Abstract
The growing centering of equity in health has elevated a conversation about how those interests should translate within the systems and sectors that influence health. In particular, the public health data system has been relatively limited in capturing the drivers and consequences of health inequity as well as the varying dimensions of equity. This article examines what it means to use equity as a guiding principle throughout the components and functions of a modern public health data system. As with other articles in this supplement, this article builds from a literature review, environmental scan, and deliberations from the National Commission to Transform Public Health Data Systems to summarize current gaps to integrate equity throughout the system. It outlines opportunities for the technology and data science sectors specifically to engage given the access that these sectors have to information that would illuminate and frame the nuances and impacts of health inequity.

Keywords: equity; public health; data

Introduction
With the amplification that COVID-19 provided in illuminating health inequities, the issue of equity has come to the fore in public health discussions, including in analyses of how to transform the public health system. One of the critical challenges has been the ability of data within the public health system to not only highlight inequities but also help in the anticipation and ability to address drivers of those inequities. Given the primacy of health equity and the devastating consequences of inequity, it is argued that equity and equity orientation should not be a feature but must be the primary purpose of the modern public health data system going forward.

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 powerlessess and lack of access to good jobs with fair pay, quality education and housing, safe environments, and health care. 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. Further, with all the information we could have about individuals and communities and their health and well-being given advancements in data content and science, it begs questions about how this data modernization can embrace equity principles from the start.

In this article, we briefly set the stage for the equity orientation of the modern 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. We describe what has been missing from the equity ambitions of the current data system and what it will take to develop a more equity-centered data system. The equity challenge can be viewed as an opportunity for leaders in data science and technology, who hold critical information about health inequities and drivers of inequities. This could be brought to bear for the move toward an equity-centered, modern public health data system.

Methods
In 2020, the Robert Wood Johnson Foundation formed the National Commission to Transform Public Health Data Systems to review significant challenges to the current public health data system, and "provide recommendations to policymakers, health care organizations and institutions, service providers, and philanthropy" on the potential solutions to overcome these challenges.4

In support of this effort, RAND conducted a supporting analysis that included an environmental scan to identify key issues, points of consideration, tradeoffs and tensions, and current activities related to public health data, data systems, and data modernization efforts. This effort included a targeted scan of published research papers and reports, reviews of websites and working documents describing coordinated activities (e.g., data interoperability), and recent initiatives. Additional searches included the use of "big data" in public health, data privacy, and ethics of public health data collection. Although the team primarily focused on public health data, it also identified seminal articles and reports from other sectors or disciplines whose findings could apply to public health data systems.

RAND simultaneously conducted semi-structured interviews with 112 experts and thought leaders on the main topics before the Commission. Individuals represented diverse sectors, including public health and health care, technology and data science, research and policy, journalism, and law. The interviews also included experts in data, data use, equity, community engagement, and research translation who work outside the traditional health sector. The project was reviewed and approved by the RAND Human Subjects Protection Committee.

In this article, we highlight relevant findings from this supporting analysis and then implications for the data science and technology sectors, with a consideration of recommendations that emerged from the final Commission report.

Findings
When applying an equity lens 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. For instance, the goal of Healthy People 2030 is 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 considering COVID-19.5−7 Although the Healthy People objectives and others like them are laudable, there is a challenge related to equity. Is health equity truly a throughline of a national scorecard like this, such that equity is understood as both process and outcome? Are the components of equity, such as procedural and distributive equity, assessed and to what extent are drivers of inequity, such as history, cumulative stresses, and structures, monitored as part of the public health data system?

In short, insight from stakeholders and the current landscape revealed that the modern public health data system must be able to capture all forms of equity, and yet currently falls short.2

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.2

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 (vs. 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;

---

1 An equity lens is defined 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).
• Distributive equity: how social welfare and need is balanced; and
• Contextual equity: how pre-existing social conditions influence equity.

We talk about each type of equity and the current gaps in the following sections. We then turn to the implications for those in the data science and technology sector.

A lack of procedural equity has impeded 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. 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 other emergent threat presents.

This type of procedural equity steps beyond trust to include 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.

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, we reference the inclusion of the perspectives, ideas, and lived experiences of those impacted by public health decisions. By participation, 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 tradeoffs of resources and values, and values are often driven by ethics and social and cultural experiences. 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.

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. Ideally, with more attention to procedural equity, data or digital colonialism, a concern that the private sector is harvesting data without reciprocal public benefit, could be less of an issue.

