

Sexual Orientation, Gender Identity, and Sex Development: Recommendations for Data Collection and Use in Clinical, Research, and Administrative Settings

Background

When considering the use of sexual orientation and gender identity (SOGI) data in your work, there are many nuances in language, culture, generation, and individual identity that are important to consider when tailoring your approach. These guidelines will introduce some key considerations for collecting SOGI data. We recommend viewing the [Foundational Concepts guide](#) as an introduction to the terms and the types of data that may be collected.

Purpose of Data Collection and Ethical Considerations

When collecting SOGI data, “Why” should be at the forefront of your decision-making process. Before reviewing the relevant guidelines, consider your reason for collecting SOGI data, and how it best supports your goals and objectives, as well as LGBTQIA+ people and communities. Given the vulnerability of members of LGBTQIA+ communities to prejudice and discrimination, it is important that data be collected ethically and in the best interest of these communities.

To protect LGBTQIA+ people and communities, it is important to:

- Collect the minimum amount of information necessary to support your goals and objectives.
- Remember that information shared with you may not be information your participants want to share with everyone. Carefully consider your survey population size when asking SOGI data questions. Collecting this information among a small group of participants can accidentally disclose an individual’s LGBTQIA+ status, even when the survey is anonymous.
- Consider that all participants may not want to answer all questions related to their sexual orientation, gender identity, or sex development.

Lesbian, Gay, Bisexual, Transgender, Queer, Intersex, and Asexual + (LGBTQIA+) is an umbrella term referring to some of the identities that exist related to sexual orientation, gender identity, and sex development. The + recognizes that many additional identities exist, and this list is not exhaustive. It is also possible for individuals to hold multiple identities.

Some actions you can take to uphold these principles include:

- Being transparent about whom the information is being shared with, how it is shared, and what security measures are in place.
- Making all questions optional, with response options for participants who choose not to disclose the requested information.
- Assuring participants that no retaliation will occur if they choose not to answer particular questions.

Additionally, providing the best care and service requires appropriate training for those creating, administering, and analyzing these questions, as well as creating a warm, welcoming space for all identities and experiences. Steps to promoting inclusivity include:

- Asking SOGI questions to all participants to avoid making assumptions and introducing bias into your data.
- Welcoming clarifying questions. It is okay to respectfully ask for clarification if there are terms used that are unclear or require additional information.
- Showing a commitment to learning from the experience when a communication misstep occurs. Appropriate first steps include a brief, sincere apology and self-correction, and a willingness to understand more.

Overarching Considerations

When collecting SOGI information, it is important to clearly and correctly identify which aspects of the participant you are inquiring about. Sexual orientation (including identity, attraction, and behavior), gender identity, and sex development are most commonly queried, but which aspects are assessed depends on your goals and purpose.

Keeping your goals and ethical use of SOGI data in mind, you can inquire about:

- **Sexual Orientation:** how a person is emotionally or physically attracted to others
- **Gender Identity:** a person's inner sense of identity as a girl/woman, boy/man, another gender, or absence of gender
- **Sex Development:** a person's physical sex characteristics

It is important to remember that even among these terms there may be ambiguity that may need to be resolved to ensure the correct information is collected. Sex can refer to a participant's biological sexual characteristics, which may be affected by various medicines and surgeries, or can be clarified as "sex assigned at birth," which refers to the sex a doctor assigned on the original birth certificate.

It is additionally important to remember that an individual's gender identity is defined by how that person views themselves, and not by how they appear to the world (their gender expression) or how others perceive that person. As patients and participants disclose their identities to you, they are all valid whether or not they fit any preconceived notion of what someone with that identity looks or acts like.

Formulation of Written Materials

When considering the materials that will be provided to participants, there are key considerations related to language.

It is best to communicate with all participants using gender-inclusive language to avoid making assumptions. This excludes gender-specific identifiers, such as Mr/Mrs/Miss/Ms or Maam/Sir, and uses general inclusive language instead. It is best to identify someone by name or to consider constructions without the use of pronouns (e.g., saying "the participant" or referring to someone by name, as opposed to he/she/they).

Additionally, when talking about particular communities it is important to use person-first language whenever possible (e.g., "transgender people" rather than "transgenders"); "transgender" should be used as an adjective and not a noun (i.e., "trans man" as two words rather than "transman").

When creating forms, surveys, and other written tools, recognize that these materials often reveal historical or personal biases, whether by listing "white" and "male" first in their respective categories or by using stereotypical masculine and feminine colors and imagery. Consider ordering categories by another metric (e.g., alpha-

betical or by prevalence) and using gender-neutral color tones and imagery.

