

Best Practices for Collecting Self-Reported Patient Demographic Data - *“Collecting data for us and not about us.”*

Introduction

The **Patient Experience Policy Forum (PXPF)**, supported by The Beryl Institute, is a broad-based and diverse coalition of patients, family members, caregivers and healthcare professionals led by a balanced board of patient/family partners and senior patient experience leaders. Our purpose is to advocate for and help shape policy at the national and local levels on issues that directly affect patient and family experience and elevate the human experience in healthcare. We do so through convening policy forums, educating policy makers, sponsoring advocacy events, providing communication updates and publishing calls to action.

Our structure translates into a unique perspective on patient experience policy issues. We combine the knowledge and understanding of how healthcare organizations work with the voices, preferences and values of patient/families/caregivers to create and advocate for policies that are both achievable and meaningful. Furthermore, our relationship with The Beryl Institute, comprised of 60,000+ members and affiliated individuals worldwide, and with representation from every state in the United States, connects us to a powerful platform from which to drive change and improvement in patient experience.

Our policy work is conducted by two PXPF workgroups: the Patient and Family Partnership Workgroup and the Measurement and Reporting Workgroup. This paper has been produced by the Measurement and Reporting Workgroup as part of the PXPF's advocacy focus on reducing preventable disparities in patient experience across patient populations, specifically by race, ethnicity or primary language (REaL) and sexual orientation and gender identity (SOGI). In order to reduce disparities, healthcare organizations must be able to analyze and utilize data to identify these disparities by stratifying measures of patient experience by REaL and SOGI. The fundamental building block of this ability is the collection of self-reported patient REaL and SOGI data by healthcare providers and organizations.

Data collection must be mandated in order to ensure it is conducted universally, reliably and in a standardized manner. Furthermore, a nationally coordinated effort amongst payors, regulators, and healthcare systems and practices should support this work both financially and through education, training and technical assistance to assure the best use of resources, reduction of duplicative and inconsistent efforts and actionability of data outputs.

This report describes why it is so important to collect REaL and SOGI data and highlights best practices for collection of these data. It includes a reference list and inventory of additional best practice resources.

Why Collect REaL and SOGI Data?

1. Every person, no matter where they receive healthcare, has the fundamental right to be treated fairly and equitably. This includes the right to choose how they wish to be identified. As declared in the Patient's Bills of Rights and endorsed by the American Hospital Association and other organizations, all patients, regardless of race, ethnicity, language, sexual orientation or gender identity, have the same right to the highest possible quality of care.
2. Disparities (also referred to as "inequities") in patient experience and quality of care across population groups defined in terms of REaL or SOGI are, by definition, suboptimal, unjust, and unacceptable.
3. Such inequities, in patient experience as well as other dimensions of performance — such as clinical quality and access to care — are known to exist across geographic regions, care settings and care providers.¹
4. Healthcare organizations cannot identify, and much less address, inequities related to patient demographics unless they know the key demographics of the people they serve and how their performance varies across their patient populations.
 - Although the collection of race, ethnicity and language data does not necessarily result in actions that will reduce disparities and improve care, the absence of the data guarantees that none of that will occur.²
 - In order to achieve health equity, we must be able to admit and demonstrate our baseline inequities (aka, "We can't improve what we can't measure").³
5. Collection of patient REaL and SOGI data is a prerequisite for understanding the needs of the population served, improving quality, and providing patient-centered care.
 - However, as David Williams has stated, "...we cannot necessarily assume that improving quality will reduce disparities; therefore, race-specific [ethnicity, language, sexual orientation, and gender] strategies may be necessary." Whereas Judith Hibbard has stated, "...we can improve quality by providing patient-centered care." Both tactics require healthcare organizations to know who their patients are.⁴
6. There are also numerous national and state reporting requirements that mandate collection of these data.⁵

