

Activating Health Data to Achieve Health Equity:

Recommended Strategies for REaL and SOGI Data

Shruti Kothari,

Director, Industry Initiatives,
Blue Shield of California

and

Lynne Chou O'Keefe,

Founder and Managing Partner,
Define Ventures

Understanding population health and realizing the goal of health equity require the collection, sharing, and use of comprehensive data – including race, ethnicity, and language (“REaL”) and sexual orientation and gender identity (“SOGI”) data.^{1, 2} Yet entities are not able to do these things consistently today, as we learned when the COVID-19 pandemic exposed deficiencies in our data systems. The failures were especially evident in our systems for collecting and sharing REaL and SOGI data.^{3, 4} Overall, we suffer from a lack of transparency as well as inconsistent or missing data, which compound historical injustices in our health systems.

If we are to address the barriers to care, we also must work together to reduce the barriers to REaL and SOGI data collection, use, and sharing. Collecting and analyzing these data helps health care stakeholders, government, research institutions, and business partners identify areas for action in policy, program, and product development. Sharing and using data can support patients individually at the point of care, and at the aggregate level, provide policymakers insights into population-level health needs. Finally, rich and comprehensive data can help reduce care disparities when tied to quality measures.

Blue Shield of California Industry Initiatives recently hosted events for stakeholders to foster better collaboration and understanding of their needs in developing strategies and solutions to the collection, use, and sharing of REaL and SOGI data. The first, publicly

accessible event hosted on March 29, 2022, featured three lightning rounds from experts in the field who have been working to collect, share and use demographic data and featured an expert panel of health care funders in venture capital, philanthropy, health plans and government. The second event, hosted on April 6, 2022, featured small break out discussions with 37 individuals whose roles and organizations vary from health information technology, payers, providers, and government.

In this paper, we explore the recommendations highlighted by stakeholders during these events. These strategies chart a path forward to how health care, government, research institutions, and business partners can each advance the collection, use, and sharing of REaL and SOGI data to help achieve health equity.

Comprehensive and granular REaL and SOGI data are key to meaningfully tackling health disparities



Approaches to collecting, using, and sharing data need to be a priority to achieve health equity.⁵ As part of these approaches, REaL and SOGI data represent characteristics that may reflect some of the key drivers of health, because researchers, governments, and other organizations rely on them to assess health disparities by population.⁶ Demographic data's role in addressing health disparities and the need to collect granular demographic data have been well-publicized for over a decade, identified in the Institute of Medicine's 2009 report on the subject.⁷

Some states are working to set standards within public programs related to SOGI data, such as Oregon's state Medicaid program.⁸ In many states, however, there is no requirement for health care organizations like providers and health plans to collect and share these data.^{9, 10} At the same time, although some organizations may have these data, they are collecting the data in different ways. That makes it challenging to share this vital information across organizations and do so in ways that ensure that analytic results accurately reflect characteristics of the population.

Governmental and standard-setting organizations have worked to create standards for REaL and SOGI data so that they can be captured by health care providers and shared in a structured manner. Both the Office of Management and Budget ("OMB")¹¹ and the

Centers for Disease Control and Prevention ("CDC")¹² have developed standards for race and ethnicity. The Office of the National Coordinator for Health Information Technology ("ONC") has used these data standards in versions two and three of the United States Core Data for Interoperability ("USCDI"), with which all certified electronic health technologies will likely have to comply.¹³ Similarly, the USCDI version 3 utilizes gender standards created by national standard setting organizations like SNOMED CT and HL7 Gender Harmony Project.^{14, 15}

Work is needed to ensure that standards capture the entire breadth of identities and that providers have the infrastructure and support needed to accurately capture these data. Part of that work must address the fact that widespread consensus in terminology can be difficult to obtain about race and ethnicity categories (as reported by Americans in polls).¹⁶ A similar issue exists in the context of SOGI data, given the fluidity and expansion of gender and sexual identities.^{17, 18}

Recommendations

The stakeholder discussions resulted in four major recommendations for government agencies, research institutions, and business leaders to incorporate into their ongoing work related to collecting, sharing, and using REaL and SOGI data in health care.

