems Implementation Research that Can Inform a Path Forward	18
Conclusion.....	20
Acknowledgements	21
References	21

Introduction

These white papers, collectively developed for the National Commission, outline a series of key actions and critical considerations for transforming the public health data system into one that meets the principles in the preamble—that is agile, can connect multiple forms of data, translate those data into information and insight, and is fully equipped for the opportunities and challenges of the 21st century. Critical to the success of this public health data system is a focus on equity in the content and structure of data, how data are used to address systemic inequities, and the engagement of diverse stakeholders in the access to and use of data for decision-making—all with attention to issues of power, marginalization, and justice.

COVID-19 and the current national awakening about social inequities provide a long overdue moment in which national consciousness may have been “switched on” about health; as such, the public health data system modernization must be different in its ambition to capitalize on that momentum (**Figure 1**).

Figure 1. Forcing factors in motivating public health data system transformation



There is an urgent opportunity to not only pursue data modernization as we have in the past, importantly addressing issues of data content and quality and evolving data infrastructure. But it is also a time to take the more system-minded and transformation-minded^a approach, which values a more purposeful look at what we are measuring and what health values for the nation these data choices are transmitting, how equity is centered and consistent across sectors, and how the public health data system embraces equity as both an ends and a means in its vision. While the general goals of improving health and well-being are important, this is a time for laser-focused specificity to help diverse sectors and stakeholders see more clearly how health contributes to the hope, promise, and potential of the nation. As noted in the preamble and throughout the prior white papers, this is a time for public health to crystallize a “moonshot” or a more ambitious north star to provide clear goals for health, well-being, and equity, which focuses on positive health and thriving. This shared vision can help sectors and stakeholders accelerate action, commit to the process of transforming the public health data system through their own work, and join forces to amplify impact. In this final paper of this series, we review key gaps and actions identified in earlier papers, discuss opportunities for sectors and stakeholders to catalyze action around this vision, and provide insight into windows of opportunity, given federal and other national actions, which may accelerate such transformation.

Revisiting Key Actions

Earlier white papers laid out a series of key actions that address gaps in the current public health data system (**Table 1**). Collectively, these actions present an opportunity to reflect on how a transformed public health data system could, and should, be more intentional about equity throughout all aspects of the public health data system. Equity continues to be a difficult concept for many to fully understand in part because it represents both a process and an outcome, and includes various types of equity, as noted in the Why? paper. Equity matters not just for the types of data that are collected, but who is included in the public health data system, whether the nation and communities can collect and present data on heterogeneous populations that have diverse cultures, languages, and lived experiences (as opposed to lumping diverse populations together, masking important differences), and who has access to those data to support sense-making and action, particularly at the local level. Although these objectives appear seemingly straightforward, achieving and sustaining them will require the innovation, support, and collective action of diverse sectors, systems, and populations represented within these data.

^a System-minded means “seeing the system as a system, understanding system dynamics, and then shaping intentions, decisions, and actions in terms of this new perspective.” Transformation-minded means “shaping intentions in terms of deeply transformed system behavior, and *framing success* in terms of system level transformation, shaping the new types of *agency and action* that will actually result in the system veering towards altered states.” Adapted from an interview with Banny Banerjee, Stanford University ChangeLabs from an article by Leonard Teichert, *Medium* 2019.

Table 1. Key actions for a modern, equity-oriented public health data system
(from *Why?*, *What?*, and *How?* papers)

The equity orientation in the modern public health data system (Why?)
<ul style="list-style-type: none"> • A modern public health data system should be more intentional about factoring in different types of equity to truly make progress on health equity. • Given the multiple dimensions of equity, revisit how equity is presented in current equity-based datasets. • The public health data system needs to be actively reoriented to facilitate agency by actors and sectors.
The content of data to prioritize in a transformed public health data system (What?)
<ul style="list-style-type: none"> • Ensure the public health data system promotes national, regional, and local data alignment, using a parsimonious set of core public health measures, and with attention to consistent use of new forms of public health data. • Assess whether the content of public health data currently prioritizes and facilitates timely, proactive, and evidence-based decisions going forward. • Efficiently and ethically use the growing volume of public health data across sectors. • Clarify the appropriate level of precision and granularity of public health data needed for public health practitioners, policymakers at different levels, and the public to make informed and timely decisions.
The design of the modern public health data system (How?)
<ul style="list-style-type: none"> • Develop a governance structure that is equitable, leverages the diversity of data across sectors, supports timely decision-making at multiple levels, and builds in adaptive capacity. • Consider the need for an independent, trusted convener, who or which can bridge sectors to facilitate development of an equity-oriented data system. • Revise principles of data stewardship and data governance to keep pace with rapidly emerging technologies for data collection, analysis, and reuse. • Ensure that equity remains a guiding principle in data sharing efforts. • Increase participation in data sharing efforts. • Identify appropriate models for information exchange. • Strengthen interoperability of data systems. • Ensure laws strike a balance between the protection of privacy and benefits of using data for public good. • Revise language in the Privacy Rule and other statutes to encourage sharing of identifiable data with public health.

The section that follows reviews opportunities for stakeholders to leverage their expertise, reach, reputation, and resources to meaningfully transform the public health data system to one that not only tracks health and well-being but also one that can support and inform proactive and dynamic solutions to current and emerging health priorities.

A Call to Action

The recognition that the current U.S. public health data system is no longer sufficient for our emerging needs, coupled with a renewed interest in data transformation efforts, and a realization that the necessary changes are not going to come about by working incrementally or around the edges, motivates a galvanizing call to action. While there are quick wins ahead, true transformation will require meaningful engagement of diverse sectors, and an authentic commitment to tackling tougher issues that challenge the way U.S. public health leaders and collaborators work together, a collaboration that generally has eluded the nation. How this is approached offers an important set of questions.

- Given that the nation's chronic and infectious disease burden has exposed economic and social frailties, and that our collective mental health challenges remain at crisis levels, what is the nation's commitment to public health?
- How does the public health data system move beyond a reactive one to a proactive one, better aligned with the requirements for 21st century public health action?
- How do we strengthen the alignment of priorities and incentives so that all sectors and stakeholders are pushing in the same direction with respect to using data to inform policy and action?
- In addition to incentives, is there a need for something akin to a public health compact for the 21st century, for which all sectors that influence health sign on to streamline the collection, analysis, and dissemination of health data?

Below, we offer high level considerations for a diversity of sectors. These are not intended to be exhaustive or final recommendations, but rather a list of ideas for sector-specific contributions that envelop a different type of shared public health compact or national agenda. It is expected that while such ideas may inform the development of Commission recommendations, the final recommendations will flow from the vision and structure of a modern, equity-oriented public health data system, which will be more fully articulated by the Commission through its deliberations. Unless otherwise specified, considerations are not only tied to the health part of government or the private sector, but, again, build from the more holistic definition of public health offered in the preamble (e.g., inclusive of sectors such as education and transportation).

These considerations are organized into four categories mapped to the thrusts of the preamble and three other papers—**narrative**; **equity orientation**; **data content**; and **governance, partnership, and coordination**—as action in each of these areas is needed for coherence and movement.