Summary of Best Practices

1. Secure the tangible commitment of senior leadership as evidenced by the declaration of self-reported collection of REaL and SOGI data as a strategic priority, along with resources needed for execution.
2. Engage racial, ethnic, ancestral, language, sexual orientation and gender identity communities served by the organization. Explain why we are asking the REaL and SOGI questions and why we want their help and input.
 - Work with community advocates and leaders to build trust and relationships with these communities, ensuring that patient and family advisory councils are reflective of a diverse patient population and can help guide the development and implementation of these efforts.
3. Establish standardized processes/workflows for the collection of REaL and SOGI data that will be used for every patient across the organization.
 - Essence of the process must be patient self-reported data.
4. Set measurable goals regarding data collection and patient perceptions of discrimination and unfair treatment; monitor trends/progress toward the goals.
5. Establish race, ethnicity, language, sexual orientation and gender identity categories that will be used to capture the data, tailored to and with input from local communities/populations served and relevant advocacy organizations. Revisit categories and definitions every two years.
 - Ensure that granular ethnic backgrounds can be rolled up to the Office of Management and Budget (OMB) data categories.
6. Incorporate multiple opportunities for data collection into the standardized processes, including via patient portals prior to visits/admissions, at the time of registration and at least one other point of interaction with clinical staff.
 - Verify continued accuracy of REaL and SOGI data with patients at regular intervals, e.g., every two years.
7. Customize the EMR as needed to record REaL and SOGI categories relevant to the organization's patient population, utilizing IT staff and the EMR vendor as necessary.
 - Ensure that REaL and SOGI data fields are required to be completed by at least one role group, e.g., registration staff.
8. Ensure the confidentiality of the data.
 - Limit staff access to individual patient REaL and SOGI data to those who need to know, e.g., non-clinical staff do not need to know a patient's sex assigned at birth.
9. Design and implement staff training.
 - Acknowledge that both staff and patients may not be initially comfortable in asking or sharing REaL and SOGI data.
 - Key components of training include:
 1. Why these questions are being asked.

2. Nuts and bolts of how to ask the REaL and SOGI questions and how to enter the data.
 3. Responses to patients' frequently asked questions.
 4. Use of a standardized script when asking these questions.
- Incorporate input from relevant patient populations into training materials, reflecting what answers they desire and will find compelling as the reason(s) for sharing this information.
 - Create population-based approaches to communicating REaL and SOGI information for all staff regardless of access to electronic health record systems, such as environmental services or food services staff.
10. Design and implement messaging campaigns targeted to patients and communities/populations served.
- Why the questions are being asked; why REaL and SOGI data are important.
 - How the data will be used.
 - Who will have access to the data and for what purposes.
 - Messaging content and formats should be informed by the values, needs and concerns of racial, ethnic and SOGI communities served by the organization.
 - Messaging should be displayed/reinforced in multiple places at multiple times.

Important Considerations and Context

- For issues of race, language and gender identity, many registration and scheduling staff who are often responsible for collection of these data feel poorly equipped to answer the question of “Why?” In a 2006 Robert Wood Johnson Foundation report, two main barriers were identified and addressed: 1) staff reluctance and discomfort with discussing topics such as patient race, language and ethnicity; and 2) patient distrust about why the information is being collected and how it might be used in their care.⁶
- Regardless of race, ethnicity, language, sexual orientation or gender identity, adequate training must be created to support frontline staff in addressing concerns and feeling comfortable and confident in navigating these conversations with patients and families. Cultural competency skills and sensitivity training must go hand-in-hand with implementation of standards and processes.
- Many patients and families, especially those who have felt the effects of healthcare disparities and harm due to discrimination, are understandably distrustful of the healthcare system. They have not found the system to be designed with their unique needs in mind. In the mindset of “no data about us, without us,” healthcare organizations must partner with communities and patient/family groups to better understand motivations and messaging to achieve buy-in, so the organization can do better in caring for all populations.
- Likewise, in communities where White populations are the majority, there may be a decreased likelihood of capturing data across all populations, including native English speakers and those who may appear to staff visually as White or Caucasian. By ensuring universal data capture across all populations, the likelihood of accuracy increases, and the prevalence of data entries based upon assumptions decreases. However, staff may receive resistance to questions about race, ethnicity, etc. from people within the dominant racial group when that group sees itself as the “default” or the “norm.”

What are Best Practices for Collecting REaL and SOGI Data?

Getting Started

- Secure the commitment and support of senior leadership to declare that the reduction of healthcare disparities and achievement of equity among all patient populations served is a strategic priority, i.e., patients' REaL and SOGI data will be used to this end, and necessary resources will be committed.
 - » This includes ensuring that IT and Patient Access departments/leadership are committed and supportive.
- Engage racial, ethnic, sexual orientation and gender identity communities served by the organization as partners supporting the initiative—helping to form underlying understandings and language that communicates the reasons behind the “why we are asking.”
 - » Sponsor community forums, focus groups and dialogue with community leaders.
 - » Make every effort to ensure these interactions are respectful and focused on listening and learning.
- Establish a standardized process that will be utilized across the organization for collecting the data. Specifics concerning what should be included in this process are described in the following sections.
- Determine what current collection practices are the baseline and then set and monitor progress toward goals for data accuracy and completeness, e.g., < 2% non-response to REaL and SOGI questions. (“decline to answer” is a legitimate response to a question and should be counted as such, whereas the appearance of words such as “unknown” or “missing” in these data fields connotes a non-response and should not be counted as a response.)
- Establish accountability by assigning responsibility for implementing strategies for improving REaL and SOGI data collection and for monitoring progress toward goals.