1

Center individuals and their communities in the development, implementation, and ongoing improvement of programs and products.

Rationale:

Building on the “nothing about us, without us” concept,¹⁹ collection, sharing, and use of REaL and SOGI data should be centered first around the individuals who are being asked to share these data and second around the communities they represent. An ethical approach to this work requires a continuous focus on the wellbeing of the individual as well as a core belief that the value of these data in terms of driving improvements is defined by the individuals whose data it is. Centering individuals and promoting the concept of “authenticity of voice” can also serve the business interest of meeting consumer needs and preferences.

Individuals being asked to share data should be core members of the team from a project’s inception, so that their input is obtained on all aspects of the plan for data collection. This can take various forms, from hosting focus groups to establishing formal community advisory groups. Participation can include shaping how individual consent is obtained, determining what data will be collected in what formats, how data will be disaggregated, how data will be secured, and permitted data uses. These individuals can provide feedback on the degree to which the information is likely to benefit the communities being asked to share their data. To support equitable consumer participation, organizations should explore forms of compensation that acknowledge the time incurred by participants.

The importance of self-identification and reporting should also be noted. Individuals are the people best poised to provide REaL and SOGI data about themselves versus imputed data, or provider observations. This approach aligns with federal policy on the subject, provided by OMB in 1997, similarly stating that self-reporting or identification for race and ethnicity data is the preferred method of data collection.²⁰ This approach might also consider how to accommodate individuals that decline to respond to mandated requests to self-report data.

In addition, the individuals who share REaL and SOGI data should have the opportunity to review and confirm

their own data, as well as receive summaries of data they contribute and how their information is being used on a population-wide basis. Ongoing representation of diverse perspectives should be included throughout the process, as programs or products based on these data are designed, implemented, and refined.

Examples of how leaders can put this into practice:

Government agencies. Government agencies can include equal consumer input in the development of policy about data sharing, including in governance models. For example, in California, the California Health and Human Services Agency’s proposed Data Exchange Framework suggests adopting a governance model like Covered California, which includes consumer advocacy organizations on its board and allows for consumers to participate in public comment.²¹

Research institutions. Research institutions can create and/or further engage groups like patient and family advisory boards or establish new groups for input on REaL and SOGI data strategies, and they can consider requiring more patient and family engagement as part of the institutional review board (IRB) process. Research institutions can also recommend that researchers adopt and adapt community-based participatory research models featuring community-engaged data collection and data management.²²

Business leaders. Health care organizations can continue to develop new interventions supporting the unique needs of specific communities—including those that involve the collection or sharing of individual data. For example, startups can develop solutions targeted to specific populations, like telehealth programs targeted at the LGBTQ+ community. Additionally, organizations like health plans and hospitals can work to develop patient and member advisory boards or other methods of patient representation, like those developed by research institutions, to advise on program and product development.

2

Deploy trusted messengers in community engagement and data collection efforts.



Rationale:

Trust between an individual and an organization takes time to build and often results from an ongoing, meaningful relationship. Individuals who are asked to share their REaL and SOGI data are likely to be more responsive when those requests come from trusted individuals or organizations with whom they have a closer relationship. Requests from a trusted health care provider or nonprofit organization providing services in the community are likely to result in more open responses than the same request from organizations such as state or federal government agencies, or health plans, or, for another example, from staff at a mass vaccination site.

Trusted entities are essential to assist with developing appropriate content, including how questions are asked, the timing, who is asking, and how the request is communicated. For individuals to feel comfortable sharing their REaL and SOGI data, the words associated with the request also will have to be compelling to the individual and easy to understand. The reasoning for the data collection, sharing, and use also must be clearly explained. Organizations should ensure this content reflects perspectives shared by others who have lived and shared experiences with the individuals being asked for data, and that these individuals receive training in this work to ensure cultural sensitivity and prevent duplicative asks for demographic data. This can include ensuring trusted messengers are involved in program or product design (as mentioned

in Recommendation 1) and ensuring that communications are provided in the consumer's preferred language.