Federal government

Over the first few months of 2021, President Biden issued a number of executive orders, passed plans to support critical infrastructure, and established working groups that signal a commitment to strengthening the public health data system (**Table 2**). COVID-19 amplified the potential power of the federal government both in the need for consistency of public health information and the expectations around the roles of federal government for unpacking and addressing structural inequities. Despite this promise, how these investments will ultimately play out at the federal, state, and local levels is unknown, and activity is underway across federal agencies (e.g., health, education, justice) on how to operationalize and implement many of these activities. The federal government has an opportunity in this push for modernization and equity orientation, through these types of considerations:

- **Narrative:** Prioritize a smaller set of public health measures for a more unifying and inspiring national health vision, which are regularly described in federal communications, including those not typically seen as “health,” thereby ensuring a focus on other sector impacts and the connection to health.
- **Equity orientation:** Advance efforts to rebuild trust in public health data, both in how data are explained and illuminate longstanding systemic inequities and in how diverse groups communicate public health information.
- **Data:** Integrate standardized expectations about data collection, presentation, and disaggregation in federally-funded data collection efforts, including subgroup analyses and equity contextualization.
- **Governance, partnerships, and coordination:** Ensure that the bold ideas and systems transformation embedded in this orientation to the future public health data system are operationalized through a combination of policies, funding mechanisms and priorities, and guidance to state and local agencies. Develop principles for equitable data access to ensure that a diverse group of stakeholders can utilize and leverage federal data for decision-making. Strengthen public health infrastructure, including data infrastructure, through critical financial investments.

Table 2. 2021 Federal government activities related to equity and the public health data system (selected)

Type	Opportunities for public health data system transformation	Summary	Cost or Scope
Executive Orders			
Executive Orders	Executive Order On Advancing Racial Equity and Support for Underserved Communities Through the Federal Government: Equitable data working group	Order to gather data that is disaggregated by race, ethnicity, gender, disability, income, veteran status, or other key demographic variables to advance equity. ¹	Focus on federal data structures
Executive Orders	Executive Order on Ensuring an Equitable Pandemic Response and Recovery	Order to gather COVID-19-related data that is disaggregated by race, ethnicity, gender, disability, income, veteran status, or other key demographic variables to ensure an equitable pandemic response. ²	Focus on COVID-19-related data
Legislation			
Legislation	American Rescue Plan	Order to expand coverage and resources for health insurance coverage and improve access to mental health services and community-based programs that address social determinants of health (SDOH). ³	Specific to the 2021 and 2022 health insurance marketplace
Legislation	American Jobs Plan	Legislative language to improve infrastructure in rural areas, including work to deliver clean drinking water, a renewed electric grid, and high-speed broadband to all Americans. ⁴	Over \$300 billion in investments
Legislation	The Improving Social Determinants of Health Act of 2021	Legislative language to coordinate across the Centers for Disease Control and Prevention (CDC) to incorporate social determinants of health into grants and activities, award states, local, territorial, and tribal health agencies to address SDOH in target communities, award grants to nonprofit organizations and institutions of higher education to conduct research on SDOH best practices, coordinate and align SDOH activities with other agencies, and collect and analyze data related to SDOH activities. ⁵	Coordination among all CDC programs and \$50 million annually for program activities

Type	Opportunities for public health data system transformation	Summary	Cost or Scope
Working Groups/Commissions			
Working Groups/Commissions	IHEC Data Workgroup	Establish an Interdepartmental Health Equity Collaborative (IHEC) Data Workgroup that includes experts engaged in data activities and data policies that address social determinants of health and health disparities. The goals of the group are to identify existing policies and practices for improving access to data and use of data in support of policy development and to promote data projects and applications that address social determinants of health and support efforts to advance health equity. ⁶	N/A
Working Groups/Commissions	Implementation of Healthy People 2030	The Health and Human Services (HHS) Secretary's Advisory Committee on National Health Promotion and Disease Prevention Objectives for 2030 was a federal advisory committee responsible for making recommendations for developing and implementing the national health promotion and disease prevention objectives, known as Healthy People 2030; this work is ongoing and now focused on implementation.	National, state, and local
Working Groups/Commissions	Regional commissions funding rural SDOH programs	Appalachian Regional Commission funding to promote economic development, education, health care access, and physical infrastructure in distressed Appalachian communities. ⁷	\$10 million annually in support
Working Groups/Commissions	Regional commissions funding rural SDOH programs	Delta Regional Authority funding in public and transportation infrastructure, as well as business and workforce development, in rural communities in the Mississippi River Delta region. ⁷	\$30 million annually in support
Working Groups/Commissions	Regional commissions funding rural SDOH programs	Area Health Education Centers (AHEC) funding and support to train future and current rural health professionals, addressing the educational and access needs of rural communities, including the Alaska Center for Rural Health and Health Workforce offers K-12 pipeline programming in health career education. ⁷	\$28 million annually in support

Type	Opportunities for public health data system transformation	Summary	Cost or Scope
Funding Opportunities			
Funding Opportunities	Federal agencies funding rural SDOH programs	CDC grants to address SDOH, including Racial and Ethnic Approaches to Community Health (REACH) program in areas such as neighborhood infrastructure, food access, and access to opportunities to increase physical activity. ⁷	Programs in all states
Funding Opportunities	Federal agencies funding rural SDOH programs	U.S. Department of Housing and Urban Development funding to support rural affordable housing and housing quality initiatives, such as the Rural Capacity Building for Community Development and Affordable Housing Program. ⁷	Each program has approximately \$15 million funded annually
Funding Opportunities	HHS funding for data interoperability for SDOH	Office of the National Coordinator for Health Information Technology (ONC) support to address collection, access, exchange, and use of SDOH data through funding opportunities such as the Leading Edge Acceleration Projects (LEAP) in Health Information Technology by advancing health IT standards and tools to improve social determinants of health data exchange and research. ⁸	Up to \$2 million in funding for LEAP areas of interest

State government

Although federal policy and federal funding can shape a new vision of a modern, equity-oriented data system, state governments hold immense power to reinforce these signals or to counter them through their own customized policies and resource allocation decisions. Within the COVID-19 pandemic, for example, states varied considerably on the types of data collected and the ease at which those data were accessible to and shared with decisionmakers across the state working to develop a localized, coordinated response. State governments also commonly have a workforce with a more diverse expertise, including those that are highly skilled at data manipulation and informatics, which is often not available at local levels. State governments, as a result, can play a critical role in the evolution toward and sustainment of, an equity-oriented public health data system through these types of opportunities:

- **Narrative:** Align approaches to data translation into *information* and *insight* (see Principle 1, preamble)—in short, what the knowledge generation process looks like over time for state trend analysis and how data are equity—contextualized at the state versus local or national levels.
- **Equity orientation:** Assess existing state policies related to data collection, sharing and privacy through an equity lens,^b and consider modifications to strengthen the public health data system.
- **Data:** Develop processes that allow local stakeholders to easily access relevant data at a level of disaggregation that is useful for local decision-making and for supporting special populations in the state.
- **Governance, partnerships, and coordination:** Partner with local health departments, and departments that provide health data (e.g., social services) to consider new models of working together to improve efficiency and timeliness of decision-making at the state and local levels.

Local government

The governance capacity for counties, cities, municipalities, and tribes varies significantly based on population size and resources. Despite this variability, however, each represent local populations and interests and are tied to local systems and sectors like education, social services, law enforcement, health care, and business, all of which have potentially important roles to play in ensuring that relevant data are collected that signal the health and hope of the community. Although connections across systems also occur at the state and federal level, local government may spearhead collaboration and the development of locally tailored solutions to addressing both emerging and longstanding needs (see Figure 3, preamble). Local government is in the best position to ensure that the community is reflected in the public health data system and an increasing number of localities have made commitments to address equity in their community through policies, the establishment of equity offices, or the declaration of racism as a public health issue. The strengthening of the public health data system could be further supported by the following:

- **Narrative:** As with federal government efforts on national health messaging, regularly connect public health data to local communications (e.g., what is happening in the community and how it relates to overall community well-being) above and beyond that from the local health department.
- **Equity orientation:** Ensure local voices are represented not only in the data itself (e.g., from whom data are collected), but in positions of authority responsible for sense-making and local decision-making based on the data.
- **Data:** Share relevant data freely across agencies within the government (e.g., infrastructure, economic development, social services) to inform a more complete and robust understanding of emerging needs and potential solutions.

