Sex and Gender in Clinical Research
Meeting Summary

Background

The roles of sex, gender, and sexual orientation in clinical research are important but underexplored. The challenge of addressing Sexual Orientation and Gender Identity (SOGI) in clinical research is compounded by controversy over definitions and imprecise use of terms. Without endorsing one set of terms over another, here we use the terms as defined in the National Academies of Science, Engineering, and Medicine (NASEM) report on “Measuring Sex, Gender Identity, and Sexual Orientation”¹ for clarity, though we recognize that these definitions were contested by trans scholars present at the meeting on June 2:

- **“Sex”** is a multidimensional construct based on a cluster of anatomical and physiological traits that include external genitalia, secondary sex characteristics, gonads, chromosomes, and hormones.”²
- **“Gender”** is a multidimensional construct that links **gender identity**, which is a core element of a person’s individual identity; **gender expression**, which is how a person signals their gender to others through their behavior and appearance (such as hair style and clothing); and cultural expectations about social status, characteristics, and behavior that are associated with sex traits.”³
- **“Sexual orientation”** is a multidimensional construct encompassing emotional, romantic, and sexual attraction, identity, and behavior.”⁴

Discussion at the June 2, 2022, Bioethics Collaborative focused on two general issues related to sex, gender, and sexual orientation in clinical research. First, while SOGI are social constructs, disease risk, manifestation, and/or response to treatment may vary by SOGI characteristics. It is important to understand that any variation in disease risk, manifestation, or response, if it exists, is likely related to other underlying differences (e.g., hormonal status). Nevertheless, collecting and analyzing data on sex, gender, and sexual orientation may point to differences that need to be understood biologically to provide safe and effective interventions across

---

² *Id.* at 20.
³ *Id.* at 20.
⁴ *Id.* at 21.
diverse populations. Second, research should be respectful and inclusive of individuals of diverse sexes, genders, and sexual orientations. The principles of respect for persons and justice must be considered in addition to that of biologic relevance.

*Considerations When Sex, Gender, and/or Sexual Orientation May Be Scientifically Relevant*

Sex, gender, and sexual orientation are frequently discussed together. It is common to see commentaries describing the underrepresentation of LGBTQIA+ people in clinical research. While this grouping speaks to the social and political realities of lived experience, it does not appear to serve scientific inquiry. For example, the biology of trans and gender diverse (TGD) people, largely related to potential exposure to hormonal treatment, suggests that TGD should be considered as a separable category from sexual orientation, at least in the evaluation of safety and efficacy of medical interventions. It is also important to recognize that TGD itself is not a single category and may be a continuum that defies segmentation. While sexual orientation, gender identity, and gender expression are social constructs, they may be important social determinants of health and should not be ignored by clinical research.

“Sex” and “gender” are poor surrogates for a broad range of potentially scientifically-relevant variables, such as sex hormones, endocrine profiles, vascular and immune system differences, and genetic and genomic differences. An alternative approach to collecting SOGI data would be collecting data on the specific variables of interest, such as cortisol and testosterone levels; at this point in scientific understanding, however, it is unclear which variables may be relevant and how best to operationalize them. Should one collect SOGI variables at all? If differences in response to medical interventions across one or more SOGI variables are found, then biological correlates can be further explored. One attendee suggested the health (and social burden) of all people would be improved by moving away from the focus on sex and gender and instead considering these issues in the context of personalized medicine, suggesting that no grouping or predetermined approach will be helpful.

