Best Practices Guidance for Sexual Orientation and Gender Identity Data Collection

June 2025

Table of Contents

The Importance of SOGI Demographic Data Collection

The routine collection, analysis, and sharing of sexual orientation and gender identity (SOGI) by the California Department of Public Health (CDPH) is essential to the health and well-being of the lesbian, gay, bisexual, transgender, queer, intersex, asexual, and other identities (LGBTQIA+) community. More specifically, this data helps identify disparities, inform public health strategies, and support health equity efforts across the state. The collection of SOGI data begins at the community level and is used by Local Health Jurisdictions (LHJs) and other public health partners to support regional efforts. CDPH programs collect data from a variety of sources (including directly from individuals, from administrative data sources, or from third parties such as healthcare or external surveys) that may affect implementation of these guidelines. For a detailed summary of these sources, see Appendix A. At an individual level, asking about SOGI status is a critical part of establishing a good client-provider relationship, and can promote client-centered services. At the population level, collection of SOGI data allows public health to better understand our communities; identify gaps, inequities, and disparities; design and deliver more responsive services; and pursue policy, systems, and environmental changes that advance health equity.

How This Guidance Was Developed

This guidance document is provided as **Best Practices Guide** for California Department of Public Health (CDPH) programs. The foundation of the CDPH SOGI Guidance is based on best practices recommended by the National Academies of Sciences, Engineering, and Medicine; the National LGBTQIA+ Health Education Center at the Fenway Institute; and UCLA Williams Institute. All three institutions endorse the three-part SOGI question format, which includes a question on sexual orientation and a two-part question on gender identity consisting of current gender identity and sex assigned at birth on original birth certificate.

These recommendations informed the formation of the CDPH SOGI Workgroup, which produced the 2022 edition of the CDPH SOGI Data Collection Guidance. CDPH subsequently received feedback from the LGBTQIA+ community, particularly during the mpox outbreak, calling for changes such as the separation of the transgender category into “transgender man” and “transgender woman.” Additional feedback was also provided by the California Conference of Local Health Officers (CCLHO) and the County Health Executives Association of California (CHEAC).

In response to this feedback, CDPH reconvened the SOGI Workgroup to develop the 2025 update of the SOGI guidance document. This update expands upon the original guidance and includes new recommendations for data display. During the process of updating the document, CDPH actively solicited and incorporated feedback from internal programs and community partners.

The most significant changes in the 2025 update include the addition of new response categories and a recommendation to allow respondents to “check all that apply”. These

revisions better reflect the complexity and fluidity of LGBTQIA+ identities and honors communities' rights to self-identification and self-determination. State SOGI Data Collection Laws and Regulations
Listed below are multiple California laws that serves as a foundational support for the collection, analysis, and use of SOGI data.

- Effective January 1, 2016, Assembly Bill (AB) 959 (Chapter 565, Statues of 2015) amended California Government Code 8310.8 to require CDPH to include SOGI data on forms that collect ancestry or ethnic origin data.
- Effective September 26, 2020, Senate Bill (SB) 932 (Chapter 183, Statutes of 2022) amended Health and Safety Code (HSC) 120255 to require CDPH to include SOGI fields in its electronic tools for data collection on reportable diseases. It also requires healthcare providers to include SOGI data, if known, in reports to the local health officer for the jurisdiction where the patient resides.
- Effective January 1, 2025, Senate Bill 957 (Chapter 868, Statutes of 2024) expands California's requirements for collecting sexual orientation and gender identity (SOGI) data. Key provisions include:
    - Government Code Section 8310.8: Amended to require the California Department of Public Health (CDPH) to collect SOGI data from third-party sources, including local health jurisdictions, on any forms or electronic systems, unless prohibited by law.
    - Health and Safety Code Section 120440: Amended to mandate that healthcare providers report SOGI data (including sexual orientation, gender identity, and sex assigned at birth) for adults aged 18 and older to local health departments and CDPH, particularly through systems like the California Immunization Registry (CAIR), when immunization forms also collect ancestry or ethnic origin
    - New Health and Safety Code Section 15200: Added to require CDPH to submit an annual public report to the Legislature detailing its efforts to collect, analyze, and use SOGI data to improve services and outcomes for LGBTQ+ communities.

CDPH SOGI Demographic Data Collection Standards

This guide outlines best practices for collecting SOGI demographic data for all CDPH programs that handle demographic data. The goal is to promote compliance with SOGI data laws while being respectful, accurate and inclusive of LGBTQIA+ communities.