Examples of how leaders can put this into practice:

Government agencies. In developing public programs, government agencies can structure them to pay for trusted messengers such as community health workers to provide services, including the collection of REaL and SOGI data. Other state Medicaid programs, for example, can follow the lead of California's Medicaid program, Medi-Cal, to develop mechanisms to cover services by community health workers as a benefit, and agencies can provide funds directly to community-based organizations to assist with critical tasks like data collection. This is similar to funds deployed to these organizations to assist with vaccine administration during the pandemic. Government agencies should also consider providing funding and technical assistance to providers on culturally relevant data collection practices.

Business leaders. Health care organizations can look to employ or contract with organizations that employ trusted community providers who can dually serve as data collectors and data stewards. Health plans, for example, can expand care coordination teams that help to manage and coordinate member needs, and in developing rapport with members, serve as a data collector for the health plan.

3

Develop data infrastructure that can serve consumer, provider, population health, and other needs.



Rationale:

The infrastructure to collect and manage these data is an essential and substantial investment for both the public and private sectors. Yet there are significant challenges to designing and creating such an infrastructure to effectively meet a variety of users' needs. Geography, type of organization, ability to meet regulatory requirements such as in the Meaningful Use program, and available resources must all be considered in designing a data infrastructure that reduces inequities. Implementation must consider the ability for individuals who contribute data to access and change their data, an important aspect of data stewardship.

Providers often require technical and financial assistance to adopt new technology to ensure that the organization has the human resources sufficient for operations, and to educate providers on how to use the technology. Indeed, cost is a major factor in data infrastructure lagging for providers serving marginalized populations. Health plans and payers also require assistance to understand both how to better collect member data and how to use it for program development.

Examples of how leaders can put this into practice:

Government agencies. Government agencies can develop incentive programs and infrastructure programs for providers to adopt information technology capabilities and develop workforce capacity to better collect REaL and SOGI data. This is especially true for small practices and those that serve marginalized communities. This could take the form of a program like the former Meaningful Use program that created incentives for electronic health record adoption and the capture of clinical data,²³ and other American Recovery and Reinvestment Act funding that provided substantial federal matching dollars for state investments in health information exchanges.²⁴

Business leaders. Health care organizations can also make investments in community partners' acquisition of information technology and human resources. For example, health plans can provide grants to providers to cover the cost of infrastructure, software, and staff training to adopt new technology, similar to grants that health plans like Health Net made during the wake of the COVID-19 pandemic.²⁵ These grants can be used specifically to help providers adopt technology, and to establish new workflows and practices that allow them to better capture SOGI and REaL data.

4

Collect, analyze, and share back data in ways that are meaningful to communities and individuals offering data.

Rationale:

Organizations can further meet individual and community needs by evaluating interventions such as programs or products. This can help organizations stay true to the adage “we can’t improve what we don’t measure.” Equally important is sharing aggregated data back with the communities who contributed it and showing the value it has to the person who shared the data and to the community. Trusted messengers and community organizations can be employed to share this information with individuals and communities and solicit their input on program and product evaluation and improvement.

Incorporating this strategy into an iterative program or product process not only helps to reinforce trust but also serves to help organizations further tailor programs and products to consumer needs. Incorporating actions like aggregation of data by race, ethnicity, and language, for example, can help tailor interventions to specific groups’ needs, as well as improve individual and population health outcomes. REaL and SOGI data should also be connected to the underlying drivers of health such as food, housing and transportation, and into alternative payment models to ensure that payers and providers are paid in ways that incentivize addressing health disparities.

The overall goals of improving individual, community, and population health and achieving health equity using REaL and SOGI data can only be achieved if the entities collecting, using, and sharing data work in close partnership with the individuals sharing their data and with trusted messengers and community organizations. These entities themselves should work together to ensure data is collected and shared, to avoid the propagation of multiple, duplicative outreach campaigns targeting the same population(s).