^b An **equity lens** is defined in these white papers as any approach for analyzing the equity impact of actions on people or places, with attention to those historically marginalized, as well as the process of addressing any structural or systemic barriers to the opportunity to be healthy (*adapted from multiple sources*).

- **Governance, partnerships, and coordination:** Explore data sharing collaborations across government and civil society (e.g., nonprofits, businesses) that can more consistently generate data to support contextual equity considerations. Strengthen local public health infrastructure and workforce capabilities through innovative models that leverage expertise from research and industry.

Health care systems

In large part due to Center for Medicare and Medicaid Services (CMS) initiatives, health care systems have increased the collection of data on social determinants of health and other “non-clinical” factors that influence health, at the individual levels primarily. However, the motivation for collecting such data is largely one of cost savings and improved care coordination. Willingness to share data is often met with reluctance, in large part due to a lack of clarity in the Health Insurance Portability and Accountability Act (HIPAA) and other state laws governing data sharing. Although privacy laws do allow data sharing with public health, health care systems often put the onus on public health to demonstrate need. Data from health care systems also tend to lack critical information on race, ethnicity, and other characteristics. Although health care systems do sometimes share data with community partners, in general a strong and universal commitment to sharing data for the public good has not occurred. A major challenge to overcome is the misalignment of incentives: sharing data that would improve population health has the potential to lower revenue for many health systems, particularly those using fee for service models. At the same time, as payment reform models increasingly value quality over quantity, these incentives may become more aligned. Health care could be a powerful ally in a strengthened public health data system through these types of opportunities:

- **Narrative:** Support a health vision that prioritizes the production of health or positive health outcomes, and not simply prevention of disease.
- **Equity orientation:** Collect (or update/confirm) social determinants of health data at every consumer encounter, using standardized questions that allow data aggregation within communities. Connect these data on individual social needs (e.g., housing), with structural inequities confronted by consumers (e.g., neighborhood characteristics, trauma exposure, experiences of discriminatory policies) in order to offer the community (or region) useable information on distributive and contextual equity.
- **Data:** Collect demographic information, in sufficient detail, for every consumer encounter, including primary care, specialty care, and laboratory services.
- **Governance, partnerships, and coordination:** Overcome historic siloes and build partnerships to facilitate sharing of relevant health care data with public health departments in a timely and efficient manner, in a way that allows for data linkages and disaggregation of subgroups and geographic regions.

Philanthropy

Philanthropic organizations, whether national or regional in reach, can play an important role in helping to catalyze a modern, equity-oriented data system. Although philanthropy cannot, and should

not, underwrite the full scope of data modernization efforts, these organizations may offer strategic and catalytic investments to support methodologic advancements, or deeper dive thinking on critical issues that have served as important barriers to true data system transformation. Through grantmaking, philanthropic organizations can help guide a national or regional agenda, by signaling important gaps and needs worthy of additional attention (e.g., positive health and well-being, structural inequity data). Philanthropy can serve an important convening function, bringing together diverse stakeholders to share ideas, challenge norms, and generate new solutions, and to share and support the spread and uptake of evidence-based solutions. Philanthropy could support these types of opportunities:

- **Narrative:** Advance the benefits of measure and data parsimony and the need for a national, unifying health goal that can be well-understood and executed, with a focus on positive health, equity, and well-being.
- **Equity orientation:** Leverage convening and funding opportunities to meaningfully engage nontraditional partners, including non-health sectors and diverse community members. Ensure there is a priority on the role of historically marginalized populations in data decision-making.
- **Data:** Align investments to minimize duplication of funded efforts and leverage overlapping interests specific to the public health data system.
- **Governance, partnerships, and coordination:** Focus funding on the processes of using critical systems approaches, governance strategies, and disaggregation to advance thinking on complex sticking points of data system transformation. This is particularly important in addressing issues of how technology upgrades align with data needs, how data are accessed and translated, and how data are tied to other accountability mechanisms.

Technology companies

Over the past decade, the world has observed a growth in technology companies' interest in health, with the expansions of smart phones, use of wearable devices, and application of medical devices (the Internet of Bodies).⁹ The amount of health data that is collected on individuals and populations on a daily basis is astounding (see What? paper). Very little of these data, however, are consistently used to identify emerging health needs for the public good or to inform local decision-making. Technology companies also have a wealth of talent, attracting top candidates in informatics, 'big data' analytics, and data security, and are often at the cutting edge of new technologies and approaches to finding signal value within vast amounts of data. At the same time, there is a larger political and societal question about the role of big technology companies like Google, Facebook, and Twitter. Current Congressional hearings and other federal policy discussions are grappling with important questions such as how technology companies should behave and what their role in society should be. These conversations and philosophical questions may open a new door, previously less available, for companies to be more willing to leverage their data, resources, and expertise for public good, and to become powerful allies in crafting a modern, equity-oriented data system through these types of opportunities:

- **Narrative:** Work with government partners and other organizations to develop standards on which health data are communicated, particularly as it is translated into information and insight, and reinforces data integrity and credibility.
- **Equity orientation:** Engage in more consistent workforce exchanges with public health organizations to cross-pollinate innovations in health data content and application. Emphasize workforce exchanges to increase the leadership of underrepresented communities and populations.
- **Data:** Supplement on-the-ground capacity with new models of public-private partnerships or cooperative agreements that augment the quality and precision of data content and use.
- **Governance, partnerships, and coordination:** Commit to consistently share certain data with public health and feed relevant data into national datasets that may improve the agility, timeliness and comprehensiveness of our national data and signal emerging public health concerns. Foster innovation related to data integration, ‘big data’ analytics, and data security and adoption of existing methods currently in use by technology companies.

Nonprofit/Community-Based Organizations

Nonprofits and community-based organizations work directly with community members and provide a strong on-the-ground understanding of the existing and emerging health needs of the community. Many also bring a unique perspective and understanding of equity, including lived experiences and missions that are working to offset the impacts of intergenerational structural and systemic inequities within communities. Nonprofits and community-based organizations serve as an outlet for local voices in the community, both through work that may lift up community voice, or as part of organizational leadership structure. As trusted institutions in the community, nonprofits often are an important link to populations, which have been marginalized or harmed by research, policies, or programs. These organizations can provide insight into rebuilding trust and increasing local participation in the data collection, sense-making, and decision-making processes.¹⁰ If authentically engaged, nonprofits and community-based organizations can be important partners in a modern, equity-oriented data system through these types of opportunities:

- **Narrative:** Lift up community voice and perspective through the active engagement of not only data collection activities (e.g., what is collected and from whom), but also sense-making and decision-making to ensure that data reflect community context and priorities.
- **Equity orientation:** Build trust in data as a public good, voicing concerns (e.g., around privacy or how data will be used) and support the development of solutions that address such concerns to ensure that such data do not further marginalize or bring harm to populations.
- **Data:** Identify priority sources of data from national nonprofits that can feed into the public health data system consistently, both nationally and through local affiliates.
- **Governance, partnerships, and coordination:** Work locally to develop data sharing strategies with government, where appropriate and where there is health benefit, reciprocity, and equitable access.