Finally, the formulation of questions listed are themselves a reflection of the current state of understanding and terminology, and represent current recommendations as of January 2022. Over time these questions may need to be expanded or altered as culture continues to shift.

Clinical Guidelines

Engaging with SOGI data as it relates to direct patient care is first and foremost about improving patient care and the patient experience. In the interest of ensuring patients feel seen and heard, it is important to assess the electronic health record (EHR) to determine which SOGI data have already been collected. Modern EHRs should have specific areas dedicated to the collection of SOGI data, though it is possible that these data also appear elsewhere in the chart. When a patient discloses some aspect of their identity to you, be sure to request and receive their explicit permission to enter that data into their EHR.

When a patient's SOGI data are unknown, there should be a process for collecting this information.

The process should:

- Be consistent for all patients, regardless of how staff perceive the patient or what identities they presume for the patient.
- Be collected on electronic and paper registration forms, as well as during the patient interview to validate the information gathered.
- Occur at appropriate times.
 - » SOGI questions should be asked at least once a year at annual visits because patients may have discovered something new about themselves and may identify differently from the time of their last visit.
 - » Information needed to address a patient appropriately, such as pronouns, can be asked at any time when the information is unknown.
 - » Information related to the patient's gender identity, gender expression, history of gender-affirming medical care, plan to obtain gender-affirming medical care, or other related sexual health history can be discussed when clinically indicated, but should not be asked unless directly relevant to care at the time of the visit.
- Collect data for multiple identities and aspects, including name used in the current setting; pronouns; gender identity; sex assigned at birth; name and gender marker on identity documents; and for insurance related queries, the name, legal name (if different), and gender marker on insurance documents.

In addition to having a process for collecting SOGI data, it is important to create an environment welcoming to LGBTQIA+ communities. Staff should be trained on where to access and document SOGI

Recording information over time

The information patients disclose about their identities may change over time as a result of their evolving understanding of their own identity, changing identity, or comfort with disclosure. It can be helpful to collect SOGI information more than once to capture these changes. In addition to positive affirmation, it is important to obtain explicit permission before including such information in the patient's chart, and be transparent as to who will have access to this information and for what purpose.

data, and be prepared to answer patient concerns and questions about why the data is being collected, who will have access to the data, and how it will be used.

Age Considerations: Data collection should be adjusted as developmentally appropriate. Gender identity evolves through life but can solidify as early as two years old. Generally, individuals are aware of sexual orientation from puberty onward, but sexual orientation may also change over time. It is important to remember that children may not feel comfortable answering SOGI questions in front of a parent or guardian.

Guidelines for Research Surveys

When collecting information via surveys for research purposes, it is important to review your goals and objectives and the ethical considerations described previously to ensure the information is being collected appropriately.

When querying gender identity, one question is not enough to accurately capture diverse gender identities. Discrepancies often arise from the fact that a portion of individuals who may be categorized as transgender identify with a singular gender (e.g., woman or man), but not as transgender. To more accurately detect the nuances of gender identity, both a two-question and three-question model have been proposed.

Question 1: What best describes your current gender identity? (Select one [or select all that apply])

- A. Woman
- B. Man
- C. Nonbinary (e.g., genderqueer, gender non-conforming)
- D. If not listed [please specify]
- E. Prefer not to answer

Question 2: On your original birth certificate, was your sex assigned female or male? (Select one)

- A. Female
- B. Male
- C. Prefer not to answer

This two-question model allows for additional analysis to occur, comparing a participant's sex assigned at birth with their gender identity to understand which category a participant falls into for the purpose of your study. The gender identity question (Question 2) can be either "select one" or "select all that apply" depending on the goals of the research and the capacity to handle multiple responses in analyses of survey data. The write-in option for another gender is an important opportunity to affirm participants' individual identities.

There are validation issues with the two-question model, and research is currently trending towards a three-question model, which includes the two previous questions and:

Question 3: Do you identify as transgender? (Select one)

- A. No, I do not identify as transgender
- B. Yes, I do identify as transgender
- C. I am not sure if I am transgender
- D. I do not know what this question is asking
- E. Prefer not to answer

Assessing transgender identity using a separate item allows for researchers to accurately code whether a participant identifies as transgender, while also allowing participants to name their current gender identity separate from transgender status, as applicable.

If querying sex development, some researchers will want to understand whether a patient was born as "intersex." As intersex is not a gender identity, it should not be included in the previously described questions and should be assessed in a separate item if this information is important for your study aim. Since all survey participants may not understand the term "intersex," the question should include a definition of the term. For example:

Recommended Intersex Question:

Intersex is a term for people born with variations in physical sex characteristics. There are many different intersex variations. Do you have an intersex variation?