This guidance reflects the balance between scientific data collection methodologies and the fluidity of cultural understandings and language related to SOGI identities. The guidance follows best practices from LGBTQIA+ advocacy institutions such as the Fenway Institute, the National Academies of Sciences, Engineering, and Medicine (NASEM), and the California Health Interview Survey (which partners with the Williams Institute). These

institutes agree that current best practice to assess SOGI status is asking a three-part question of sexual orientation, gender identity, and sex assigned at birth. To promote universal adoption and ease of implementation, the guidance also incorporates feedback from CDPH LGBTQIA+ subject matter experts with lived and professional experience, CDPH programs, and Local Health Jurisdictions potentially affected by CDPH's adoption of SOGI data collection standards.

Based on best practices from national LGBTQIA+ advocacy organizations, it is recommended that CDPH forms collecting SOGI demographic data should include all three questions listed on page 6 and 7, with the 2-step gender identity question in the order listed, unless there are practical or logical reasons not to do so (please see below for acceptable reasons to deviate from this best practice)**.** The two-step question of gender identity then sex assigned at birth (SAAB) has been established as a best practice ever since it was proposed in 2009 by what the UCSF Center of Excellence for Transgender Health is now. It has continued to be stated as a best practice by leading LGBTQIA+ research organizations such as the National Academies for Science, Engineering, and Medicine, the National LGBTQIA+ Health Education Center of the Fenway Institute, and the UCLA Williams Institute. As per the National Academies of Sciences, Engineering, and Medicine report "Measuring Sex, Gender Identity, and Sexual Orientation" issued in 2022:

Asking respondents to separately identify their sex and their gender—in particular, their sex assigned at birth and gender identity—improves overall measurement quality and also allows researchers and other data users to identify individuals with transgender experience by comparing their sex assigned at birth to their current gender identity. This two-step gender measure has become an increasingly common and validated way to identify people with transgender experience because it identifies a wider range of transgender people than single-step methods that ask respondents whether they identify as transgender.

Please note that classifying individuals as "transgender" who select "man" or "woman" for gender identity while indicating the opposite sex for SAAB is for the purposes of aggregated data analysis only. Individuals should always be addressed according to their stated gender identity. However, for epidemiological studies, it is vital to capture the transgender population as fully as possible. For example, in a study on health inequities for the transgender population, classifying these individuals as "man" or "woman" without taking SAAB into account effectively reclassifies them as cisgender, possibly compromising the study results.

In some cases, there may be a structural reason in which the questions are not relevant; for example, surveys for infants/children would not ask about sexual orientation, and California law prohibits providers from disclosing this information for inclusion in local or state immunization registries if it is collected. The Fenway Institute guidelines say that gender identity questions can be asked around age 3 and sexual orientation asked at ages 10 to 13. However, public health programs should consider the risks and benefits of collecting SOGI data with minors as California statute does not provide any privacy for

minors less than 12 years of age and parents have a right to answer SOGI questions on their behalf. Individuals who are 12+ years of age and identify as LGBTQIA+ may not yet feel comfortable sharing this information with their parents or guardians, especially if this information could be accessible through a shared medical record or evidence of insurance benefits.

Sexual Orientation

1. Which of the following best describes your sexual orientation or sexual identity? (Select all that apply):
    a. Lesbian
    b. Gay
    c. Straight
    d. Bisexual
    e. Two- Spirit (a term used by some American Indians and Alaska Native individuals)
    f. Same Gender Loving (a term used by some Black or African American individuals)
    g. Asexual
    h. Pansexual
    i. Queer
    j. Questioning/Unsure
    k. Use another term (specify): _____
    l. Don't understand the question
    m. *(Interview only)* Respondent preferred not to answer

Gender Identity

2. What is your current gender identity? (Select all that apply)
    a. Woman/Female
    b. Man/Male
    c. Transgender
    c. Transgender Woman/Female
    d. Transgender Man/Male
    e. Two- Spirit (a term used by some American Indian/and Alaska Native individuals)
    f. Genderqueer
    g. Non-Binary
    h. Agender
    i. Gender Fluid/Diverse
    i. Questioning/Unsure
    j. Use another term (specify): _____
    k. Don't understand the question
    l. *(Interview only)* Respondent preferred not to answer

Sex Assigned at Birth

3. What sex were you assigned at birth (i.e., what would be listed on your original birth certificate)? (Select one)
    a. Female
    b. Male
    c. My sex was not listed on my original birth certificate
    d. I'm not sure
    e. *(Interview only)* Respondent preferred not to answer

How to Ask Questions About SOGI

- Ask current gender identity and sex assigned at birth separately. This provides an opportunity for transgender and gender non-conforming individuals who may not otherwise select those options to self-identify. Remember to ask the sex assigned at birth second.
- Specify that sex assigned at birth refers to the person's original birth certificate.
- Whether in an oral interview, paper form, or online form, it should always be made clear that answering these questions are voluntary. For online surveys, there should always be an option to skip these questions.
- "Respondent preferred not to answer" should only be an option when an interviewer is asking the question, as providing this option on a written form has been shown to result in decreased data quality.