Professional associations (e.g., NACCHO, ASTHO, NASEM)

Professional associations, including scientific organizations such as the National Academies of Sciences, Engineering and Medicine (NASEM), are trusted organizations and often the go-to place for information and resources among members. Even for organizations not formally affiliated with a specific member association, professional associations can serve as a resource through dissemination of information, briefs, issue statements, and other resources that are put in the public domain. For instance, the National Academy of Medicine (NAM) recently produced an analysis of lessons learned from COVID-19 for future public health infrastructure.¹¹ These dissemination channels coupled with member organizations spanning regions, states, or the nation, can be an asset for ensuring that information is disseminated quickly. Professional associations are comprised of relevant stakeholders, who bring a diversity of expertise that could be leveraged to establish a set of standards of practice, as well as standardization of measures and data collection approaches that are relevant, feasible, and sustainable. With these assets in mind, professional associations could fill an important role in the transformation towards a modern, equity-oriented data system through the following types of opportunities:

- **Narrative:** Educate members about the importance of public health data, and data for the public good, as a powerful tool for decision-making.
- **Equity orientation:** Develop national standards for ensuring equity is reflected in public health data collection, sense-making, and decision-making efforts, including standardization of how demographics are collected and the types of data that should be uniformly collected.
- **Data:** Develop strategies to improve data completeness and quality, particularly with respect to equity and drivers of health, hope, and well-being.
- **Governance, partnerships, and coordination:** Actively work with professional associations from other sectors that influence health (e.g., education, social services, environment) to align efforts, particularly around standardization to minimize variability across sectors that impedes data sharing and health sense-making at a local level.

Schools of public health and policy

A critical question surrounding a modern, equity-oriented data system is whether the public health workforce has the skills and capacity to leverage a more robust data system. While efforts like the Association of Schools and Programs of Public Health (ASPPH) *Framing the Future*¹² outlines the next stage of public health education focused on new capacities needed for the field, it is unclear how much the skills around data—from equity orientation to knowing how to translate data into information and insight—actually have been realized. For instance:

- Is the current workforce pipeline generating graduates steeped in an understanding of equity and the multi-faceted and multi-generational impacts on health?
- Are graduates taught to then apply that understanding to meaningfully engage with those who can partner on sense-making and decision-making activities to ensure those perspectives are not lost?
- Does student training include training on how to balance the precision of research with more rapid sense-making with data that may be necessary in a public health emergency and for which data may be less precise?
- Does student instruction include a value on the diversity of public health data, including qualitative data and authentic lived experiences of those within the community in which they will be working?

Similar questions can be posed of faculty and researchers, who have an opportunity at this point to assess their own biases and approaches to working and consider ways in which their work could contribute to a modern, equity-oriented public health data system. Further, this is not about just schools of public health, but other allied fields, including schools of public policy.

- **Narrative:** Shift orientation away from monitoring and reporting of risk factors and disease outcomes towards one that promotes positive health, hope, and well-being.
- **Equity orientation:** Help students monitor and evaluate structural and systemic inequities and understand what data systems must include that signals improvements in those inequities.
- **Data:** Train students at schools of public health and public policy about decision-making under deep uncertainty,^{c,13} as well as sense-making from diverse data sources, some of which may be incomplete or evolving during public health emergencies.¹⁴⁻¹⁶
- **Governance, partnerships, and coordination:** Train the next generation of public health practitioners and researchers to meaningfully partner with local stakeholders, and to value lived experience and community expertise as much as formal training.

^c Deep uncertainty refers to the level of uncertainty of the event itself or to the level of uncertainty we have about our own knowledge. Deep uncertainty relates to a situation in which events (e.g., catastrophic) cannot be addressed by assigning probabilities to its occurrence.

How Far Away Are We?

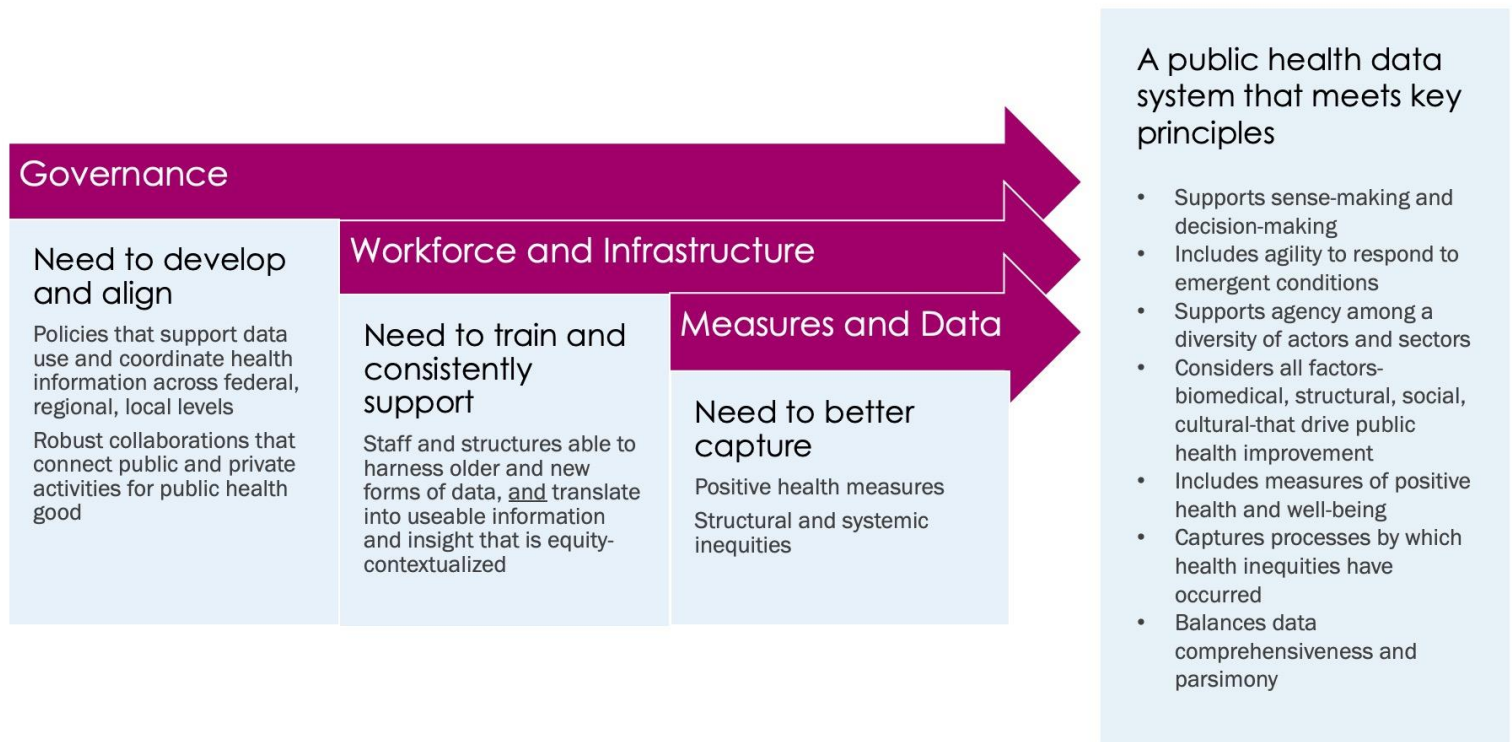
The white papers offered key actions needed for full modernization and called out a wide range of efforts, innovations and collaborations that are underway. *With all this activity, it is reasonable to ask, how far away are we really from achieving this vision?*

The answer depends, in part, on what aspect of the system we are considering (see **Figure 2**). Perhaps the biggest challenge is that of governance, where the chasm between where we are and where we want to be is both wide and deep. Governance is critical to this effort if we are aiming for true transformation, as it signals the development and alignment of policies across sectors and federal, state, and local levels that support data sharing and use, and the establishment of robust collaborations that connect public and private activities for public good. The current distributed public health system is comprised of thousands of different organizations, and this number will only increase as data from other sectors is used for public good. Underpinning the governance issue are several fundamental questions: who is responsible, what are they responsible for, and what authority do they have to implement that responsibility? In other words, whose job is it to ensure that all of the fish in the ocean are swimming in the same direction?