From Whom is Data Collected?

- Patients and/or their caregivers should self-identify and self-report; this is considered the industry “gold standard.”⁷ Data should never be entered into the record based on staff or provider observation or “best guess.”

What Data Should be Collected?

- Determine the data categories that will be used. Define each category and embed definitions in data collection tools. As part of this process, seek the input of advocacy organizations and local populations served (e.g., ethnic background/ancestry groups, LGBTQ+) regarding how members prefer to be identified.
- Once data categories and definitions have been established, ensure they are used consistently across the entire organization in a standardized manner.
- Always include a “decline to answer” or “prefer not to say” option.
- Data categories, standards and terms should be revised, as necessary, whenever national stan-

dards are modified by governmental entities (e.g., Office of Management and Budget, Institute of Medicine) or by nationally recognized advocacy organizations. In the absence of such modifications, data categories, standards and terminology should be revisited regularly, e.g., every two years, with relevant advocacy groups and local populations served to ensure that terminology continues to align with evolving understanding of race, ethnicity, sexual orientation and gender identity and is not perceived as stigmatizing.

- The Appendix contains wording and recommended response categories for ethnicity, race, language, sexual orientation and gender identity questions.

Additional Considerations for Collecting Data

- Ask REaL and SOGI questions at periodic intervals, e.g., every two years; do not ask repeatedly, e.g., at every visit.
- Periodically validate collected data. Options include but are not limited to: 1) random selection of a sample of patients for interviews to compare stated REaL and SOGI responses to responses recorded in the medical record; or 2) less precisely and comprehensively, comparisons of stratification of patients by race and ethnicity using medical records as the source. Look at stratification of the geographic population served by race and ethnicity using U.S. census data (e.g., patient race profile based upon medical record review of 8% African American, 2% Asian, 90% White, compared to race profile of the population served based upon geographically defined service area of 12% African American, 3% Asian, 85% White, from census data.)
- Allow patients to distinguish language preference for clinical vs. non-clinical conversations.
- For patients with limited English proficiency, a command of basic, conversational English may result in an inaccurate assessment of English proficiency and a consequent lack of access to needed qualified medical interpretation services. A process should be in place to ensure that electronic medical record information is updated with language preference as soon as the need for language services is identified, such as during admission or any time thereafter. However, a study conducted in 2014 indicated that even with this process, an estimated 15% of patients needing interpreter services may still not be correctly identified.⁸
- Information about gender identify versus sex assigned at birth needs to be matrixed to protect patient privacy and safety. For example, documentation for admission and discharge should reflect the gender identity of the individual rather than legal sex wherever possible. Such documents should never display sex assigned at birth. These discussions should be kept private with the provider and any necessary billing staff to address specific healthcare concerns/screenings and ensure accurate patient identification with payors. Furthermore, any out-facing gender identity should reflect the patient's identity preference, i.e., use of patient preferred honorifics (Mr., Ms., first names for non-binary patients) when calling patients into an exam room; use of preferred identity on wristbands; use of preferred pronouns when addressing patients.

When Should Data be Collected and by Whom?

- There is no single answer to these questions. The type of care setting, space limitations, patient preferences and other factors may influence when REaL and SOGI questions are asked and by whom. The critical factor is that the process should be standardized and utilized consistently across all departments/units of the healthcare organization (every time, every patient.)