Allow for diverse measures of sexual orientation and gender identity in response options

- Offer "Check all that apply" options, wherever possible. CDPH recognizes that encouraging and enabling respondents to "Check all that apply" may make data analysis more difficult. In making this decision to support a wider range of responses, CDPH embraces core values of inclusivity and the right of self-determination as we work toward creating more equitable systems and policies that allow us to honor those values and the needs of our diverse communities.
- Data systems are often not built to accommodate more than one response; *if your system does not allow for multiple responses, clearly ask respondents to "*select the option that best describes your identity." This allows the individual to self-identify as best as possible, rather than data entry personnel assuming which answer is to be recorded.
- Include culture specific gender identities: two-spirit and same gender loving. In the responses section for the question, explain that 'two-spirit' is a term used by some Native American or Alaska Native individuals and that 'same gender loving' is a term used by some Black or African American individuals.
- When using write-in options, do not use the terms "other" or "different" or define people as being outside of the dominant group (such as "not straight") as these can be alienating and othering to individuals who identify outside of the provided answer options.
- Ensure appropriate privacy protections of SOGI information collected for all individuals.

Sample Scripts

Some individuals may not understand why their provider or the LHJ is asking about SOGI, and some healthcare providers or public health personnel may be unsure how to ask these questions. The following sample scripts provide an example on how we can ensure that those asking the questions and those responding to them understand the context in which the information is being collected.

Explain Why You are Asking

Provide individuals with clear guidelines for how and why the collected information will be used. This can help encourage comprehensive and accurate data gathering. Sample phrasing:

- "I have a few simple questions to ask you about yourself. I am gathering this information to ensure that you are receiving the best quality of service."
- "We ask all our clients about their sexual orientation and gender identity so we can make sure that you get the service you need. This is meant to be a comfortable and safe environment, where you can show up as your whole self."
- "This information helps us better understand the communities that we serve. It helps us meet individuals' needs and preferences and provide the best service possible."
- "Asking for your sex assigned at birth helps staff ask the right questions and provide effective service based on the needs of the individual."

Note That Questions are Voluntary

Let individuals know that answering these questions is voluntary. They do not need to answer questions if they do not want to.

- "You may decline to answer any of these questions. Any response you provide will help us to better serve you."

Explain How Data Will Be Protected

Let individuals know their information will be kept confidential. If it will be shared, let them know how and in which context.

- "We protect all data collected to the greatest extent permissible under the law. We want to provide a safe environment and provide the highest level of service to meet your specific needs."
- "Test results for some conditions will be shared with the local health department so they can help us make sure you get the care you need. The local health department is required to keep your information confidential as well."

Define Terms if Clients are Unsure What They Mean

- Explain in plain language terms, understanding that not all terms will translate exactly in every language.

- "Sex assigned at birth means the sex (such as female or male) listed on your original birth certificate when you were a baby."
- "Sexual orientation describes your physical and emotional attraction to others."
- "Gender identity is your internal sense of your own gender, regardless of the sex you were assigned at birth. Gender identity is personal to each individual."

SOGI Data Display

The following section provides guidance and considerations for the *display* of SOGI data once it has been collected.

Combining Categories When Displaying SOGI Data

Because individuals use terms that they identify with, asking how someone refers to themselves and reflecting that back are best practices. The ideal standard for displaying SOGI categories is to display all of the categories for which data was collected, including those with no responses. As long as privacy is protected per the California Health and Human Services Data Deidentification Guidelines (DDG) (and all other applicable law), presenting this level of granularity helps promote data transparency and the visibility of individual identities within the LGBTQIA+ community while addressing data privacy concerns. However, there may be a need to combine more granular categories into larger ones, such as:

- Counts that are too small to analyze or publicly display, especially when considering unique combinations that occur when individuals select multiple categories in the data collection questions.
- Need to bridge more recent datasets (which tend to be more granular) to older datasets (which tend to be less granular) to do trend analysis.
- Need to match data to less granular denominator datasets to calculate rates or other denominator-based metrics.

If combining categories is required, refer to Tables 1 and 2, which show how the most granular categories on the left are combined into the least granular categories on the right. When combining smaller categories into larger categories, avoid duplication within the larger category. For example, if someone identifies as both genderqueer and non-binary, do not double count them in the "Another Gender Identity" larger category.

The most accurate method to capture an individual's identity is to assign each combination to its own unique category. For example, someone who identifies as both genderqueer and transgender would be placed in a "Genderqueer and transgender" category. However, this method would result in very small category numbers for many data sets, leading to both re-identification risk and insufficient power for statistical analysis. The aggregation of more granular categories into less granular categories should also be

used for reclassifying an individual who selects multiple categories into a single category as displayed in tables 1 and 2.

Alternatively, if there are combinations that are not on the below table or the resulting numbers are still too small to display, one can combine all combinations into a "More than one category selected" category. Data reports should clearly explain how the data were collected and combined to provide transparency to readers. This should include a rationale for collapsing any category.