The current system is quite misaligned with respect to workforce capabilities and infrastructure needed to be able to harness both new and older forms of data and to translate that into information and insight that is equity contextualized. Here it is important to call out the ongoing defunding of public health infrastructure, which has resulted in legacy systems that fail to meet current and future needs and has stripped many public health departments of staff and resources. As noted in the preamble, a transformed public health *data* system must sit upon a modern public health system, which recognizes the health needs of the future. Achieving a modern, equity-oriented data system is going to require widespread acknowledgement of what is actually driving health and well-being in the community today, what is that “moonshot” that fulfills American’s health promise and not merely disease mitigation and cost containment, and what investments in both workforce and infrastructure will be necessary to achieve this vision.

Finally, a public health data system that meets the key system principles required for a modern, equity-oriented public health system (described in the preamble and in the right box of **Figure 2** below) has to be able to capture both positive health outcomes and structural and systemic inequities. Although we still have far to go, current efforts working to address these measures and data gaps are underway. It will be important to acknowledge, however, that while the *development of measures and strategies to collect data to populate those measures* are critically important, this comprises only one aspect of a public health data system. To implement the use of new measures, but not address gaps in infrastructure and workforce, and ignore systems-level governance considerations will fall well short of achieving a modern, equity-oriented public health data system.

Figure 2. How far away are we from achieving a public health data system that meets key principles?



Considerations from Systems Implementation Research that Can Inform a Path Forward

Although the Commission is not tasked with developing implementation guidance, consideration of how such recommendations may be implemented can inform their development, increasing the likelihood that the recommendations will achieve the intended result. A growing body of implementation research, particularly at the programmatic and policy levels, uses process theories or implementation theories to explain factors that facilitate or impede implementation. These theories and frameworks tend to focus on factors at the leadership, practice, resource, systems, and community levels that either facilitate progress and successful outcomes or impede results.¹⁷ **Table 3** summarizes some of the key features identified in these theories as important for systems change implementation efforts. Some of these features come from the Stewards of Change, Human Service Systems 2.0 Framework, while other features are cited in systems change literature. These features are critical reminders of necessary components for effective implementation of systems change.

Table 3. Key features in systems change implementation for Commission consideration

Key Features
A governance structure clarifies the purpose of data and how data are translated into sense-making, decision-making, performance, and accountability.
Policies and plans manage silo problems (e.g., lack of connection, communication).
Workforce development and training is well-aligned to the timing and pace of system changes.
A technology framework aligns with systems changes (e.g., updates in technology map to updates in the overall system mission and functions).
Systems level change management considers social context, and how individuals in that system learn to change over time (i.e., using social learning theory, active sense-making processes as different parts of system changes take root) ¹⁸

In addition to the features in **Table 3**, there are two other considerations for the Commission in terms of how the entirety of the public health data system transformation should be viewed. First, the public health data system should be viewed as a **Complex Adaptive System**, rather than trying to shoehorn something more linear. A complex adaptive system is a “a dynamic, self-similar collectivity of interacting, adaptive agents and their artefacts.”¹⁹ Said more operationally, when trying to address complex challenges with a vision of big scope, such as health, hope, and well-being, the systems change plan must squarely address issues of **interconnectedness**, **diffusion of authority**, and **systemic uncertainty**. The interconnectedness comes from the multiple sectors and actors; the diffusion emerges from the tiered approach to public health from national to local; and the uncertainty is core to a system that seeks to be more agile and adaptive, as well as proactive rather than reactive. In addition to these complexities for data system change is that health is intrinsically connected to notions of personal responsibility, social values, and human rights,²⁰ which add

18 Transforming Public Health Data Systems: Who and what next? Opportunities for a new kind of modernization

complexity to many of the equity and ethical choices noted in prior papers regarding systems design for whom and by whom.

A second point for the Commission before proceeding is what **philosophy of systems change** should be pursued given the calls for transformation and system-mindedness noted throughout these white papers. In short, the Commission must consider the entry points for transformation:

- Should this transformation in the public health data system or systems adopt a new, universal guidance for change (*national movement*)?
- Should transformation build more quickly on regional or local innovations and push those transformations from demonstration projects to something that could be conceived as a tipping point for the mobilization of more transformation (*coordinated acceleration*)?
- Should transformation call on the most impactful influencers for big change first, then move through tiers of other influencers (*stepped change*)?

Conclusion

There are many ways to approach this “golden” moment that the public health community and the nation finds itself because of the events of 2020—one way pushes the public health field outward, one pushes the field to look inward. Both ways are key to this process of transforming the public health data system to a system that is actually modern and equity-oriented.

In an article on the importance of not wasting the experience of COVID-19 for public health, Brownson et al. (2020) stated, “The pandemic reminds everyone of the value and necessity of public health.”²¹ Outwardly, the criticality of public health to the nation’s future has been understated for too long despite the urgency of public health and its data for the social and economic hope of the United States. While this is a moment for public health to advance the importance of its role in American society, there is a concurrent field introspection that also is needed. In an article on new standards for publishing on racial health inequities, Boyd et al. (2020) pushed scientific publications to unpack data on inequity in ways that go deeper and explain the processes by which inequity (e.g., racism) have occurred,²² a point also made in one of the key principles for a transformed public health data system noted in the preamble and in **Figure 2**. In recent writings, Petteway (2021) and Sirry (2021) each thoughtfully examined white supremacy within public health, challenging assumptions about what data are selected, how those data are contextualized, and how data are applied, including via public health’s essential functions (assurance, assessment, policy development).^{23,24} It is in this convergence of outward (public health values and importance) and inward (public health redefining in ways that center equity) where the Commission must situate its recommendations to transform the public health data system.

Public health data and the data systems that support collection and application of those data are a leading symbol of how we realize our national health mission going forward—how we choose to work in a deeper covenant with diverse sectors and to lead with the inputs of populations most marginalized reflects our *values*; the choice of measures represents our *priorities*, and the data infrastructure we develop to act on those priorities in timely and proactive ways demonstrates our *will*.

Acknowledgements

Each paper in this series of white papers has been reviewed by Mahshid Abir of the RAND Corporation and the University of Michigan and Karen Smith of Public Health Strategies. We appreciate these reviewers for their insights and contributions. We also thank Burness for their editorial and communication support for each paper.