Examples of how leaders can put this into practice:

Government agencies. Government agencies funding health projects can both require the collection and

reporting of REaL and SOGI data and require grantees to report these data (and evaluation and research publications) back to the public. For example, the government can replicate data platforms like ClinicalTrials.gov and similarly require government grantees to report data and publications back on these platforms, as do the National Institutes of Health and the Food and Drug Administration. Government agencies can also require health plans and providers to collect and report on REaL and SOGI data. For example, California Senate Bill 1033 (2022) would require commercial health plans in the state to collect standardized SOGI and REaL data, among other data, and report these data to the state’s insurance regulatory agencies. Government agencies can also consider how they can regulate the use of REaL and SOGI data to prevent harms and ensure that these data are not used in discriminatory or other adverse ways. These agencies, like the California Health and Human Services Agency, can convene stakeholder advisory groups to determine how best to prevent harms while fostering data collection practices.

Research institutions. Research institutions can develop means for sharing data back with communities and populations with whom they conduct research. For example, researchers can utilize approaches like Data Walks, where researchers actively share data and research findings back with communities.²⁶ Researchers can help to translate what findings mean and collaborate with affected communities to then determine next steps for program or policy improvement.

Business leaders. Business leaders can develop user-friendly solutions to share data back with individuals. As an example, health care providers can develop patient portals that not only meet regulatory requirements to share clinical data back with patients but also involve patient advisors to ensure that portals are user-friendly and continuously improve by incorporating consumer feedback. These portals can offer patients options to review and correct their data, including their REaL- and SOGI-related data.

Conclusion and Call to Action

Government agencies, health systems, research institutions, and business leaders all play important roles in efforts to collect, share, and use REaL and SOGI data to achieve health equity. Critical to these efforts is the inclusion of and close partnerships with the individuals who are being asked to share their data as well as organizations they trust. Ongoing evaluations of and revisions to policies, programs, and products are needed to ensure they are meeting the needs of the individuals, communities, and populations they are designed to serve. Achieving health equity demands that policy and program planners have the data needed to meet the needs of our historically marginalized communities, and that those communities are participants in shaping our data infrastructure.

Acknowledgments

We would like to thank all the partners who provided input on this action plan. The contributors represent a diverse group of healthcare organizations, including but not limited to providers, community health organizations, government, academia, health technology companies and health plans. In particular, we would like to thank the following individuals:

- **Alethea Varra**, Vice President, Clinical Care, Lyra Health
- **Alice Hm Chen**, Chief Medical Officer, Covered California
- **Andy Slavitt**, Former Senior Advisor to the Biden Administration COVID Response
- **Cary Sanders**, Senior Policy Director, California Pan-Ethnic Health Network
- **Dan Chavez**, Executive Director, Santa Cruz Health Information Organization
- **Daniel Stein**, President, Stewards of Change Institute
- **Gabriella Barbosa**, Managing Director of Policy, The Children's Partnership
- **Halima Ahmadi-Montecalvo**, Senior Director, Research and Evaluation, Unite Us
- **Jaime Bland**, Chief Executive Officer, CyncHealth
- **Jessica Kahn**, Partner, McKinsey & Company
- **Julia Skapik**, Chief Medical Information Officer, National Association of Community Health Centers
- **Katie Heidorn**, Executive Director, Insure the Uninsured Project
- **Kia Wilson**, Community Health Advocate, Blue Shield of California
- **Kristi Ebong**, Head of Partnerships and Market Development, Define Ventures
- **Kymie Thomas**, Program Specialist, Community Outreach and Patient Empowerment
- **Lisa Bari**, Chief Executive Officer, Civitas Networks for Health
- **Mark Savage**, Managing Director, Digital Health Strategy & Policy, Savage & Savage LLC
- **Naomi Senkeeto**, Managing Director of Health Equity Policy, Blue Cross Blue Shield Association
- **Naveen Rao**, Founder & Managing Partner, Patchwise Labs
- **Nelly Ganesan**, Executive Director, Health Equity, Morgan Health
- **Peter Robertson**, Director of Insights, Analytics & Data Management, Purchaser Business Group on Health
- **Robby Knight**, Co-Founder & Chief Executive Officer, Soda Health
- **Rocco Perla**, Co-Founder, The Health Initiative
- **Sandra Hernández**, President and Chief Executive Officer, California Health Care Foundation
- **Shireen McSpadden**, Director of Homelessness and Supportive Services, City and County of San Francisco