Not every variable will have a 100% completion rate. Therefore, it might be necessary to consider both gender identity and sex assigned at birth when combining categories into broader groups.

Table 1: Combining Sexual Orientation Categories

Sexual Orientation	Gender ID	Sex assigned at Birth	Coded As Sexual Orientation #1	Coded As Sexual Orientation #2
Straight	All Gender ID's	All options	**Straight**	**Straight**
Lesbian/Gay/Two-Spirit/Same Gender Loving/Queer	Female/Woman and Transgender Female/Woman	All options (if gender id is not missing) Female (if Gender ID is missing)	**Lesbian**	**Lesbian, Gay or Bisexual**
Lesbian	All Gender ID's or missing	All options		
Gay/Two-Spirit/Same Gender Loving/Queer	Male/Man and Transgender Male/Man	All options (if gender id is not missing) Male (if Gender ID is missing)	**Gay**	
Gay	Male/Man and Transgender Male/Man**,** Two-Spirit, Genderqueer, Agender, Non-binary or missing	All options		
Bisexual	All Gender ID's or missing	All options	**Bisexual**	

Asexual	All Gender IDs, or missing	All options	**Another Sexual Orientation**	**Another Sexual Orientation**
Pansexual	All Gender IDs, or missing	All options		
Use Another Term	All Gender IDs, or missing	All options		

Consider a dataset with 5 individuals who selected Woman/Female as gender identity and Gay for sexual orientation, 3 who selected Lesbian and have missing information on their gender ID, 1 who selected Transgender Man/Male and Gay, and 1 who selected genderqueer and Gay and 10 that selected straight and man/male and genderqueer as gender identity . According to the table above, this dataset could be categorized as 10 straight, 8 lesbian and 2 gay individuals. Based on the small number rule this would be combined into 10 straight and 10 Lesbian or Gay individuals.

Table 2: Combining Gender Identity Categories

Selected Gender ID	**Selected Sex assigned at Birth**	**Coded As Gender ID (more detailed)**	**Coded As Gender ID (less detailed)**
Woman/Female	Woman/Female, Unknown or Missing	Woman/Female	**Cisgender**
Man/Male	Man/Male, Unknown or Missing	Man/Male	
Woman/Female	Man/Male	Transgender Woman/Female	**Transgender**
Man/Male	Woman/Female	Transgender Man/Male	
Transgender (Selected with or without Woman/Female, Man/Male Non-binary, Two-Spirit, Genderqueer or Use another term)	Woman/Female, Man/Male, Unknown or Missing	Transgender	
Transgender Woman/Female (Selected with or without Man/Male, Non-binary, Two-Spirit, Genderqueer or Use another term)	Man/Male, Unknown or Missing	Transgender Woman/Female	

Transgender Man/Male (Selected with or without Woman/Female, Non-binary, Two-Spirit, Genderqueer or Use another term)	Woman/Female, Unknown or Missing	Transgender Man/Male	
Two- Spirit	Male, Female, Unknown or Missing	Two- Spirit	**Another Gender Identity (or Transgender)**
Genderqueer	Male, Female, Unknown or Missing	Genderqueer	
Non-Binary	Male, Female, Unknown or Missing	Non-Binary	
Agender	Male, Female, Unknown or Missing	Agender (can be combined with non-binary or Transgender)	
Use another term	Male, Female, Unknown or Missing	Use another term	**Another Gender Identity**
Questioning/Unsure	Male, Female, Unknown or Missing	Questioning/Unsure	

Consider a dataset with 10 individuals who selected Female for both Gender Identity and Sex Assigned at Birth, 3 who selected Female for Gender Identity and Male for Sex Assigned at Birth, 1 who selected Transgender, and 1 who selected Genderqueer. According to the table above, this dataset could be categorized as 10 Cisgender, 4Transgender, and 1 Another Gender Identity, or alternatively as 10 Cisgender and 5 Transgender or another gender identity.

Using the Two-Step Gender Identity Question to Inform Data Display

The two-question format to define a person's gender identity ("current gender identity" and "sex assigned at birth") in the SOGI Data Collection Guidance helps capture both individuals who self-identify as transgender and those who identify as man/male or woman/female and were assigned a different sex on their original birth certificate. Because there are multiple possible responses to each part of this two-part question, Table 2 provides guidance on how to interpret various possible combinations of answers.