References

1. Biden Jr. JR. Executive Order On Advancing Racial Equity and Support for Underserved Communities Through the Federal Government. The White House. <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/20/executive-order-advancing-racial-equity-and-support-for-underserved-communities-through-the-federal-government/>. Published January 20, 2021. Accessed April 12, 2021.
2. Biden Jr. JR. Executive Order on Ensuring an Equitable Pandemic Response and Recovery. <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/21/executive-order-ensuring-an-equitable-pandemic-response-and-recovery/>. Published January 21, 2021. Accessed April 12, 2021.
3. American Public Health Association. APHA applauds passage of American Rescue Plan Act of 2021 as a step toward healing and recovery. <https://www.apha.org/News-and-Media/News-Releases/APHA-News-Releases/2021/american-rescue-plan>. Published 2021. Accessed April 12, 2021.
4. The White House. Fact Sheet: The American Jobs Plan. <https://www.whitehouse.gov/briefing-room/statements-releases/2021/03/31/fact-sheet-the-american-jobs-plan/>. Published March 31, 2021. Accessed April 12, 2021.
5. Barragán ND. Barragán Reintroduces Bill to Address Health Disparities by Improving the Social Determinants of Health. <https://barragan.house.gov/barragan-reintroduces-bill-to-address-health-disparities-by-improving-the-social-determinants-of-health/>. Published January 21, 2021. Accessed April 12, 2021.
6. U.S. Department of Health and Human Services OoMH. IHEC Data Workgroup. <https://www.minorityhealth.hhs.gov/omh/browse.aspx?lvl=3&lvlid=124>. Published 2021. Accessed April 12, 2021.
7. Rural Health Information Hub. Grant Funding for Programs that Address Social Determinants of Health. <https://www.ruralhealthinfo.org/toolkits/sdoh/6/grant-funding>. Published 2020-2021. Accessed April 12, 2021.
8. HealthIT.gov. Leading Edge Acceleration Projects (LEAP) in Health Information Technology (Health IT) Notice of Funding Opportunity (NOFO). <https://www.healthit.gov/topic/onc-funding-opportunities/leading-edge-acceleration-projects-leap-health-information>. Published 2021. Accessed April 12, 2021.
9. Gardner M. The Internet of Bodies Will Change Everything, for Better or Worse. In: *The RAND Blog* October 29, 2020.
10. Secretary's Advisory Committee on National Health Promotion and Disease Prevention Objectives for 2030 Report #2: *Recommendations for Developing Objectives, Setting Priorities, Identifying Data Needs, and Involving Stakeholders for Healthy People 2030*. Online: Healthy People;2018.
11. DeSalvo K, Hughes B, Basset M, et al. *Public Health COVID-19 Impact Assessment: Lessons Learned and Compelling Needs*. Washington, DC: Discussion Paper, National Academy of Medicine;2021.
12. Association of Schools and Programs of Public Health. Framing the Future. <https://www.aspph.org/teach-research/framing-the-future/>. Accessed April 12, 2021.
13. Lempert RJ, Popper SW, Bankes SC. *Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis*. Santa Monica, C.A.: RAND Corporation;2003. MR-1626-RPC.
14. Barbero C, Gilchrist S, Schooley MW, Chiqui JF, Luke DA, Elyer AA. Appraising the Evidence for Public Health Policy Components Using the Quality and Impact of Component Evidence Assessment. *Global Heart*. 2015;10(1):3-11.
15. Barbero C, Gilchrist S, Shantharam S, Fulmer E, Schooley MW. Doing More with More: How “Early” Evidence Can Inform Public Policies. *Public Adm Rev*. 2017;77(5):646-649.
16. Seidler A, Nußbaumer-Streit B, Apfelbacher C, Zeeb H, COVID-19 fdQ-ARRdKPHz. [Rapid Reviews in the Time of COVID-19 - Experiences of the Competence Network Public Health COVID-19 and Proposal for a Standardized Procedure]. *Gesundheitswesen*. 2021;83(3):173-179.
17. Nilsen P. Making sense of implementation theories, models and frameworks. *Implementation Science*. 2015;10(1):53.

18. Castillo JM. The Intersection between Systems Change, Implementation Science, and Human Beings: A Call to Investigate People and Context in Future Systems-Level Consultation Research. *Journal of Educational and Psychological Consultation*. 2020;30(4):402-411.
19. Braithwaite J, Churrua K, Long JC, Ellis LA, Herkes J. When complexity science meets implementation science: a theoretical and empirical analysis of systems change. *BMC Medicine*. 2018;16(1):63.
20. Chandra A, Miller CE, Acosta JD, Weiland S, Trujillo M, Plough A. Drivers Of Health As A Shared Value: Mindset, Expectations, Sense Of Community, And Civic Engagement. *Health Affairs*. 2016;35(11):1959-1963.
21. Brownson RC, Burke TA, Colditz GA, Samet JM. Reimagining Public Health in the Aftermath of a Pandemic. *American Journal of Public Health*. 2020;110(11):1605-1610.
22. Boyd RW, Lindo EG, Weeks LD, McLemore MR. On Racism: A New Standard For Publishing On Racial Health Inequities. In. *Health Affairs Blog* July 2, 2020.
23. Petteway RJ. Dreams Of A Beloved Public Health: Confronting White Supremacy In Our Field. In. *Health Affairs Blog*, February 9, 2021.
24. Alang S, Hardeman R, Karbeah J, et al. White Supremacy and the Core Functions of Public Health. *American Journal of Public Health*. 2021;111:815-819.

Transforming Public Health Data Systems

Supplemental Materials

This document is part of a group of supplemental materials for use by the National Commission to Transform Public Health Data Systems, an independent commission convened by the Robert Wood Johnson Foundation (RWJF). These materials were used to inform Commission discussions and recommendations. The materials and corresponding papers represent insights from literature review, a diverse set of stakeholder perspectives, and inputs from RWJF grantee projects related to the topic of the Commission. These materials were developed in support of the Commission by the RAND Corporation, a nonprofit research organization, within its division, RAND Social and Economic Well-Being. Corresponding authors are [Anita Chandra](#) and [Laurie Martin](#).

**National Commission to
Transform Public Health
Data Systems**

May 2021

Federal Government Activities Related to Equity and the Public Health Data System

This table summarizes recent and ongoing activities in the U.S. federal government related to equity and the public health data system. Some of these activities were summarized in the final, Who and What Next? paper as exemplars of the types of executive orders, working groups, and plans to support critical infrastructure offered by the Biden administration. This table, though not exhaustive, provides an expanded set of activities, and highlights the diversity of federal agencies engaged in such efforts. Such action provides strong signal value of the potential for meaningful engagement and action at the federal level to address equity in the public health data system. This table can be reviewed in combination with the full set of white papers or used separately by the Commission and Robert Wood Johnson Foundation staff.

Table 1. 2021 Federal government activities related to equity and the public health data system

Type	Department	Opportunities for Public Health Data System Transformation	Summary	Cost or Scope
Executive Orders and Federal Guidelines				
Executive Order	Office of Management and Budget; Domestic Policy Council	Executive Order On Advancing Racial Equity and Support for Underserved Communities Through the Federal Government: Equitable data working group	Order to gather data that is disaggregated by race, ethnicity, gender, disability, income, veteran status, or other key demographic variables to advance equity. ¹	Focus on federal data structures
Executive Order	Health and Human Services	Executive Order on Ensuring an Equitable Pandemic Response and Recovery	Order to gather COVID-19-related data that is disaggregated by race, ethnicity, gender, disability, income, veteran status, or other key demographic variables to ensure an equitable pandemic response. ²	Focus on COVID-19-related data
Executive Order	Cross-Department	Support for gender-based policymaking	Establishment of a Gender Policy Council to work across agencies to ensure gender-based consideration for all policies, particularly for individual and family supports and labor practices. ³	N/A
Executive Order	Labor	Support for workers, particularly workers of color	Clarification of rules around worker safety to ensure that those who cannot work for safety reasons are eligible for unemployment insurance. ⁴	Focus on all workers
Executive Order	Agriculture	Support for low-income families	Expanded allotments for lowest-income households (predominantly Black and brown families) for their SNAP benefits. ⁴	Expand support to an additional 12 million people

Type	Department	Opportunities for Public Health Data System Transformation	Summary	Cost or Scope
Executive Order	Agriculture	Support for low-income families	Support of programs to build networks of benefit delivery to reduce burden and time to receive urgent support. ⁴	N/A
Executive Order	Education	Support for students and families affected by COVID-19	Working through the Department of Education to provide students with safe education options and access to services students normally receive in school during remote learning. ⁵	N/A
Executive Order	Education	Support for students and families affected by COVID-19	Working with the Director of the Institute of Education Sciences in the Department of Education to collect data to understand the impact of the COVID-19 pandemic on students and educators, disaggregated by student demographics, language, and family income. ⁵	N/A
Executive Order	Education	Support for students and families through shared data and better policymaking	Support data collection activities related to safe reopening of schools and provide support to State, local, Tribal, and territorial leaders and the public, disaggregated by demographics, to support decision-making. ⁵	N/A
Federal Guidelines	Health and Human Services	Expand access to opioid treatment	Rules eliminate training requirement and allow wider range of health workers to offer buprenorphine treatment, which will make treatment for opioid disorder widely available. ⁶	\$125 billion in additional funding
Legislation				
Legislation	Cross-Department	American Rescue Plan	Order to expand coverage and resources for health insurance and improve access to mental health services and community-based programs that address social determinants of health. ^{7,8}	Specific to the 2021 and 2022 health insurance marketplace
Legislation	Treasury	American Rescue Plan: Support for children and early education	Expanded child tax credit that includes more money to families with children and expands access to the credit to cover more people, particularly Black and Hispanic families, and expanded funding for schools in lower income areas. ⁸	Increase the scope and money associated with each child for the tax credits
Legislation	Cross-Department	American Jobs Plan	Legislative language to improve infrastructure in rural areas, including work to deliver clean drinking water, a renewed electric grid, and high-speed broadband to all Americans. ⁹	Over \$300 billion in investments