- In order to maximize response rates, there should be multiple opportunities for patients to be asked and to respond to these questions.
 - » Many patients may be most comfortable answering these questions via private self-report on a patient portal prior to arrival for a visit/admission. This option should be made available to all patients.
 - » Additional self-report options include check-in kiosks and check-in paperwork.
 - » Recognizing that some patients will not answer these questions prior to arrival, these patients should be provided the opportunity to self-identify during the check-in/registration process. In order to maximize patients' comfort in answering these questions, the registration interaction should be as private as possible. Data collected about sex assigned at birth should be done by the provider and only when medically necessary.
- Despite best efforts at process design and implementation, some patients will undoubtedly be missed during the registration process as well. Therefore, best practice calls for a second opportunity to collect these data face-to-face. For example, some organizations have found success in having behavioral health support specialists ask these questions during their encounters with patients.
- Preferred language and need for an interpreter, ideally, should be collected prior to arrival (such as with scheduling or via the patient portal) and confirmed upon admission or patient registration by the patient. This enables greater readiness and coordination of resources upon arrival, so patients and families feel welcomed and their needs anticipated. The registration staff should not be entering data for these fields based on their own perceptions of the patient's identity or language preferences.
- For questions of gender identity and sex assigned at birth, information should be self-reported by the individual either via the portal, prior to arrival, or during a private conversation with the provider when it pertains to clinical decision-making.
- In the case of health plans, REaL and SOGI information should be collected when the individual enrolls (as permitted by state law.)
- As noted in the previous section, patients' race, ethnicity, preferred language, sexual orientation and gender identity should be verified periodically, e.g., every two years. These questions should not be asked at every encounter.

How Should Data be Collected and Stored?

- REaL and SOGI data should be entered and stored in the electronic health record (EHR).
 - » Corresponding fields within the EHR should be structured as "required"—i.e., the question cannot be skipped or a warning should appear. This requirement should be defined by job role (e.g., required for registration staff but not for schedulers).
 - Identify all other roles who should have access to update the EHR fields, even if it is not a mandatory part of their workflow.
 - » The patient portal and EHR should interface so that data entered into one populates identical fields in the other.
- The EHR should be customized to accommodate the accurate recording of patient self-report-

ed responses to REaL and SOGI questions.

- » The order in which fields correspond to REaL and SOGI, as well as names attached to data fields, should be consistent with best practice (e.g., ask ethnicity questions first) and incorporate nomenclature that promotes patient preference and comfort.
- » Definitions of REaL and SOGI values should be embedded in the patient portal and the EHR to provide direction and guidance to the patient/registration staff, respectively, when answering/asking these questions.
- Staff responsible for asking patients REaL and SOGI questions (e.g., registration staff) should receive training in how to effectively do so. Such training helps to ensure both the capture of accurate data and alleviation of discomfort that staff may feel in asking these questions. Training should be periodically repeated to account for staff turnover and the need for reinforcement. Training modules should include:
 - » The rationale for collecting REaL and SOGI data—why these data are important.
 - » How to ask the questions and enter the responses into the patient record.
 - » How to answer common questions concerning the collection of REaL and SOGI data.⁹
- Standardized scripts for use by staff asking REaL and SOGI questions are considered best practice.
- Messaging to patients, family and the public is critically important to securing accurate and complete data. This should address: 1) why REaL and SOGI questions are being asked; 2) how these data are used; and 3) who will be able to see/access the data, i.e., its level of confidentiality.
 - » Research has found the following “why we collect it” themes to be particularly effective: 1) helps to assure that every patient receives the best possible care; 2) helps us to continuously improve the quality of care being given to all patients; and 3) helps us provide care free of racism and discrimination to all patients.^{10, 11}
 - » This messaging should also directly and transparently describe who has access to the REaL and SOGI data provided by patients and how those data will be used, with assurances that the data are confidential and won't be accessible at the individual patient level beyond these types of staff.
 - It may be added that this includes reporting to the government, which is done with aggregated data and not at the patient level.
 - » Messaging content and formats should be informed by the values, needs and concerns of racial, ethnic, sexual orientation and gender identity communities served by the organization, identified through the community partnering work described under “Getting Started.”
 - » This messaging and the rationale for asking should be embedded in the portion of the patient portal where REaL and SOGI questions are asked, i.e., appear as text on the screen and in the script used by any staff who ask REaL and SOGI questions.
 - » In addition, it is helpful to reinforce this messaging through displays (e.g., posters) in areas such as patient registration, waiting rooms, examination rooms and other patient-facing spaces.

- » Other opportunities for conveying the rationale for collecting REaL and SOGI data should be utilized as well (e.g., emails and texts to patients, newsletters, etc.)
 - One example of a messaging platform used by many leading healthcare organizations is “We Ask Because We Care,” created by the Robert Wood Johnson Foundation as part of its Aligning Forces for Quality Initiative.¹²
- Individual data storage should be highly secure but also facilitate the reporting and data aggregation needs and requirements of the complex healthcare system, e.g., requirements to report aggregated race, ethnicity and other demographic data to various state and federal entities.