Protecting the Privacy and Confidentiality of SOGI Data

Protecting the privacy of individuals and preventing the re-identification of individuals is of the utmost importance to CDPH. Therefore, CDPH requires the application of the California Health and Human Services Agency (CalHHS) Data De-Identification Guidelines (DDG) to all publicly disclosed data containing demographic information from departments under its purview. This is a necessary measure to prevent inadvertent re-identification of individuals within small data cells, thereby ensuring compliance with privacy protection requirements outlined in the California Information Practices Act (IPA) and the Health Insurance Portability and Accountability Act (HIPAA). It is especially important in the

context of SOGI data to not inadvertently disclose an individual's SOGI status against their will. Note that the DDG is only mandatory for CalHHS Departments—local health jurisdictions may use different guidelines to protect data privacy, so long as they comply with IPA and HIPAA.

It is important to note the DDG does not prevent the display of SOGI data, but only masks numbers in certain situations. The DDG employs a risk assessment algorithm to evaluate the likelihood of re-identification on an individual level for each dataset. If the risk score is above a certain threshold, the DDG mandates that data counts below 11 (or metrics derived from counts, such as rates) are either suppressed or granularity of the data decreased (e.g., combining categories or years) until the risk score falls below the threshold. Data counts above 11 (or metrics derived from such counts, such as rates) are not restricted by the DDG and can be displayed.

Note that data that has been suppressed is different than data that should be interpreted with caution due to statistical instability. While the former will be redacted, the latter may not. As examples, see the California Health Interview Survey document on how the two are handled differently, and the Excel tables on the CDPH County Health Status Profiles website which also differentiate between the two.

SOGI Denominator Data Sources

As of 2024, the United States Census, American Community Survey, and California Department of Finance do not collect SOGI data. Therefore, denominator data (meaning the number of people with a given demographic in the overall population of a geographic area) must be estimated from other surveys. A CDPH taskforce evaluated three surveys as the most suitable candidates for providing denominator data: The California Health Interview Survey, the U.S. Census Bureau Household Pulse Survey (HPS), and the Behavioral Risk Factor Surveillance System (BRFSS). For a detailed assessment of the pros and cons of using each survey, the SOGI categories they provide, and how to obtain the data, please refer to the Sexual Orientation and Gender Identity Data Sources Document.

Be aware that estimates of California's LGBTQIA+ population size can vary widely and likely underestimate the actual number. Additionally, younger respondents are more likely to identify as LGBTQIA+. Each successive generation is roughly twice as likely as the previous one to identify as "LGBTQ+", with over 20% of Gen Z adults identifying as such based on polling results from 2023 (LGBTQ+ Identification in U.S. Rises to 9.3%LGBTQ+ Identification in U.S. Rises to 9.3%). .

Note: that more granular denominator estimates for non-binary categories such as genderqueer, and Two-Spirit, are usually unavailable. In such cases, only numerator counts can be displayed.

Improving Completeness of SOGI information and Displaying Data with High Proportion of Incompleteness (Missing or Unknown Values)

An assessment of SOGI data collection practices across CDPH programs conducted between 2022 and 2024 found that a large proportion of the data collected by CDPH programs does not include SOGI status. CDPH is required to display SOGI-related metrics to advance health equity, but this becomes difficult if a large proportion of SOGI data is missing and runs the risk of inaccurate and/or misleading conclusions being drawn from incomplete data. For programs dependent on third party data (e.g., EHRs) or conditions that are not investigated by local health departments (e.g., cancer), highlighting this fact may help increase understanding in the community among reasons for data missingness and improve reporting.

Another area in need of evaluation and improvement is proportion of respondents who selected not to answer. If response rates indicate potential bias, nonresponse bias analyses should be conducted, with weighting adjustments and other techniques used as needed. The results should be reported to provide data to users, so they understand the data's limitations when interpreting it. Table 3 offers an example of how to display missing SOGI data as a separate category.

Table 3: Sample CDPH Data Table: Cumulative Mpox Cases by Gender*

Gender	Persons[1]	Percent[2]
Male	5,903	95.6
Female	151	2.4
Transgender Female	87	1.4
Transgender Male	19	0.3
Genderqueer/Non-Binary	16	0.3
Unknown	88	-

* Data collection system did not include options for Two-spirit when these data were collected; this option has since been added. 1 As sex assigned at birth was often not reported, counts of cases in cisgender categories could not be reliably distinguished among the "Male" and "Female" categories of data.
2 Percentages are calculated with the denominator excluding unknowns. Source: https://www.cdph.ca.gov/Programs/CID/DCDC/Pages/Mpox-Data.aspx
For further guidance, please see the Federal Committee on Statistical Methodology's paper "A Framework for Data Quality", specifically pages 71-72.

Cross-tabulating sexual orientation and gender identity

When appropriate, data on gender and sexual orientation should be cross tabulated to account for the nuances and intersections between these identities. Using the data from the Mpox study referenced in Table 3, a sample data table for individuals identifying as

female is provided in Table 4. Similar tables should be created for individuals of other gender identities. Please note that this example includes terminology—such as 'Heterosexual'—that is not recommended by current guidance.