Type	Department	Opportunities for Public Health Data System Transformation	Summary	Cost or Scope
Legislation	Labor	American Jobs Plan: Support for families and children	Support for day care and preschool in low-income areas, and support for the movement of people to the workforce that were limited by child care responsibilities. ⁹	\$200 billion
Legislation	Labor	American Jobs Plan: Support for and investment in home and community-based health workers and long-term care	Supporting the training of and protections for expanded home and community-based care workers, including supports for unionization and wage supports. ⁹	\$400 billion in range of programs and support for home and community-based infrastructure development
Legislation	Education	American Jobs Plan: Support for education	Investing in technology and school support, especially in education deserts and among community colleges that support local jobs. ⁹	\$200 billion
Legislation	Cross-Department	The Improving Social Determinants of Health Act of 2021	Legislative language to coordinate across the Centers for Disease Control and Prevention (CDC) to incorporate social determinants of health into grants and activities, award states, local, territorial, and tribal health agencies to address SDOH in target communities, award grants to nonprofit organizations and institutions of higher education to conduct research on SDOH best practices, coordinate and align SDOH activities with other agencies, and collect and analyze data related to SDOH activities. ¹⁰	Coordination among all CDC programs and \$50 million annually for program activities
Working Groups, Advisory Committees, and Regional Commissions				
Working Group	Cross-Department	Interdepartmental Health Equity Collaborative (IHEC) Data Workgroup	Establish an IHEC Data Workgroup that includes experts engaged in data activities and data policies to address SDOH and health disparities. The goals of the group are to identify existing policies and practices for improving access to data and use of data in support of policy development and to promote data projects and applications that address SDOH and support efforts to advance health equity. ¹¹	N/A

Type	Department	Opportunities for Public Health Data System Transformation	Summary	Cost or Scope
Advisory Committee	Health and Human Services	Implementation of Healthy People 2030	The Health and Human Services (HHS) Secretary's Advisory Committee on National Health Promotion and Disease Prevention Objectives for 2030 was a federal advisory committee responsible for making recommendations for developing and implementing the national health promotion and disease prevention objectives, known as Healthy People 2030; this work is ongoing and is now focused on implementation.	National, state, and local
Regional Commissions	Cross-Department	Regional commissions funding rural SDOH programs	Appalachian Regional Commission funding to promote economic development, education, health care access, and physical infrastructure in distressed Appalachian communities. ¹²	\$10 million annually in support
Regional Commissions	Cross-Department	Regional commissions funding rural SDOH programs	Delta Regional Authority funding in public and transportation infrastructure, as well as business and workforce development, in rural communities in the Mississippi River Delta region. ¹²	\$30 million annually in support
Regional Commissions	Cross-Department	Regional commissions funding rural SDOH programs	Area Health Education Centers provide funding and support to train future and current rural health professionals, addressing the educational and access needs of rural communities, including the Alaska Center for Rural Health and Health Workforce, which offers K-12 pipeline programming in health career education. ¹²	\$28 million annually in support
Funding Opportunities				
Funding Opportunity	Housing and Urban Development	Federal agencies funding rural SDOH programs	Department of Housing and Urban Development funding to support rural affordable housing and housing quality initiatives, such as the Rural Capacity Building for Community Development and Affordable Housing Program. ¹²	Each program has approximately \$15 million funded annually
Funding Opportunity	Office of the National Coordinator for Health Information Technology (ONC)	HHS funding for data interoperability for SDOH	ONC support to address collection, access, exchange, and use of SDOH data through funding opportunities such as the Leading Edge Acceleration Projects (LEAP) in Health Information Technology by advancing health IT standards and tools to improve SDOH data exchange and research. ¹³	Up to \$2 million in funding for LEAP areas of interest

Type	Department	Opportunities for Public Health Data System Transformation	Summary	Cost or Scope
Funding Opportunity	Labor	Data systems to identify gaps in pay by race	Legislation to create a new Small Business Opportunity Fund to support domestic activities, make permanent targeted tax credits, and require reporting data and trends on racial economic gaps in employment and salaries.	Approximately \$50 billion in private funding through a range of programs to expand capitol access

References

1. Biden Jr. JR. Executive Order On Advancing Racial Equity and Support for Underserved Communities Through the Federal Government. The White House. <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/20/executive-order-advancing-racial-equity-and-support-for-underserved-communities-through-the-federal-government/>. Published January 20, 2021. Accessed April 12, 2021.
2. Biden Jr. JR. Executive Order on Ensuring an Equitable Pandemic Response and Recovery. <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/21/executive-order-ensuring-an-equitable-pandemic-response-and-recovery/>. Published January 21, 2021. Accessed April 12, 2021.
3. Biden Jr. JR. Executive Order on Establishment of the White House Gender Policy Council. The White House. <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/03/08/executive-order-on-establishment-of-the-white-house-gender-policy-council/>. Published March 8, 2021. Accessed April 2, 2021.
4. The White House. Fact Sheet: President Biden's New Executive Actions Deliver Economic Relief for American Families and Businesses Amid the COVID-19 Crises. The White House. <https://www.whitehouse.gov/briefing-room/statements-releases/2021/01/22/fact-sheet-president-bidens-new-executive-actions-deliver-economic-relief-for-american-families-and-businesses-amid-the-covid-19-crises/>. Published January 22, 2021. Accessed April 4, 2021.
5. Biden Jr. JR. Executive Order on Supporting the Reopening and Continuing Operation of Schools and Early Childhood Education Providers. The White House. <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/21/executive-order-supporting-the-reopening-and-continuing-operation-of-schools-and-early-childhood-education-providers/>. Published January 21, 2021. Accessed March 23, 2021.
6. Mann B. As Opioid Deaths Surge, Biden Team Moves to Make Buprenorphine Treatment Mainstream. National Public Radio. <https://www.npr.org/2021/04/27/990997759/as-opioid-deaths-surge-biden-team-moves-to-make-buprenorphine-treatment-mainstre>. Published April 27, 2021. Accessed April 29, 2021.
7. American Public Health Association. APHA applauds passage of American Rescue Plan Act of 2021 as a step toward healing and recovery. <https://www.apha.org/News-and-Media/News-Releases/APHA-News-Releases/2021/american-rescue-plan>. Published 2021. Accessed April 12, 2021.
8. House TW. President Biden Announces American Rescue Plan. The White House. <https://www.whitehouse.gov/briefing-room/legislation/2021/01/20/president-biden-announces-american-rescue-plan/>. Published January 20, 2021. Accessed March 4, 2021.
9. The White House. Fact Sheet: The American Jobs Plan. <https://www.whitehouse.gov/briefing-room/statements-releases/2021/03/31/fact-sheet-the-american-jobs-plan/>. Published March 31, 2021. Accessed April 12, 2021.
10. Barragán ND. Barragán Reintroduces Bill to Address Health Disparities by Improving the Social Determinants of Health. <https://barragan.house.gov/barragan-reintroduces-bill-to-address-health-disparities-by-improving-the-social-determinants-of-health/>. Published January 21, 2021. Accessed April 12, 2021.
11. U.S. Department of Health and Human Services OoMH. IHEC Data Workgroup. <https://www.minorityhealth.hhs.gov/omh/browse.aspx?lvl=3&lvlid=124>. Published 2021. Accessed April 12, 2021.
12. Rural Health Information Hub. Grant Funding for Programs that Address Social Determinants of Health. <https://www.ruralhealthinfo.org/toolkits/sdoh/6/grant-funding>. Published 2020-2021. Accessed April 12, 2021.
13. HealthIT.gov. Leading Edge Acceleration Projects (LEAP) in Health Information Technology (Health IT) Notice of Funding Opportunity (NOFO). <https://www.healthit.gov/topic/onc-funding-opportunities/leading-edge-acceleration-projects-leap-health-information>. Published 2021. Accessed April 12, 2021.