How should performance be monitored?

- Processes for regularly monitoring performance and progress toward REaL and SOGI data collection goals should be designed and implemented.
- Examples of types of goals include:
 - » Percentages of patient records, by department and across the entire organization, in which REaL and SOGI data field responses are “unknown,” “blank,” or “decline to answer.”
 - » Accuracy of the data, as determined by verification with a representative sample of patients, and by comparisons with other sources where these data are captured, such as patient surveys.
 - » Percentage of the workforce (staff and clinicians) trained in the collection of REaL and SOGI data and who can answer the questions of why REaL and SOGI data are being collected and how those data are being used.
 - » Percentages of quality metrics, patient experience metrics, and other performance metrics that are routinely stratified by REaL and SOGI for analytic purposes to identify disparities.
 - » Trends over time in percentages of patients, by REaL and SOGI categories, who feel known, respected, understood and comfortable with revealing their respective identities and who believe they received culturally safe care when interacting with the organization, as assessed by patient surveys or interviews.
 - » Conversely, trends over time in percentages of patients, by REaL and SOGI categories, who feel that they experienced discrimination in their interactions with the organization or received unfair treatment, as assessed by patient surveys or interviews.
- Dashboards and other reporting formats should be utilized to impart performance information to all levels and departments of the healthcare organization, including senior management.¹⁴
 - » Consider augmenting quantitative information with patient narratives and stories, particularly (for this purpose) stories that include experiences/reactions related to being asked REaL and SOGI questions.
- As needed, interventions to improve performance should be implemented. Staff assigned responsibility for meeting REaL and SOGI data collection goals should be given the resources and authority to do so.

Acknowledgements

Thank you to members of the PXP Measurement and Reporting Workgroup for their contributions to this document:

Nicole Cable

Lena Cuthbertson

Maria Fernandez

Stephanie Fishkin

Francis Fullam

Helen Haskell

Mitchell Kellaway

Sofie Morgan

Liza Nyeko

Robert Parrish

Dale Shaller

Sara Toomey

Gail Watkins

References

1. Centers for Medicare & Medicaid Services, Office of Minority Health. (2021). Racial, Ethnic and Gender Disparities in Health Care in Medicare Advantage. Retrieved from <https://www.cms.gov/about-cms/agency-information/omh>
2. Institute of Medicine. 2009. *Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement*. Washington, DC.
3. Mass General Brigham. (January, 2022). *Experiences Collecting Self-Reported Race, Ethnicity and Other Demographic Data*. Presentation to PXP Measurement and Reporting Workgroup.
4. Robert Wood Johnson Foundation Conference on Disparities and Quality of Care. (2005). <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1797091/>
5. Hasnain-Wynia, R., Pierce, D., Haque, A., Hedges Greising, C., Prince, V., Reiter, J. (2007) *Health Research and Educational Trust Disparities Toolkit: Why Collect Race, Ethnicity and Primary Language*. <http://www.hretdisparities.org>
6. Regenstein M., Sickler D. (2006). *Race, Ethnicity and Language of Patients: Hospital Practices Regarding Collection of Information to Address Disparities in Health Care*. RWJ Foundation and National Public Health and Hospital Institute.
7. Jarrin OF, Nyandege AN, Grafova IB, Dong X, Lin H. Validity of Race and Ethnicity Codes in Medicare Administrative Data Compared With Gold-standard Self-reported Race Collected During Routine Home Health Care Visits. *Med Care*. 2020;58(1):e1. doi:10.1097/MLR.0000000000001216
8. Boscolo-Hightower A., Rafton S., et al. (2014). Identifying Families With Limited English Proficiency Using a Capture-Recapture Approach. *Hospital Pediatrics*, 4(1), 16-22. <https://doi.org/10.1542/hpeds.2013-0040>
9. Hasnain-Wynia, R., Pierce, D., Haque, A., Hedges Greising, C., Prince, V., Reiter, J. (2007) *Health Research and Educational Trust Disparities Toolkit*. www.aha.org/system/files/2018-01/equity-care-toolkit-eliminating-health-care-disparities-2015.pdf#page=4
10. Hasnain-Wynia, R., Race, Ethnicity and Language Data Collection: Nuts and Bolts. <https://hcup-us.ahrq.gov/datainnovations/raceethnicitytoolkit/ca11.pdf>
11. Agency for Healthcare Research and Quality. (2016, July). Race and Ethnicity Data Improvement Toolkit. Healthcare Cost and Utilization Project (HCUP). FAQs About Collection of REaL Data. <https://hcup-us.ahrq.gov/datainnovations/raceethnicitytoolkit/ca5.jsp>
12. Robert Wood Johnson Foundation. (2010, September 13). *We Ask Because We Care*. Aligning Forces for Quality. <http://forces4quality.org/node/4185.html>
13. AHA Institute for Diversity and Health Equity. (2020, December). *Health Equity, Diversity & Inclusion Measures for Hospitals and Health System Dashboards*. https://ifdhe.aha.org/system/files/media/file/2020/12/ifdhe_inclusion_dashboard.pdf
14. Institute of Medicine. 2009. *Race, Ethnicity and Language Data: Standardization for Health Care Quality Improvement*. Washington D.C.: The National Academies Press. <https://doi.org/10.17226/12696>