Table 4: Sample Cross-Tabulated Data Table: Cumulative Mpox Cases within Females by Sexual Orientation

Female	Persons[1]	Percent[2]
Total	151	2.5
Bisexual	12	10.5
Gay, Lesbian, or Same-Gender Loving	4	3.5
Heterosexual/Straight	98	86
Other	0	0
Unknown	37	-

* Two-spirit was not among the options at the time data was collected; this option has since been added.
[1] As sex assigned at birth was often not reported, counts of cases in cisgender categories could not be reliably distinguished among the "Male" and "Female" categories of data.
[2] Percentages are calculated with the denominator excluding unknowns.

References

En " LGBTQIA+ health education center. LGBTQIA+ Health Education Center. (2024, August 26). https://www.lgbtqiahealtheducation.org/glossary/en/

UCSF Center of Excellence for Transgender Health. 2009. *Recommendations for Inclusive Data Collection of Trans People in HIV Prevention, Care & Services*. https://prevention.ucsf.edu/transhealth/education/data-recs-long

The GenIUSS Group. 2014. *Best Practices for Asking Questions to Identify Transgender and Other Gender Minority Respondents on Population-Based Surveys*. J.L. Herman (Ed.). Los Angeles, CA: The Williams Institute. https://williamsinstitute.law.ucla.edu/publications/geniuss-trans-pop-based-survey/

Glossary of terms. HRC. (n.d.). https://www.hrc.org/resources/glossary-of-terms

Glossary of terms: LGBTQ. GLAAD. (2023, May 1). https://glaad.org/reference/terms

National Academies of Sciences, Engineering, and Medicine. 2022. *Measuring Sex, Gender Identity, and Sexual Orientation*. Washington, DC: The National Academies Press. https://doi.org/10.17226/26424.

National LGBTQIA+ Health Education Center: A Program of the Fenway Institute. 2022. *READY, SET, GO! A Guide for Collecting Data on Sexual Orientation and Gender Identity.* Boston, MA. https://www.lgbtqiahealtheducation.org/publication/ready-set-go-a-guide-for-collecting-data-on-sexual-orientation-and-gender-identity-2022-update/

Appendix A: SOGI Data Source Strengths and Challenges

SOGI demographic data is collected by LHJs and CDPH from a variety of sources. Depending on the reporting sources, SOGI data may be complete, incomplete, or missing altogether. The table below outlines the various types of sources from which the department receives SOGI data. These include individual provider reports, LHJ case investigations, interviews, electronic health records (EHRs), individual self-report (program survey is conducted outside of clinical settings), lab reports, and federally mandated forms. The table also highlights the specific strengths of each collection method, and the challenges associated with these methods.

TABLE 6: SOGI Data Sources, Strengths, and Challenges

Source	Strengths	Challenges
Laboratory reports	CDPH maintains a list of laboratory reportable conditions. Laboratories report test results electronically into the California Reportable Diseases Information Exchange (CalREDIE). CalREDIE is a secure database that CDPH and LHJs can access SOGI data fields in the patient tab for all diseases and conditions.	Laboratories typically do not collect SOGI data on their paper laboratory requisition forms or online platforms that providers use to order laboratory tests. If laboratories do not provide options to enter SOGI information on their forms, then they cannot send it to LHJs or CDPH via CalREDIE. Laboratories are not required to collect SOGI data per statute.
Individual provider reports	CDPH maintains a list of provider reportable conditions, which includes a list of required variables, including name, address, telephone number, date of birth, race, ethnicity, and SOGI. Providers submit case reports to their local health department on the Confidential Morbidity Report (CMR) form, which includes SOGI variables. LHJs review these reports and enter them into CalREDIE or, for some LHJs that do not use CalREDIE, their local data systems. Non-CalREDIE LHJs then submit this data electronically to CDPH.	Despite state regulations requiring providers to report SOGI data, and the inclusion of SOGI demographic data fields on the CMR form, health care providers may not understand SOGI terminology and often do not feel comfortable asking their patients about SOGI. Even if providers do ask, they may not include this information in the CMR. If providers do not collect or report SOGI information, LHJs may not receive it and are thus unable to send it to CDPH. Provider reporting is also incomplete; a provider report is not submitted for all incidents reported by laboratories.