Transforming Public Health Data Systems

Supplemental Materials

This document is part of a group of supplemental materials for use by the National Commission to Transform Public Health Data Systems, an independent commission convened by the Robert Wood Johnson Foundation (RWJF). These materials were used to inform Commission discussions and recommendations. The materials and corresponding papers represent insights from literature review, a diverse set of stakeholder perspectives, and inputs from RWJF grantee projects related to the topic of the Commission. These materials were developed in support of the Commission by the RAND Corporation, a nonprofit research organization, within its division, RAND Social and Economic Well-Being. Corresponding authors are [Anita Chandra](#) and [Laurie Martin](#).

**National Commission to
Transform Public Health
Data Systems**

May 2021

Glossary of Working Definitions for Terms Used in White Papers

Common-pool resource. A common-pool resource is a resource that benefits a group of people, usually by consumption. Common-pool resources (e.g., forests, fisheries) are susceptible to overuse and provide diminished benefits to everyone if each individual pursues his or her own self-interest.

Confidentiality. This term refers to the obligations of individuals or groups, which receive or use information, to respect the privacy interests of individuals, who are subjects of the data.

Data disaggregation. The collection, provision, and reporting of data by subcategories.

Data equity. The collection, provision, and reporting of demographic data that take into account the equity consequences and sociopolitical nature of the category definitions.

Deep uncertainty. Deep uncertainty refers to the level of uncertainty of the event itself or to the level of uncertainty we have about our own knowledge. Deep uncertainty relates to a situation in which events (e.g., catastrophic) cannot be addressed by simply assigning probabilities to its occurrence.

Equity. This term is defined in shorthand as the fair and just access to opportunity. Three types of equity are used in the white papers:

- *Distributive equity* focuses on allocation and resource management decisions, with attention to the balance of costs, risks, and benefits. Usually, distributive equity considers how decisions are made and benefits are distributed based on dimensions of need and social benefit.
- *Procedural equity* addresses how the concept of fairness is included in approaches and policies, once the equity parameters are set (i.e., for whom are we improving equity).
- *Contextual equity* is the backdrop of both procedural and distributive equity because this form of equity accounts for the political, economic, social, and intergenerational factors in which populations engage with society, its systems, and its benefits. This includes contextual variables such as access (e.g., access to capital) and power (e.g., the ability to gain and maintain access to resources).

Equity lens. This term is defined in these papers as any approach for analyzing the impact of actions on people or places, with attention to those historically marginalized, as well as the process of addressing any structural or systemic barriers (*adapted from multiple sources*).

Equity-oriented. In the context of the white papers, this means a public health data system that centers equity, or the fair and just access to the opportunity to be healthy, in the content and structure of data; the engagement of diverse stakeholders in the access to and use of data for decision-making with attention to issues of power, marginalization, and justice; and in how data are used to address systemic inequities.

Health equity. This term means that everyone has a fair and just opportunity to be as healthy as possible. This requires removing obstacles to health such as poverty, discrimination and its

consequences, and powerlessness as well as lack of access to good jobs with fair pay, quality education and housing, safe environments, and health care (Braveman, 2017).

Health inequities. This term refers to systematic differences in the opportunities groups have to achieve optimal health, leading to unfair and avoidable differences in health outcomes. The dimensions of social identity and location that organize or “structure” differential access to opportunities for health include race and ethnicity, gender, employment and socioeconomic status, disability and immigration status, geography, and more.

Health information exchange models. We use several terms in the papers to explain the structural types:

- *Centralized.* Data are sent to a centralized database or repository by a network of users. The data can then be accessed by participants in accordance with the policies and procedures governing that repository.
- *Federated.* Data are stored in separate data repositories or nodes, which are interconnected in a federation. The nodes have a central management framework and set of agreed upon principles and standards that creates a uniform environment so that the member networks can share data.
- *Hybrid.* Data are stored in separate data repositories or nodes, but a central repository contains a limited set of standardized data on every individual in the network along with a record locator service, which enables tracking of data across nodes.
- *Distributed.* Data are stored in a distributed network, where users maintain control of their data, but provide access to other users by enabling user privileges.

Governance. This term refers to the structures and processes by which people in societies make decisions and share power, creating the conditions for ordered rules and collective action.

- *Adaptive governance.* Adaptive governance refers to an “evolution of the rules and norms that promote the satisfaction of underlying human needs and preferences given changes in understanding, objectives, and the social, economic, and environmental context.”

Information. This term is used to refer to useable knowledge that can support decision-making, as part of the *data to information to insight* continuum.

Insight. This term is used for information that is contextualized in history and structures, and with action for decision-making. This is the last stage of the *data to information to insight* continuum.

Modern. This term is used to describe a data system that is agile, can connect multiple forms of data and translate those data into *information* (defined as useable knowledge) and *insight* (defined as information contextualized in history and structures), and is fully equipped for the health opportunities and challenges of the 21st century.

Parsimony. This term is the principle that the simplest explanation that can explain the data is to be preferred. *Data* parsimony follows in terms of being minimalist or sparing in the number of variables, indicators, or measures needed to explain a concept, cause, phenomenon, or outcome.

Participation. Participation means the active role of those impacted by public health decisions in the actual process of arraying decision options and providing meaningful input on those decisions.

Power. This term means the authority to shape expectations, decisions, and outcomes in public health.

Precise. In the context of data precision, the term used in the white papers, this means data that are exact and/or sharply defined.

Privacy. This term refers to an individual's rights to control the acquisition, uses, or disclosures of his/her/their identifiable health data.

Public health data. These are any data that can be used to understand, assess, and monitor the health and well-being of a community, population, and/or nation, with the purpose of prospectively determining health assets and needs and retrospectively capturing intergenerational trajectories of health over time.

Public health data system. This is the system or systems of actors and sectors with data and agency to make decisions to advance the health and well-being of a community, population, and/or nation, with consideration of how opportunities to be healthy are equitably distributed. The public health data system crosses federal, regional, and local boundaries.

Security. This term refers to technological or administrative safeguards or tools designed to protect identifiable health data from unwarranted access or disclosure.

Sense-making. This term is used to describe the cognitive processes by which people make meaning from data and experiences.

Structural inequities. This term refers to personal, interpersonal, institutional, and systemic drivers—such as racism, sexism, classism, able-ism, xenophobia, and homophobia—that make those identities salient to the fair distribution of health opportunities and outcomes. Policies that foster inequities at all levels (from organization to community to county, state, and nation) are critical drivers of structural inequities.

System-minded. This term is used in the papers to mean “seeing the system as a system, understanding system dynamics, and then shaping intentions, decisions, and actions in terms of this new perspective.” (*Adapted from an interview with Banny Banerjee, Stanford University ChangeLabs from article by Leonard Teichert, Medium 2019*)

Transformation-minded. This term is used in the papers to mean “shaping intentions in terms of deeply transformed system behavior, and *framing success* in terms of system level transformation, shaping the new types of *agency and action* that will actually result in the system veering towards altered states.” (*Adapted from an interview with Banny Banerjee, Stanford University ChangeLabs from article by Leonard Teichert, Medium 2019*)

Voice. Voice means the inclusion of the perspectives, ideas, and lived experiences of those impacted by public health decisions.