15. National LGBTQIA+ Health Education Center. (2022). *Ready, Set, Go!: A Guide For Collecting Data On Sexual Orientation and Gender Identity*. https://www.lgbtqiahealtheducation.org/wp-content/uploads/2022/05/TFIE-64_Updates2022_ReadySetGo_10_18_22.pdf

Bibliography

1. Agency for Healthcare Research and Quality. (2010, May). *Race, Ethnicity and Language Data: Standardization for Health Care Quality Improvement*. (Publication: 10-0058-EF). <https://www.ahrq.gov/research/findings/final-reports/iomracereport/reldata5.html>
2. Agency for Healthcare Research and Quality. (2016, July). Race and Ethnicity Data Improvement Toolkit. Healthcare Cost and Utilization Project (HCUP). www.hcup-us.ahrq.gov/datainnovations/raceethnicitytoolkit/case_for_re.jsp. <https://www.hcup-us.ahrq.gov/datainnovations/raceethnicitytoolkit/nm1.pdf>
3. AHA Institute for Diversity and Health Equity. (2020, December). *Health Equity, Diversity & Inclusion Measures for Hospitals and Health System Dashboards*. https://ifdhe.aha.org/system/files/media/file/2020/12/ifdhe_inclusion_dashboard.pdf
4. AHA Institute for Diversity and Health Equity. (2012, March). *Health Equity Resource Series: Data-Driven Care Delivery: Data Collection, Stratification and Use*. https://ifdhe.aha.org/system/files/media/file/2021/04/ifdhe_real_data_toolkit_1.pdf
5. Ali, P., (2015). Speaking My Patient's Language: Providing Language Concordant Care to Patients With Limited English Proficiency. Technical Report.
6. Ali, P., Language Barriers and their impact of Provision of Care to patients with limited English Proficiency: Nurses Perspective. *Journal of Clinical Nursing*, 27 (5-6). e1152-e1160.
7. Baker, D. W., Hasnain-Wynia, R., Kandula, N. R., Thompson, J. A., & Brown, E. R. (2007). Attitudes toward Health Care Providers, Collecting Information about Patients' Race, Ethnicity, and Language. *Medical Care*, 45(11), 1034-1042. <http://www.jstor.org/stable/40221578>
8. Boscolo-Hightower A., Rafton S., et al. (2014). Identifying Families With Limited English Proficiency Using a Capture-Recapture Approach. *Hospital Pediatrics*, 4(1), 16-22. <https://doi.org/10.1542/hpeds.2013-0040>
9. Centers for Medicare & Medicaid Services, Office of Minority Health. (2021). Racial, Ethnic and Gender Disparities in Health Care in Medicare Advantage. Retrieved from <https://www.cms.gov/about-cms/agency-information/omh>
10. Centers for Medicare & Medicaid Services, (2022). *Inventory of Resources for Standardized Demographic and Language Data Collection*. <https://www.cms.gov/about-cms/agency-information/omh/downloads/data-collection-resources.pdf> (This document is a highly comprehensive inventory of resources related to the collection of REaL and disparities data.)
11. Centers for Medicare and Medicaid Services, (2022). *Sexual and Gender Minority Clearinghouse*. <https://www.cms.gov/About-CMS/Agency-Information/OMH/resource-center/hcps-and-researchers/data-tools/sgm-clearinghouse>
12. Chin MH. (2015). Using patient race, ethnicity, and language data to achieve health equity. *J Gen Intern Med*. Jun;30(6):703-5. doi: 10.1007/s11606-015-3245-2. PMID: 25722093; PMCID: PMC4441661.
13. The Commonwealth Fund. (2008, January 14) *Collecting Race, Ethnicity, and Language Data: A How-To Guide*. <https://www.commonwealthfund.org/publications/other-publication/2008/jan/collecting-race-ethnicity-and-language-data-how-guide>