Endnotes

- ¹ National Research Council (US) Panel on DHHS Collection of Race and Ethnic Data; Ver Ploeg M, Perrin E, editors. Eliminating Health Disparities: Measurement and Data Needs. Washington (DC): National Academies Press (US); 2004. Appendix D, The Role of Racial and Ethnic Data Collection in Eliminating Disparities in Health Care. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK215740/>.
- ² Streed Jr CG, Grasso C, Reisner SL, Mayer KH. Sexual Orientation and Gender Identity Data Collection: Clinical and Public Health Importance. *American Journal of Public Health* 2020;110(7):991-993. Available from: <https://doi.org/10.2105/AJPH.2020.305722>.
- ³ James CV, Lyons B, Saynisch PA, Hudson Scholle S. Modernizing Race and Ethnicity Data in Our Federal Health Programs. *The Commonwealth Fund Blog*. 2021. Available at: <https://www.commonwealthfund.org/blog/2021/modernizing-race-and-ethnicity-data-our-federal-health-programs>.
- ⁴ Johnson A. Race and ethnicity data missing for nearly half of coronavirus vaccine recipients, federal study finds. *Washington Post*. February 1, 2021. Available at: <https://www.washingtonpost.com/health/2021/02/01/covid-vaccine-race-ethnicity-data/>.
- ⁵ McAvey KC, Reginal A. *Unlocking Race and Ethnicity Data to Promote Health Equity in California: Proposals for State Action*; 2021. Available at <https://www.manatt.com/insights/white-papers/2021/unlocking-race-and-ethnicity-data-to-promote-health>.
- ⁶ For example, see Lee WC, Veeranki SP, Serag H, Eschbach K, Smith KD. Improving the Collection of Race, Ethnicity, and Language Data to Reduce Healthcare Disparities: A Case Study from an Academic Medical Center. *Perspect Health Inf Manag*. 2016;13(Fall):1g. Published 2016 Oct 1. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5075235/>. See also Health Disparities by Race and Ethnicity in California: Pattern of Inequity, *California Health Care Foundation*. 2021. Available at <https://www.chcf.org/publication/2021-edition-health-disparities-race-ethnicity-california/>.
- ⁷ IOM (Institute of Medicine). 2009. Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement. Washington, DC: The National Academies Press.
- ⁸ Zylla E, Lukanen E. Collection of Sexual Orientation and Gender Identity (SOGI) Data: Considerations for Medicaid and Spotlight on Oregon. *State Health and Value Strategies*. 2021. Available at: https://www.shvs.org/wp-content/uploads/2021/10/SOGI-Data-Collection-in-Medicaid_SHVS-Issue-Brief_Revised.pdf.
- ⁹ Federal Action is Needed to Improve Race and Ethnicity Data in Health Programs. Grantmakers in Health and National Committee for Quality Assurance. 2021. Available at: <https://www.gih.org/wp-content/uploads/2021/10/GIH-Commonwealth-Fund-federal-data-report-part-1.pdf>.
- ¹⁰ LGBTQ-Inclusive Data Collection: A Lifesaving Imperative. Human Rights Campaign Foundation. 2019. Available at: <https://assets2.hrc.org/files/assets/resources/HRC-LGBTQ-DataCollection-Report.pdf>.
- ¹¹ 62 Fed. Reg. 210 at 58782 (October 30, 1997). Available at <https://www.govinfo.gov/content/pkg/FR-1997-10-30/pdf/97-28653.pdf>.
- ¹² Centers for Disease Control and Prevention. Race and Ethnicity Code Set Version 1.0. 2000. Available at <https://www.cdc.gov/phn/resources/vocabulary/documents/cdc-race-ethnicity-background-and-purpose.pdf>.
