MULTI-REGIONAL
CLINICAL TRIALS
THE MRCT CENTER OF
BRIGHAM AND WOMEN'S HOSPITAL
and HARVARD

Representation
in Research
LGBTQIA+

Tool 3: SOGI Data Collection Checklist

All study participants should be able to see themselves reflected in the research data. Yet, standard demographic variables of sexual orientation and gender identity (SOGI) have rarely been reported for clinical trials. We therefore don't know whether LGBTQIA+ people are able to participate in clinical trials, the extent to which clinical trials reflect sexual and gender identify diversity, or whether the safety and efficacy of tested products differs for any LGBTQIA+ participants. To begin to address this gap that impacts respect for persons, study generalizability, and beneficence, and to support research teams, sites, and sponsors in following this recommendation we developed the SOGI Data Collection Checklist.

This SOGI Data Collection Checklist provides prompts to consider (note especially those prompts in bold) for promoting affirming and robust data collection, which have been drawn from published guidance and the insights of working group members (convened by the MRCT Center) who have been leading in this the fields of piloting survey methodology and SOGI data collection. In addition, please note that this SOGI Data Collection Checklist is meant to be utilized in tandem with the SOGI Data Privacy Checklist.

Checklist

- ☐ **I will always create an option for participants to opt out of answering any or all SOGI data questions and make clear in all participant-facing materials that this option may be exercised at any time during the trial.**
- ☐ **I will only collect the data necessary for the defined purpose (as outlined in the checkpoint below) and to answer the study question/s.**[1, 2]
- ☐ Before creating data collection plans and templates, I will clearly define the purpose of the data collection (e.g., to answer a scientific question, to confirm study eligibility, to support the 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).[a]

☐ Before creating data collection plans and templates, I will work with experts in survey (and particularly SOGI survey) methodology and analytics[b] and follow recommended guidance for statistical surveys.[3, 4, 5,c]

☐ Before creating data collection plans and templates, **I will transparently discuss when/where to potentially 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 involve people who are LGBTQIA+, including people who are transgender, non-binary, or intersex.** I have an efficient process for securely documenting this input and sharing updates back with community members.[d]

---

[a] "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

[b] We caution that data collection can easily go awry, and the results can end up being detrimental to the people from whom data were 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.

[c] 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 Unites 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.

[d] Please note that it can be helpful to add a date stamp and version number to all SOGI data collection development materials, to keep a list of reviewers associated with each version, and to track the substantiveness of community input into those materials.

☐ Before creating data collection plans and templates, 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.[e]

☐ I will collect/recommend [to sites/research teams] to collect SOGI data along with other demographic data, like age, race, and ethnicity, and **ensure that SOGI data is self-reported (by the participant or potentially by a family member in household surveys) and not assumed/filled out by the researcher.**

☐ **I will clearly explain in all participant-facing materials the potential risks and benefits of SOGI data collection and how the participant's data will be stored and protected (and the situations in which we might not be able to protect their privacy), 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 reviewed and mapped proposed data elements to select the options that best align with the data collection purpose and study question/s, current guidance and discussion on SOGI data collection,[6, 7, 8, 9, 10, 11, 12, 13, 14, 15,16, 17] and the lived experiences of the population of interest. Proposed data elements may include:

- 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, Two-Spirit[18] (or other culturally specific term/s, where appropriate), "I use a different term").
- Terminology (i.e., how terms used in data collection questions and response choices [such as sexual orientation, lesbian, bisexual] are defined.

---

[e] 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 Unites 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. [Previously] available from: https://dpcpsi.nih.gov/sgmro/measurement-and-data/surveys-and-measures

    - 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 and common/name used if they choose to do so.[f] I have recommended that staff who interact with participants periodically check-in with them about their pronouns, legal name, and common/name used. 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, timelines, and accountability mechanisms for:

    - Readily accessible and regularly completed staff training (e.g., cultural competence, how to ask SOGI questions to adults and to children, how to respond to questions from participants).
    - Piloting and refining the data collection forms, portals, and/or apps and all participant-facing materials related to data collection, as appropriate.
    - 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, graphics/design, or copyrighted materials.
    - Evaluation, analysis, and dissemination of results.

---

[f] A common/name may also be known as a chosen name, nickname, familiar name, or informal name. Also, please note, it is important to discuss with your institution how to best protect and store information about a participant's pronouns and common/name (and any changes in those). This may vary from a hand-written note in the participants file, to electronic means of data storage, depending upon the political environment and organization's electronic security capabilities.

☐ I have established (or confirmed that the site and research team are providing) private and safe settings for participants to fill out or orally respond to data collection questions at all stages of the trial.[g]

☐ I have used culturally respectful language,[19] precise (and defined) terminology, and gender-neutral color palettes and imagery, where appropriate and feasible.[20, 21]

---

[g] Please note, 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 nor document 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.