14. Cowden, J.D., Flores, G., Chow, T. et al. (2020). Variability in Collection and Use of Race/Ethnicity and Language Data in 93 Pediatric Hospitals. *J. Racial and Ethnic Health Disparities*, 7, 928–936. <https://doi.org/10.1007/s40615-020-00716-8>
15. Hasnain-Wynia R, Baker DW. (2006). Obtaining data on patient race, ethnicity, and primary language in health care organizations: current challenges and proposed solutions. *Health Serv Res*. Aug;41(4 Pt 1):1501-18. doi: 10.1111/j.1475-6773.2006.00552.x. PMID: 16899021; PMCID: PMC1797091
16. Hasnain-Wynia, R., Pierce, D., Haque, A., Hedges Greising, C., Prince, V., Reiter, J. (2007) *Health Research and Educational Trust Disparities Toolkit*. www.aha.org/system/files/2018-01/equity-care-toolkit-eliminating-health-care-disparities-2015.pdf#page=4 (This is a highly useful resource, containing easy to find best practices.)
17. Hasnain-Wynia, R., Race, Ethnicity and Language Data Collection: Nuts and Bolts. <https://hcup-us.ahrq.gov/datainnovations/raceethnicitytoolkit/ca11.pdf> (Presentation on staff training regarding REaL data collection.)
18. Health Research & Educational Trust. (2013, August). *Reducing health care disparities: Collection and use of race, ethnicity and language data*. Chicago: Health Research & Educational Trust. http://www.hpoe.org/Reports-HPOE/Equity_Care_Report_August2013.PDF (This is an excellent resource with easy-to-find best practices.)
19. Institute of Medicine. 2009. *Race, Ethnicity and Language Data: Standardization for Health Care Quality Improvement*. Washington D.C.: The National Academies Press. <https://doi.org/10.17226/12696>.
20. Jorgensen, S., Thorlby, R., Weinick, R.M. et al. (2010). Responses of Massachusetts hospitals to a state mandate to collect race, ethnicity and language data from patients: a qualitative study. *BMC Health Serv Res*. 10, 352. <https://doi.org/10.1186/1472-6963-10-352>
21. Lee, W. C., Veeranki, S. P., Serag, H., Eschbach, K., & Smith, K. D. (2016). Improving the Collection of Race, Ethnicity, and Language Data to Reduce Healthcare Disparities: A Case Study from an Academic Medical Center. *Perspectives in Health Information Management*, 13(Fall), 1g.
22. Mass General Brigham. (January, 2022). *Experiences Collecting Self-Reported Race, Ethnicity and Other Demographic Data*. Presentation to PXP Measurement and Reporting Workgroup.
23. Maulik, J. (2009, November 4). *Collecting Race, Ethnicity and Language Data: Why is it Important?*. Health Research and Educational Trust. http://www.hpoe.org/HPOE_Live_PPT/Disparities_HRET_11_4_2009_webinarpresentation
24. National LCBTQIA+ Health Education Center. (2022). *Ready, Set, Go! A Guide for Collecting Data on Sexual Orientation and Gender Identity*. https://www.lgbtqihealtheducation.org/wp-content/uploads/2022/05/TFIE-64_Updates2022_ReadySetGo_10_18_22.pdf (This is a key resource to guide collection of SOGI data , includes processes, scripts, videos, scenario-based training, and more.)
25. The Persons With Lived Experience Advisory Committee. (2020, October 08). *Putting Patients First: Toward Equity-Informed Care*. Equity in Palliative Approaches to Care. <https://my.visme.co/view/4d86k3d9-ppf20>