- ¹³ U.S. Dep't of Health and Human Services, Office of the Nat'l Coordinator for Health Information Technology. United States Core Data for Interoperability, Version 2. 2021. Available at <https://www.healthit.gov/isa/sites/isa/files/2021-07/USCDI-Version-2-July-2021-Final.pdf>.
- ¹⁴ Ibid.
- ¹⁵ McClure R et al. Gender harmony: improving standards to support affirmative care of gender-marginalized people through inclusive gender and sex representation. *J Amer Med Inform Assoc*. 2022;29(2): 354-63. Available at <https://academic.oup.com/jamia/article/29/2/354/6382238?login=true>.
- ¹⁶ Cohn D, Brown A, Lopez MH. Only about half of Americans say census questions reflect their identity very well. Pew Research Center. 2021. Available at: <https://www.pewresearch.org/social-trends/2021/05/14/only-about-half-of-americans-say-census-questions-reflect-their-identity-very-well/>.
- ¹⁷ Garrett-Walker JJ, Montagno MJ. Queering labels: Expanding identity categories in LGBTQ+ research and clinical practice. *Journal of LGBT Youth*. March 11, 2021. Available at: <https://www.tandfonline.com/doi/abs/10.1080/19361653.2021.1896411>.
- ¹⁸ Kronk CA et al. Transgender data collection in the electronic health record: Current concepts and issues. *J Amer Med Inform Assoc*. 2022;29(2):271-84. Available at <https://academic.oup.com/jamia/article/29/2/271/6364772>.
- ¹⁹ Charlton, J. I. (1998). *Nothing About Us Without Us: Disability Oppression and Empowerment* (1st ed.). University of California Press. Available at <http://www.jstor.org/stable/10.1525/j.ctt1pnqn9>.
- ²⁰ 62 Fed. Reg. 210 at 58782 (October 30, 1997). Available at <https://www.govinfo.gov/content/pkg/FR-1997-10-30/pdf/97-28653.pdf>.
- ²¹ California Health and Human Services Agency, Center for Data Insights and Innovation. Data Exchange Framework Stakeholder Advisory Group Meeting #7. April 7, 2022. Available at <https://www.chhs.ca.gov/wp-content/uploads/2022/04/Data-Exchange-Framework-Meeting-7-%E2%80%9393-Presentation.pdf>.
- ²² Cene CW et al. Data Collection and Management in Community Engaged Research: Lessons Learned From Two Community-Based Participatory Research Partnerships. *Prog Community Health Partnersh*. 2015;9(3):413-22. Doi: 10.1353/cpr.2015.0057.
- ²³ Centers for Medicare & Medicaid Services. Medicare & Medicaid EHR Incentive Program. 2010. Available at https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/mu_stage1_reqoverview.pdf.
- ²⁴ Centers for Medicare & Medicaid Services. State Medicaid Director Letter #11-004. May 18, 2011. Available at <https://www.medicaid.gov/sites/default/files/Federal-Policy-Guidance/downloads/SMD11004.pdf>.
- ²⁵ Health Net. California Medi-Cal Providers Receive Grants from Health Net Totaling \$13.4 Million to Expand Telehealth Services. June 2, 2020. Available at https://www.healthnet.com/content/healthnet/en_us/news-center/news-releases/2020-06-02_medi-cal_providers_receive_grants_from_health_net.html.
- ²⁶ Murray B, Falkenburger E, Saxena P. Data Walks: An Innovative Way to Share Data with Communities. Urban Institute. 2015. Available at <https://www.urban.org/sites/default/files/publication/72906/2000510-Data-Walks-An-Innovative-Way-to-Share-Data-with-Communities.pdf>.