Without distributive equity represented in the public health data system, it is difficult for the data system to fully support its sense-making (or the cognitive processes by which people make meaning from data and experiences) 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 benefits are distributed based on the 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 (e.g., insurance status, quality of health services, prenatal care, discrimination) of a particular health outcome (e.g., maternal mortality) with the health outcome so that decision makers are considering the investments needed in those drivers together to influence an outcome, and what is then possible...
given resource constraints.\textsuperscript{16–19} A transformed public health data system could offer information and, ultimately, insights, which help with tradeoffs for realizing concepts, such as targeted universalism and efforts to address historical inequities.

Contextual equity regarding pre-existing social conditions and its influence on equity is a significant gap

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).\textsuperscript{8}

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.

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.\textsuperscript{2}

Implications

During the stakeholder analysis, review of existing data systems and dashboards, and Commission deliberations, it not only became clear that the current public health data system falls short in representing the comprehensive dimensions of equity, but also that many sectors must be part of realizing the equity-centered expectations of the next-generation public health data system.\textsuperscript{2,4} In short, this cannot be solved through a governmental public health lens alone.

Capturing the complexities of health inequity—whether the intersectionalities of health impacts by overlapping population characteristics or the ability to contextualize why those health inequities have emerged—will require new forms of data and new ways to merge data, visualize those data, and tell stories from the data to drive data sense-making and decision making. Specifically, technology companies, with access to information about community context and historical patterns, and data scientists, with analytic tools to connect seemingly disparate, cross-sector drivers of health inequity, are needed to advance procedural, distributive, and contextual equity.

There are a few areas that are ripe for stronger data science and tech sector engagement. For instance, one of the Commission’s recommendations was squarely focused on equity by asking technology companies to support public health data system transformation in under-resourced areas of the country with the largest health inequities, by either direct financial support or indirect collaborations.\textsuperscript{4} Further, the Commission argued that public health innovation could be fostered by promoting methods\textsuperscript{4}; some of those methods are key to understanding equity.

We highlight two major implications for this sector here.

Data science and technology companies can help to augment transparency and power in data access, use, and decision making

Equity requires participation and justice, and there have been challenges to those themes when considering topics such as “big data” and algorithmic bias. As noted in the role of procedural equity, the issue of representation in health decision making must be a key part of a modern, equity-oriented public health data system, but there remain concerns about the extent of participation, accountability, and power\textsuperscript{19} (or the authority to shape expectations, decisions, and outcomes in public health).

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.\textsuperscript{2} 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 (characteristics of “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. Kalluri wrote, “Don’t ask if AI is good or fair, ask it if it shifts power.”\textsuperscript{20}
This sentiment targeting AI was focused on ensuring that the tool does not exacerbate inequities through problems such as algorithmic bias, but this theme can also be widened and attributed to all 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. 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. 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 whether the benefits of these “big data” accrue to corporations while the burden is 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.

Data science and technology companies can expand their partnership with public health agencies and other public health leaders in informing the public health and equity narrative.

In addition to this issue of what is collected and analyzed and how those most impacted are included in data sense-making, there is the topic of data story. Communicating the intricacies of equity and its role in health outcomes is not an easy task. The ability to explain the impacts of systemic and structural racism requires analytic storytelling, and other tools that quite simply are not often in the resources or realm of governmental public health. Given this, there are two clear roles for the data science and technology sector going forward.

One is to provide supporting platforms and bridges to other public health actors in communities to aid in explaining the historical links of racism and other inequities to health outcomes today, in ways that ensure the public can both understand the information and apply it to redress distributive equity issues. The illumination of disparities only (i.e., just describing gaps between groups) has left public health falling short, but information that explains the drivers of health inequity can aid local response by tying those inequities to policies and other paths for intervention.

The second area for the data science and tech sector is the sustained engagement and training of the public health workforce in how to center equity in the design of data platforms, such that these stories of historical and current inequity can be explained. This means revisiting how data dashboards are organized to allow for contextual equity considerations, and how those in data science and technology learn to communicate those findings in ways that can lead to social change and structural action.

Conclusion

Although the task of public health data transformation is daunting and requires a level of cross-sector collaboration often not observed, the opportunities for the data science and tech sector to engage in positive ways to address health equity are clear. Moving ahead, more operational discussions about equity-centered data and representation will require new terms of collaboration and data governance that elevates public and private sector cooperation to a new level of engagement.

Author Disclosure Statement

No competing financial interests exist.

Funding Information

This article was supported under a grant from the Robert Wood Johnson Foundation. The views expressed are solely the authors.