- A. Yes, I have an intersex variation
- B. No, I do not have an intersex variation
- C. I am not sure if I have an intersex variation
- D. I do not know what this question is asking
- E. Prefer not to answer

When querying sexual orientation, the specific dimension assessed (identity, attraction, or behavior) depends on the goals and objectives of the survey. In some cases, it may make sense to assess multiple sexual orientation dimensions.

Sexual orientation identity is the label an individual uses to denote the emotional or physical attraction they feel for others. As identity labels continue to evolve, it is important to include commonly used identity labels in the response options, while being mindful of not creating confusion by using terms that are not well known.

Recommended Sexual Orientation Identity Question:

Which of the following best describes your current sexual orientation? (Select one [or select all that apply])

- A. Straight/heterosexual
- B. Bisexual
- C. Gay or Lesbian
- D. Pansexual
- E. Queer
- F. Asexual
- G. If not listed above [please specify]
- H. Prefer not to answer

Individuals may identify with one or more than one of these labels, and understanding of the terms may vary by population. This question can either be asked as "select one" or "select all that apply" depending on the study goals and analysis.

Sexual orientation identity only refers to how a participant identifies. It does not provide information on the participant's sexual behaviors or attractions.

If your study is focused on sexual behaviors (e.g., HIV transmission rates) then it would be more appropriate to assess the behavior dimension of sexual orientation rather than identity. Often, more than one question is needed to accurately understand both the types of sexual behaviors engaged in, as well as the gender identities of the sexual partners. This is a budding area of research, and no current recommendations exist for best practices. When considering the wording for these questions, it is important to remember what information is important for your research.

Additionally, appropriate use of the terms "gender" and "sex" and clearly defining key terms such as

“intercourse” will be key to ensuring that participants understand what is being queried, and for you to obtain the information needed for your study.

Attraction refers to sexual or romantic feelings towards another person, which can change through the course of someone’s life. A question that could be asked regarding attraction is:

Recommended Attractions Question:

Who are you currently attracted to? (Select all that apply)

- A. Women
- B. Men
- C. Nonbinary people (e.g., genderqueer, gender non-conforming)
- D. People of another gender identity [please specify]
- E. None of the above
- F. Prefer not to answer

Administrative Surveys

There are a variety of reasons why organizations may wish to request and use SOGI information for administrative purposes. A few examples are:

- Climate surveys to assess member experiences, such as perceptions of safety, support, and satisfaction.
- Matching roommates in college student housing assignments of multi-occupancy rooms.
- During a selection process when the aim is to have a diverse cohort for a particular educational, community service, or other program.

One common administrative situation where issues related to gender identity can arise is during the process of hiring employees. Currently, some US states only have female and male as legal identities, whereas others have begun to include other legal gender markers (e.g., X). Furthermore, when collecting data on behalf of applicants and new hires, the gender marker collected for official state or federal purposes may differ from the identity applicants wish to use and have affirmed with their institution. If your institution offers health benefits, the gender marker associated with an individual’s health insurance may also differ.

Whatever the purpose, keep in mind that collecting such information will help to enrich the experience of LGBTQIA+ individuals in your organization.

Be mindful to ensure that the SOGI information being asked is appropriate and limited only to what is needed to reach your objectives. For example, asking questions about physical sex characteristics or about legal identity in a climate survey would be intrusive and irrelevant. Finally, when collecting SOGI data for administrative surveys, there may be legal considerations involved. It is your responsibility to seek and consult with those who have expertise in this area such as your organization’s legal counsel, human resources professionals, or others. Some questions that might be relevant to your administrative objectives for surveys or other data collection needs are:

What best describes your current gender identity? (Select one)

- A. Woman
- B. Man
- C. Nonbinary (e.g., genderqueer, gender non-conforming)
- D. If not listed above [please specify]
- E. Prefer not to answer

Do you identify as transgender? (Select one)

- A. No, I do not identify as transgender
- B. Yes, I do identify as transgender
- C. I am not sure if I am transgender
- D. I do not know what this question is asking
- E. Prefer not to answer

Which of the following best describes your current sexual orientation? (Select one)

- A. Straight/heterosexual
- B. Bisexual
- C. Gay or Lesbian
- D. Pansexual
- E. Queer
- F. Asexual
- G. If not listed above [please specify]

Additional questions for legal reporting requirements and/or insurance purposes:

What is the sex marker on record with the United States Social Security Administration? (Select one)

- A. Female (F)
- B. Male (M)

What sex marker would you like listed on your insurance card? (Select one). If “Prefer Not to Answer” is selected, you may be contacted prior to enrollment in an insurance plan.

- A. Female (F)
- B. Male (M)
- C. Prefer not to answer