

Sexual Orientation and Gender Identity (SOGI) Data Collection Checklist*

[*Use with the Sexual Orientation and Gender Identity (SOGI) Data Privacy Checklist]

- I will work with experts in survey (and particularly SOGI survey) methodology and analytics^a and follow recommended guidance for statistical surveys.^{1, 2, 3}
- I have defined the purpose of the data collection (e.g., to answer a scientific question, to confirm study eligibility, to support inclusion of diverse participants, to understand patient or caregiver barriers and types of support needed, to understand epidemiology, to assess research or advisory team composition, for administrative purposes)⁴ and will communicate this purpose to survey participants in consent forms.
- I will only collect the data necessary for the defined purpose (as outlined in the checkpoint above) and to answer and the study question/s.^{5, 6}
- I recommend that SOGI data is collected along with other demographic data, like age, race, and ethnicity, and that it is self-reported (filled out by the participant or potentially by a family member in household surveys and not by the researcher). That said, I will not require demographic data collection, inclusive of SOGI data, and will ensure that individuals, sites, and organizations can always opt-out.
- I have documented and clearly explained in all participant-facing materials the potential risks of SOGI data collection and how the participant's data will be protected,⁷ why SOGI data collection is essential to the project or study question, how collecting SOGI data may benefit the participants, and the intended use of the SOGI data.
- I have defined the population of interest through reference to (or determination of) the epidemiology, the burden of disease, and other factors (e.g., social determinants of health) relevant to the data collection purpose and study question/s subject to data availability.⁸
- I have vetted the decision to collect SOGI data, the collection method, the level of subgroup detail/disaggregation, and the data collection forms and survey format (e.g., in-person oral, paper, online app) with working groups and/or advisory boards that are inclusive of people who are lesbian or gay, and people who are transgender, non-binary, or intersex. Please note that it can be helpful to add a date stamp and version number to all materials, to keep a list of reviewers associated with each version, and to track the substantiveness of community input into those materials.
- I have reviewed and mapped proposed data elements to select the options that best align with the data collection purpose and study question/s, current guidance on SOGI data

^a We caution that data collection can easily go awry, and the results can end up being detrimental to the people from whom data was collected. Survey research methods, and the associated preparation for analysis, are skills that can take many years of experience and study to develop. Therefore, we list working with experts in these fields first on this checklist, in addition to a thorough grounding in current guidance.

collection,^{9, 10, 11, 12, 13, 14, 15, 16, 17} and the lived experiences of the population of interest. Proposed data elements may include:

- Terminology
 - Data collection questions (e.g., sexual orientation, gender identity, sex, whether a person is intersex)
 - Response choices (e.g., [for sexual orientation] lesbian or gay, straight, bisexual, other)
 - Response formats (e.g., open-response, checking-off multiple options)
 - Question and response choice ordering (and implications for skip patterns on subsequent survey questions)
- I have advised sites to create space in the data collection forms for participants to report their pronouns if they choose to do so. I have recommended that staff who interact with participants periodically check-in with them about their pronouns, legal name, and chosen/nickname. For example, the researcher or clinical staff member could say, “Some people update their names from time to time. Would you like to report any changes?” If the participant reports a change, the staff interacting with the participant should inform them of where this information will be viewable (e.g., by the research team, by lab personnel, and/or on patient identification wristbands used at health facilities).
- I have prepared with my team operational plans, budgets, and timelines for:
- Piloting and validating the data collection forms, portals, and/or apps and all participant-facing materials related to data collection, as appropriate.
 - Evaluation, analysis, and dissemination of results.
 - Translation of the data collection forms, portals, and/or apps and all participant-facing materials related to data collection as needed.
 - Licensing fees for any software (potentially including Electronic Health Records systems), graphics/design, or copyrighted materials.
 - Staff training (e.g., cultural humility, how to ask SOGI questions to adults and to children, how to respond to questions from participants.)
 - Private and safe settings for participants to fill out or orally respond to data collection questions at all stages of the trial. This will vary depending upon the survey methods and the context. In each case, please be mindful of ethical and equity considerations (e.g., meetings times, locations, signage, visibility of entrances and rooms, gender identity of facilitator/s). For countries and local regions where there are punitive laws and social environments that discriminate against people who are LGBTQIA+, sites, IRBs, and/or researchers might choose not to ask the optional SOGI questions developed for common protocols or consents. Finally, please see the data privacy checklist for SOGI data privacy considerations specific to pediatric populations.
- I have used culturally respectful language,¹⁸ precise (and defined) terminology, and gender-neutral color palettes and imagery, where appropriate and feasible.