26. Regenstein M., Sickler D. (2006). *Race, Ethnicity and Language of Patients: Hospital Practices Regarding Collection of Information to Address Disparities in Health Care*. RWJ Foundation and National Public Health and Hospital Institute.
27. Robert Wood Johnson Foundation Conference on Disparities and Quality of Care. (2005).
28. Robert Wood Johnson Foundation. (2010, September 13). *We Ask Because We Care*. Aligning Forces for Quality. <http://forces4quality.org/node/4185.html>
29. Rodriguez, J., Casillas, A., Cook, B., Marlin, R. (2021). The Language of Equity in Digital Health: Prioritizing the needs of Limited English Proficient Communities in the Patient Portal 2.0. *Journal of Health Care for the Poor and Underserved*, Vol. 32, No. 2. 211-219.
30. Shapiro, A., Meyer, D., Riley, L., et. al. (2021, September 1). Building the Foundations for Equitable Care. *New England Journal of Medicine Catalyst*. <https://catalyst.nejm.org/doi/full/10.1056/CAT.21.0256> (This article is a succinct, yet highly informative overview of New York-Presbyterian's journey related to collection of self-reported REaL data.)
31. Ulmer C, McFadden B, Nerenz D. (2009). *Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement*. Agency for Health Care Research and Quality. <https://www.ahrq.gov/research/findings/final-reports/iomracereport/reldata5.html>

Appendix

Specific Race, Ethnicity, Language, Sexual Orientation and Gender Identity Questions

These questions, their sequence, and the response categories corresponding to each question are believed to be best practice as of late 2022. The broad ethnicity and race response categories reflect current Office of Management and Budget (OMB) standards. The ethnic background and language response categories, as stated below, reflect Institute of Medicine recommendations. The sexual orientation and gender identity response categories are those recommended by the National LGBTQIA+ Health Education Center, affiliated with the Fenway Institute in Boston, a nationally recognized and highly respected organization for topics related to LGBTQIA+ healthcare.

Before utilizing these questions and response categories, the user is strongly encouraged to conduct appropriate research to determine whether they continue to be viewed as "current best practice."

- Ask ethnicity questions first
 - » What is your ethnicity?
 - Hispanic or Latino
 - Not Hispanic or Latino
 - Don't know
 - Decline/prefer not to answer
 - » What is your ethnic background or ancestry?
 - Use local census data and knowledge of population served to narrow the number of categories listed to those that are locally relevant (subset of 500+ Institute of Medicine (IOM) categories.)¹
 - Ensure that categories can be "rolled up" to the OMB ethnicity and race categories (e.g., granular "Dominican" can roll up to be counted under "Hispanic"; "Japanese" can roll up to be counted under "Asian."
 - » Ethnicity: Hispanic or Latino; Not Hispanic or Latino
 - » Race: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, Other, White

(Continued on next page >)

- Ask race questions after ethnicity questions
 - » Which category or categories best describe your race (mark every category that applies to you (multiple races)?
 - American Indian or Alaska Native
 - Asian
 - Black or African American
 - Native Hawaiian or Other Pacific Islander
 - White
 - Don't know
 - Decline/prefer not to answer
 - » Whenever possible when reporting, identify all racial categories; when not feasible, count patients with multiple races under "Two or More Races."
- Ask language questions
 - » How well do you speak English?
 - Very well
 - Well
 - Not well
 - Not at all
 - Don't know
 - Decline/prefer not to answer
 - » Would you like an interpreter?
 - Yes
 - No
 - Don't know
 - Decline/prefer not to answer
 - » What language do you prefer when speaking to doctors, nurses and other health care staff?
 - Use local census data and knowledge of population served to narrow the number of categories listed (subset of 600+ IOM categories).¹
 - » What written language do you prefer for reading medical information and instructions?
 - Use local census data and knowledge of population served to narrow the number of categories listed (subset of 600+ IOM categories).¹

(Continued on next page >)

- Ask sexual orientation question²
 - » Do you think of yourself as: (Check all that apply) or (Please choose the option that best describes you. Currently our system allows only one option.)
 - Lesbian or gay
 - Straight or heterosexual
 - Bisexual
 - Queer
 - Pansexual
 - Asexual
 - Something else: _____
 - Don't know
 - Decline/prefer not to answer
- Ask gender identity question²
 - » What is your current gender identity? (Check all that apply) or (Please choose the option that best describes you. Currently our system allows only one option.)
 - Female/woman/girl
 - Male/man/boy
 - Nonbinary, genderqueer, or not exclusively female or male
 - Transgender female/woman/girl
 - Transgender male/man/boy
 - Another gender: _____
 - Don't know
 - Decline/prefer not to answer
- Ask sex assigned at birth question²
 - » What sex were you assigned at birth, as listed on your original birth certificate? (Check one.)
 - Female
 - Male
 - X/Another sex: _____
 - Don't know
 - Decline/prefer not to answer