References

1. Krieger N. The Commonwealth Fund. COVID-19, Data, and Health Justice. New York, NY; 2020. Available from: https://www.commonwealthfund.org/blog/2020/covid-19-data-and-health-justice [Last accessed: 2021].
2. Chandra A, Martin L, Acosta J, et al. Transforming Public Health Data Systems, Why? The Focus on Equity in the Modern Public Health Data System. Robert Wood Johnson Foundation, National Commission to Transform Public Health Data Systems: Princeton, NJ; 2021.
3. Braveman P, Arkin E, Orleans T, et al. What Is Health Equity? Robert Wood Johnson Foundation: Princeton, NJ; 2017.
4. National Commission to Transform Public Health Data Systems. Charting a Course for an Equity-Centered Data System: Recommendations from the National Commission to Transform Public Health Data Systems. Robert Wood Johnson Foundation: Princeton, NJ; 2021. Available from: https://www.rwjf.org/en/library/research/2021/10/charting-a-course-for-an-equity-centered-data-system.html [Last accessed: April 1, 2022].
5. Centers for Disease Control and Prevention. Paving the Road to Health Equity Online. Atlanta, GA; 2020. Available from: https://www.cdc.gov/minorityhealth/publications/health_equity/index.html [Last accessed: 2022].
6. Dover DC, Belon AP. The health equity measurement framework: A comprehensive model to measure social inequities in health. Int J Equity Health 2019;18(1):36; doi: 10.1186/s12939-019-0935-0.

7. Institute for Healthcare Improvement. Health Leads. A Measurement Framework for Health Equity. Boston, MA; 2019. Available from: https://healthleadsusa.org/resources/a-measurement-framework-for-health-equity/ [Last accessed: 2021].

8. McDermott M, Mahanty S, Schreckenberg K. Examining equity: a multidimensional framework for assessing equity in payments for ecosystem services. Environ Sci Policy 2013;33(November 2013):416–427; doi: 10.1016/j.envsci.2012.10.006.

9. Shore DA. Communicating in times of uncertainty: The need for trust. J Health Commun 2003;8(S1):13–14; doi: 10.1080/713851977.

10. Frewer L. Risk perception, social trust, and public participation in strategic decision making: Implications for emerging technologies. Ambio 1999; 28(6):569–574. Corpus ID: 89053581.

11. Casebeer A, Mackean G, Abelson J, et al. Engaging the public voice in health care decision-making. In: Organizing and Reorganizing. (McKee L, Ferlie E, Hyde P. eds.) Springer: London; 2008, pp. 123–138.

12. Church J, Saunders D, Wanke M, et al. Citizen participation in health decision-making: Past experience and future prospects. J Public Health Policy 2002;23(1):12–32. PMID: 12013713.

13. Wiggins A, Willbanks J. The rise of citizen science in health and biomedical research. Am J Bioeth 2019;19(8):3–14; doi: 10.1080/15265161.2019.1619859.

14. Boyd RW, Lindo EG, Weeks LD, et al. Health Affairs. On Racism: A New Standard For Publishing On Racial Health Inequities. Washington, DC; 2020. Available from: https://www.healthaffairs.org/doi/10.1377/forefront.20200630.939347/ [Last accessed: 2020].

15. Couldry N, Mejias UA. Data colonialism: Rethinking big data’s relation to the contemporary subject. Televis New Med 2019;20(4):336–349; doi: 10.1177/1527476418796632.

16. Davis R. Measuring What Works to Achieve Health Equity: Metrics for the Determinants of Health. Prevention Institute: Oakland, CA; 2015.

17. The HOPE Initiative. Measures to Advance Health and Opportunity Online. The HOPE Initiative. Available from: https://www.hopeinitiative.org/ [Last accessed: 2021].

18. Ausura C. Rhode Island Department of Health. Rhode Island Health Equity Measures. Providence, RI; 2021. Available from: https://health.ri.gov/data/healthequity/ [Last accessed: 2021].

19. Gore R, Parker R. Analysing power and politics in health policies and systems. Glob Public Health 2019;14(4):481–488; doi: 10.1080/17441692.2019.1575446.

20. Kalluri P. Don’t ask if artificial intelligence is good or fair, ask how it shifts power. Nature 2020;583(169); doi: 10.1038/d41586-020-02003-2.

21. Vayena E, Madoff L. Navigating the ethics of big data in public health. In: The Oxford Handbook of Public Health Ethics. (Mastroianni AC, Kahn JP, Kass NE. eds.) Oxford University Press: New York, NY; 2019; pp. 354–368.

22. Mittelstadt BD, Floridi L. The ethics of big data: Current and foreseeable issues in biomedical contexts. Sci Eng Ethics 2015;22(2):303–314; doi: 10.1007/s11948-015-9652-2.

Cite this article as: Chandra A, Martin LT, Acosta JD, Nelson C, Yeung D, Qureshi N, Blagg T (2022) Equity as a guiding principle for the public health data system. Big Data 10:S1, S3–S8, DOI: 10.1089/big.2022.0204.