References

Tool 3: SOGI Data Collection Checklist

1. 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*. U.S. Bureau of Labor Statistics. Available from: https://www.bls.gov/osmr/research-papers/2017/pdf/st170200.pdf
2. Reisner, S. L., Deutsch, M., Bhasin, S., Bockting, W., Brown, G. R., Feldman, J., Garofalo, R., Kreukels, B., Radix, A., Safer, J. D., Tangpricha, V., T'Sjoen, G., & Goodman, M. (2016). Advancing methods for US transgender health research. *Current Opinion in Endocrinology, Diabetes and Obesity, 23*(2), 198–207. Available from: https://pubmed.ncbi.nlm.nih.gov/26845331/
3. US Office of Management and Budget. (2006). *Standards and Guidelines for Statistical Surveys*. The White House. Available from: https://www.whitehouse.gov/wp-content/uploads/2021/04/standards_stat_surveys.pdf
4. US Office of Management and Budget. (2016). *Statistical Policy Directive No. 2 Addendum: Standards and Guidelines for Cognitive Interviews*. The White House. Available from: https://www.whitehouse.gov/wp-content/uploads/2021/04/final_addendum_to_stat_policy_dir_2.pdf
5. US Census Bureau. (2023). *Statistical Quality Standards*. U.S. Department of Commerce. Available from: https://www2.census.gov/about/policies/quality/quality-standards.pdf
6. National Academies of Sciences, Engineering, and Medicine. (2022). *Measuring sex, gender identity, and sexual orientation*. The National Academies Press. Available from: http://nap.edu/26424
7. 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
8. Fenway Health. (2022). *Ready, set, go! A guide for collecting data on sexual orientation and gender identity*. Available from: https://www.lgbtqiahealtheducation.org/publication/ready-set-go-a-guide-for-collecting-data-on-sexual-orientation-and-gender-identity-2022-update/

9. 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

10. 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

11. Heidari, S., Babor, T. F., 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

12. Raising the bar on sex and gender reporting in research. (2022). *Nature Medicine, 28*, 1099. Available from: https://doi.org/10.1038/s41591-022-01860-w

13. U.S. National Institute for Health Sexual & Gender Minority Research Office. (2023). *Sexual & gender minority measurement & data: Surveys and measures*. U.S. Department of Health and Human Services. Available from: https://dpcpsi.nih.gov/sgmro/measurement-and-data/surveys-and-measures

14. Cheloff, A. Z., Jarvie, E., Tabaac, 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

15. Pratt-Chapman, M.L., Tredway K., Wheldon, C.W., Streed, C.G. Jr., Scout, N.F.N, Ose, J., & Jackson, S.S. (2024). Strategies for Advancing Sexual Orientation and Gender Identity Data Collection in Cancer Research. JCO Oncology Practice, 20(7), 883-889. Available from: https://ascopubs.org/doi/pdf/10.1200/OP.23.00629

16. Clinical Data Interchange Standards Consortium. Sexual Orientation & Gender Identity eCRF Considerations. Available from: https://www.cdisc.org/kb/ecrf/sexual-orientation-gender-identity-sogi

17. Keuroghlian AS, Goldhammer H. Creating Welcoming and Affirming Clinical Studies for LGBTQIA+ Participants Toolkit. The Fenway Institute and Johnson & Johnson; 2024. Available from: https://www.lgbtqiahealtheducation.org/collection/clinical-studies-toolkit/

18. Government of Canada. (2024). *2SLGBTQI+ Glossary of Terms.* Available from: https://www.canada.ca/en/women-gender-equality/free-to-be-me/2slgbtqi-plus-glossary.html

19. Fenway Health. (2024*). Cultural Adaptation of Measures and Tools for Sexual Orientation and Gender Identity (SOGI) Data Collection.* Available from: https://www.lgbtqiahealtheducation.org/publication/cultural-adaptation-of-measures-and-tools-for-sexual-orientation-and-gender-identity-sogi-data-collection/

20. Alpert, A. B., Hernandez, T., Ruddick, R., Manzano, C., Adams, S., Rivers, L., Ramos-Pibernus, A., Sevelius, J., Poteat, T., Obedin-Maliver, J., Cicero, E. C., Radix, A., Lett, E., Operario, D., Chapman, B., Wittink, M., Cupertino, A., Murchison, G. R., Nunez-Smith, M., & Cartujano-Barrera, F. (2025). Disentangling gender, sex, and biology: a mixed methods study of gender identity data collection tools. *Patient Education and Counseling, 130*, 108473. Available from: https://pubmed.ncbi.nlm.nih.gov/39471780/

21. Practice Committee of the American Society for Reproductive Medicine. (2024). *Inclusive language and environment to welcome lesbian, gay, bisexual, transgender, queer, questioning, intersex, and asexual+ patients. Fertility and Sterility, 121*(6), 954–960. Available from: https://doi.org/10.1016/j.fertnstert.2024.01.031