-
- ¹ US Office of Management and Budget. (2006). Standards and Guidelines for Statistical Surveys. Available from: https://www.whitehouse.gov/wp-content/uploads/2021/04/standards_stat_surveys.pdf
- ² US Office of Management and Budget (2016). Statistical Policy Directive No. 2 Addendum: Standards and Guidelines for Cognitive Interviews. Available from: https://www.whitehouse.gov/wp-content/uploads/2021/04/final_addendum_to_stat_policy_dir_2.pdf
- ³ US Census Bureau. (2023). Statistical Quality Standards. Available from: <https://www2.census.gov/about/policies/quality/quality-standards.pdf>
- ⁴ “Administrative data” is defined by the US Office of Management and Budget (OMB) as: “Refers to administrative, regulatory, law enforcement, adjudicatory, financial, or other data held by agencies and offices of the government or their contractors or grantees (including States or other units of government) and collected for other than statistical purposes. Administrative data are typically collected to carry out the basic administration of a program, such as processing benefit applications or tracking services received. These data relate to individuals, businesses, and other institutions.” OMB Guidance for Providing and Using Administrative Data for Statistical Purposes. Available from: <https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2014/m-14-06.pdf>
- *Please note that administrative data is different from survey data and requires different analytical techniques. For a summary of differences between these and examples of using different types of data collection for different purposes, see the OMB White Paper (and especially Table 1) on “Using Administrative and Survey Data to Build Evidence.” Available from: https://obamawhitehouse.archives.gov/sites/default/files/omb/mgmt-gpra/using_administrative_and_survey_data_to_build_evidence_0.pdf
- ⁵ Holzberg J, Ellis R, Virgile M, Nelson D, Edgar J, Phipps P, Kaplan R. (2017). Assessing the feasibility of asking about gender identity in the current population survey: Results from focus groups with members of the transgender population. Available from: <https://www.bls.gov/osmr/research-papers/2017/pdf/st170200.pdf>
- ⁶ Reisner SL, Deutsch M, Bhasin S, Bockting W, Brown GR, Feldman J, Garofalo R, Kreukels B, Radix A, Safer JD, Tangpricha V, T’Sjoen G, Goodman M. (2016). Advancing methods for US transgender health research. *Curr Opin Endocrinol Diabetes Obes*, 23(2): 198-207.
- ⁷ US Department of Health and Human Services. 2023. Summary of the HIPAA Security Rule. Available from: <https://www.hhs.gov/hipaa/for-professionals/security/laws-regulations/index.html>
- ⁸ The NIH has collated examples of surveys that have sexual orientation and/or gender identity measures, examples of (and summary findings on) sexual orientation identity measures in the United States and Other Countries, examples of (and summary findings on) two-step gender identity measures in national and international surveys, and a summary of findings nonbinary sex measure and evaluation criteria. Available from: <https://dpcpsi.nih.gov/sgmro/measurement-and-data/surveys-and-measures>
- ⁹ National Academies of Sciences, Engineering, and Medicine. (2022). Measuring sex, gender identity, and sexual orientation. Available from: <http://nap.edu/26424>
- ¹⁰ The White House. (2023). Recommendations on the best practices for the collection of sexual orientation and gender identity data on federal statistical surveys. Available from: <https://www.whitehouse.gov/wp-content/uploads/2023/01/SOGI-Best-Practices.pdf>
- ¹¹ Fenway Health. (2022). Ready, set, go! A guide for collecting data on sexual orientation and gender identity. Available from: <https://www.lgbtqihealtheducation.org/publication/ready-set-go-a-guide-for-collecting-data-on-sexual-orientation-and-gender-identity-2022-update/>

-
- ¹² Federal Committee on Statistical Methodology. (2021). Why Do Federal Agencies Ask About Sexual Orientation and Gender Identity on Surveys? Available from: <https://www.fcsm.gov/assets/files/docs/FCSM%2021%2001%20062221.pdf>
- ¹³ Executive Office of the President of the United States. Subcommittee on Sexual Orientation, Gender Identity, and Variations in Sex Characteristics (SOGI) Data Subcommittee on Equitable Data of the National Science and Technology Council. (2023). Federal Evidence Agenda on LGBTQIA+ Equity. Available from: <https://www.whitehouse.gov/wp-content/uploads/2023/01/Federal-Evidence-Agenda-on-LGBTQI-Equity.pdf>
- ¹⁴ Heidari S, Babor TF, De Castro P, Tort S, Curno M. (2016). Sex and Gender Equity in Research: Rationale for the SAGER Guidelines and Recommended Use. *Research Integrity and Peer Review* 1:2. Available from: <https://researchintegrityjournal.biomedcentral.com/articles/10.1186/s41073-016-0007-6>
- ¹⁵ Raising the Bar on Sex and Gender Reporting in Research. (2022). *Nature Medicine* 28, 1099. Available from: <https://www.nature.com/articles/s41591-022-01860-w>
- ¹⁶ US National Institute for Health Sexual & Gender Minority Research Office. (2023). Sexual & Gender Minority Measurement & Data: Surveys and Measures. Available from: <https://dpcpsi.nih.gov/sgmro/measurement-and-data/surveys-and-measures>
- ¹⁷ Cheloff, A.Z., Jarvie, E., Tabacac, A.R., Katz-Wise, S.L., Fygetakis, L.M., & Keuroghlian, A. (2022). Sexual Orientation, Gender Identity, and Sex Development: Recommendations for Data Collection and Use in Clinical, Research, and Administrative Settings. Harvard Medical School, Dean's LGBT Advisory Committee. Available from: <https://dicp.hms.harvard.edu/sites/default/files/2022-10/SOGI%20Data%20Collection.pdf>
- ¹⁸ Fenway Health. (2024). Cultural Adaptation of Measures and Tools for Sexual Orientation and Gender Identity (SOGI) Data Collection. Available from: <https://www.lgbtqihealtheducation.org/publication/cultural-adaptation-of-measures-and-tools-for-sexual-orientation-and-gender-identity-sogi-data-collection/>