LHJ provider interviews	LHJs review incoming CMRs, and laboratory reports to identify high priority cases where urgent public health action is needed to interrupt disease transmission or assure patients receive appropriate services. As part of case investigations, LHJs can contact the provider to ask for the patient's SOGI data. If the provider has it then the LHJ can report this information to CDPH using the SOGI fields in CalREDIE or CalConnect, a data system developed for COVID-19 and adapted for tracking other diseases and conditions.	LHJs receive thousands of CMRs, and laboratory reports each year, sometimes tens of thousands in large LHJs. In most LHJs, it would be impossible to follow up on every single reported disease or condition case to contact the provider (or the patient) to collect SOGI data. LHJs prioritize the cases with the greatest potential for intervention to prevent transmission and/or assure timely care. This is especially true when case volumes increase and during public health emergencies.
LHJ patient interviews	For high priority diseases and conditions, such as mpox, syphilis, and HIV, LHJs may contact individuals to collect information and assure services/care. When conducting these case investigations, LHJs can establish trust with the individual and ask them about their SOGI. Establishing this trust helps people feel comfortable sharing their information. For conditions where LHJs routinely conduct case investigations, such as the high priority conditions listed above, SOGI data variables are collected routinely.	Patient interviews are time intensive and prioritized for only high priority diseases conditions. LHJs may have difficulty locating and interviewing some individuals, especially people who are unstably housed or transitory. Individuals also may not feel comfortable disclosing SOGI data for conditions not related to sexual health, such as childhood lead poisoning. Cultural and linguistic differences based on race/ethnicity, age, and subculture result in lack of clarity about the questions being asked.

Electronic health records (EHRs)	Some healthcare providers, such as those in health clinics that specialize in serving LBGTQIA+ communities, routinely ask their patients about SOGI and collect this information in their EHR. If the LHJ has access to the EHR, they may be able to collect this information when conducting a case investigation and are able to record it in CalREDIE or their local data systems to share with CDPH. CDPH is starting to roll out electronic transmission of certain data elements directly from EHRs into CalREDIE for select reportable diseases and conditions, known as electronic case reporting (eCR). Data collection systems used by CDPH outside of CalREDIE (federal enterprise systems and others) are similarly being modernized to leverage eCR. This has the potential to greatly increase SOGI data completeness.	EHR vendors use national standardized data elements established by the federal U.S. Department of Health and Human Services (HHS), Office of the National Coordinator for Health Information Technology (ONC), United States Core Data for Interoperability (USCDI). There are multiple versions of SOGI standards from USCDI. Until January 2026, EHR systems can still adhere to USCDI version 1, which does not contain any SOGI standards. By January 2026, EHR systems will be required to update to USCDI version 3, which does include SOGI standards. Until then, programs that collect data from EHR systems may be missing SOGI data from a substantial proportion of their records due to these circumstances. Some national standard setting organizations are still in the process of updating to include LGBTQIA+ options.
Patient self-reporting	Some individuals may feel more comfortable answering a survey than answering a question from a provider or LHJ.	Some individuals may not complete the SOGI questions if they do not know how it will be used, do not think the question is relevant to the situation, or fear identification and discrimination.
Program surveys (outside of a clinical setting)	Many public health programs conduct community surveys outside of a clinical setting. Programs using surveys can explain why the question is being asked and how the data will be used to improve services.	Not all public health programs address sexual health or may be familiar with asking SOGI questions. Individuals may not understand why these questions are being asked.
Federally mandated forms	Depending on their guidelines for data collection, federally funded projects may require specific forms with SOGI variables to be completed.	Federally mandated forms typically cannot be adapted locally to reflect local SOGI collection data standards, limiting SOGI data collection. Federally mandated forms

		cannot be adapted to reflect state or local SOGI data collection standards.

APPENDIX B: GLOSSARY OF TERMS

Acknowledge The Diversity of LGBTQIA+ Terminology

Terminology used to describe communities that fall under the LGBTQIA+ umbrella changes over time and can differ across demographics, cultures, and generations. Because individuals use terms that they identify with, asking how someone refers to themselves and reflecting that back are best practices. Health systems and the electronic data tools and paper forms they use often move more slowly than do shifting and evolving terms used within communities. For this reason, LGBTQIA+ individuals often encounter outdated or even offensive terms in medical, psychological, or legal contexts. By reviewing and updating these data systems and forms using terminology that is culturally responsive, providers and public health programs can minimize trauma, stigma, and health inequities. The definitions are based on information from LGBTQIA+ community organizations, including the GLAAD, the Human Rights Campaign, and the National LGBTQIA+ Health Education Center (at the Fenway Institute)

Glossary

Agender: A person that "lacks" a gender. Agender people see themselves as neither a man nor a woman, or both. They're gender-neutral and often describe themselves as gender-free or genderless. This term is under the non-binary, transgender umbrella. Agender is often considered a type of non-binary identity, but not all non-binary individuals identify as agender.

Asexual: A complete or partial lack of sexual attraction or lack of interest in sexual activity with others.

Bisexual: A person emotionally, romantically, or sexually attracted to more than one gender, though not necessarily simultaneously, in the same way or the same degree. Usually used for attraction to both men and women.

Gay: A person who is emotionally, romantically, or sexually attracted to members of the same gender.

Gender Fluid: is an identity under the multigender, nonbinary, and transgender umbrellas. Genderfluid individuals have different gender identities at different times. A genderfluid individual's gender identity could be multiple genders at once and then switch to none at all, or move between single gender identities, or some other combination therein. For some genderfluid people, these changes happen as often as several times a day and for others, monthly, or less often. Some genderfluid people regularly move between only a few specific genders, perhaps as few as two (which could also fit under the label bigender), whereas other genderfluid people never know what they'll feel like next.

Gender Identity: One's innermost concept of self as male, female, a blend of both or neither. How individuals perceive and refer to themselves can be the same or different

from their sex that was assigned to them at birth. Note: an individual's gender identity is independent of their sexual orientation.

Genderqueer: Genderqueer people typically reject notions of the binary and static categories of gender and embrace a fluidity of gender identity. People who identity as "genderqueer' may see themselves as being both man and woman, neither man or woman or as falling completely outside these categories. Genderqueer can also be used as an umbrella term encompassing identities such as agender, bigender, pangender, non-binary or gender-fluid.

Intersex: is an umbrella term for differences in sex traits or reproductive anatomy. Intersex people are born with these differences or develop them in childhood. There are many possible differences in genitalia, hormones, internal anatomy, or chromosomes, compared to the usual two ways that human bodies develop.

Lesbian: A woman who is emotionally, romantically, or sexually attracted to other women.

Non-binary: An adjective describing a person who does not identify exclusively with the gender binary. Non-binary people may identify as being both a man and a woman, somewhere in between, or as falling completely outside these categories. Some, but not all, non-binary people also identify as transgender.

Pansexual: Describes someone who is emotionally, romantically, or sexually attracted to people of any gender though not necessarily simultaneously, in the same way or to the same degree.

Queer: Used as an umbrella term for diverse genders or sexualities. People use queer to describe their own gender and/ or sexuality if other terms do not fit. For some people, especially some older LGBTQIA+ community members, "queer" has negative connotations, because of its use as a discriminatory term. This term was adopted/reclaimed in the late 1980s/early 1990s by some people who used it in an empowering way to describe themselves.

Questioning: Rather than be locked into a certainty, some people are still exploring or questioning their gender or sexual orientation. People may not wish to have one of the other labels applied to them yet, for a variety of reasons, but may still wish to be clear, for example, that they are non-binary or non-heterosexual.

Sex Assigned at Birth: The assignment of Male or Female as sex on an original birth certificate. This is the sex that a doctor, midwife, or other adult uses to describe a child at birth based on their external anatomy. Children born with variation of sex characteristics are usually referred to as intersex. Not every intersex individual uses this term.

Same Gender Loving: A term used within the Black and African American community instead of lesbian, gay or bisexual to express attraction to and love of people of the same gender.

Sexual Orientation: An inherent or immutable enduring emotional, romantic, or sexual attraction to other people. Note: an individual's sexual orientation is independent of their gender identity.

Straight: A person who is emotionally, romantically, or sexually attracted to members of the opposite gender. Men, women, and non-binary people may use this term to describe themselves.

Transgender: An umbrella term for people whose gender identity and/or expression is different from cultural expectations based on the sex they were assigned at birth. Being

transgender does not imply any specific sexual orientation. Therefore, transgender people may identify as straight, gay, lesbian, bisexual, etc. Some transgender individuals take steps to change their outside to match their gender identity, but that is not a necessity to be considered transgender. Some transgender people identify inside of the gender binary (as either male or female) while others identify as transgender along the gender spectrum.

Transgender Man: A man who was assigned female at birth. Transgender men have a male gender identity, and many trans men undergo medical and social transition to alter their appearance in a way that aligns with their gender identity or alleviates gender dysphoria. As with transgender women, not all transgender men experience gender dysphoria. This term can include Trans-masculine identifying persons.

Transgender Woman: A woman who was assigned male at birth. Transgender women have a female gender identity, and many transgender women undergo medical and social transition to alter their appearance in a way that aligns with their gender identity or alleviates gender dysphoria. As with transgender men, not all transgender women experience gender dysphoria. This term can include Trans-feminine identifying persons.

Two-Spirit: A term referring to an individual's gender identity used by some Native American and Alaska Native individuals. Some use as both a gender identity and a sexual orientation and/or identity.

Use another term: This option should be used for sexual and gender identities not listed.

APPENDIX C: LIST OF ABBREVIATIONS

Abbreviation	Definition
BRFSS	Behavioral Risk Factor Surveillance System
CalHHS	California Health and Human Services Agency
CDPH	California Department of Public Health
DDG	De-Identification Guidelines
EHR	Electronic Health Record
HIPAA	Health Insurance Portability and Accountability Act
HPS	Household Pulse Survey
IPA	California Information Practices Act
LGBTQIA+	Lesbian, Gay, Bisexual, Transgender, Queer, Intersex, Asexual, and Plus
LHJ	Local Health Jurisdiction
SOGI	Sexual Orientation and Gender Identity