Unless a clinical trial is designed to enroll a sufficient number of participants that vary by sex, gender, and/or sexual orientation, it is unlikely that the study will be statistically powered to answer questions relating to SOGI. For this reason, TGD individuals are currently excluded from many clinical trials, creating differential access to novel interventions. Attendees discussed whether collecting “some” data is preferable to collecting no data in these situations. Most attendees agreed that collecting potentially relevant data is preferable but cautioned against overinterpretation of incomplete datasets; any interpretation of the data would be directional, and not conclusive, but informative for future research pursuits. One person challenged researchers to consider the social consequences of not collecting SOGI data. Abstaining from collecting these data may contribute to societal stigmatization of certain populations by failing to acknowledge and recognize their existence, with the (unintended) consequence of contributing to differential health outcomes and health inequities.
Attendees also made scientific arguments in favor of collecting data. First, collecting data on individuals of diverse sexes, genders, and sexual orientations allows researchers to examine assumptions about dosage, pharmacokinetics and pharmacodynamics, and risk-benefit profiles. One attendee recommended routinely collecting data on sex in preclinical development of investigational products, and sex and gender early in clinical development, to begin to examine assumptions about equivalence, even if studies lack the statistical power to answer these questions definitively. Second, even incomplete data can lead to hypothesis-generation for further research. Third, incomplete data from multiple studies can be aggregated to result in a sufficiently large dataset to allow statistical analyses with sufficient power to answer research questions. Attendees suggested establishing registries that collect and aggregate data on sex, gender, and sexual orientation and health outcomes following regulatory approval of an investigational product. Registries on SOGI, with particular reference to TGD people, would inform drug safety and efficacy in the various populations. The concept of post-approval registries is analogous to pregnancy exposure registries that collect health information from mothers and babies following intrauterine exposure to prescription medicines and vaccines. Regulators may need to require or incentivize such post-approval analyses.

Attendees discussed two challenges with collecting SOGI data. First, research sponsors may be overwhelmed by the seemingly infinite number of questions that could be asked about sex, gender, and sexual orientation. Understanding epidemiology is an important grounding exercise: data on disease prevalence, severity, course, and treatment and their variation by sex, gender, and sexual orientation may provide some clarity regarding whether and when to ask certain questions. Second, there may be a tension between accuracy and standardization when determining how to collect SOGI data. On the one hand, standardizing data collection methods makes it easier for researchers to use these methods and facilitates data aggregation and comparison. On the other hand, definitions of sexual orientation and gender vary geographically, culturally, temporally, and by language, as does the culturally-acceptable terminology. In this context, it is important to remember that language—and culture—change over time. The fluidity of these definitions poses a challenge to longitudinal studies but should not deter researchers from collecting SOGI data or conducting long-term studies focused specifically on groups within the LGBTQIA+ community.

**Ethical Considerations**

A primary ethical consideration is privacy and confidentiality when collecting data on sex, gender, and sexual orientation. Re-identification of transgender and/or nonbinary individuals, for example, sometimes only requires geographical information and a few relevant variables. Over 70 countries currently have anti-homosexuality laws⁵ and some prohibit non-heterosexual behavior and/or certain types of gender expression. Concerns, however, about protecting

---

participants’ identities are not limited to those countries. Even in countries where same-sex marriage and diverse gender expression are legal, members of these communities are subject to discrimination. For example, Texas House Bill 25 prohibits trans children enrolled in public schools from competing on teams consistent with their gender identity, and seventeen other states have similar laws. Data on sex, gender, and sexual orientation can be used against individuals legally, socially, politically, and personally. Particularly in settings that may expose participants to risk, special care should be taken to ensure that (1) the data are necessary, (2) data minimization principles apply, (3) data collection, transfer, and storage are secure, and (4) the participants are informed of and understand the risk before agreeing to provide any information.

Bioethics Collaborative attendees identified ways to facilitate the inclusion of individuals of diverse sexes, genders, and sexual orientations in research. Building relationships with populations that have historically been excluded from research, such as transgender communities, is foundational, and establishing advisory discussions and community input to the research is helpful. Engagement with local health organizations is important and can facilitate the siting of trials within diverse communities. Training can be offered to acclimate researchers to asking questions about sex, gender, and sexual orientation in a respectful manner. Every study should have inclusion/exclusion criteria that are inclusive, gender-neutral, and relevant to the study. Eligibility criteria that exclude specific populations should require ethical and/or scientific justification. IRBs are the natural mechanism for ensuring that studies uphold these standards.

Finally, while a clinical research protocol may not intend to collect and record SOGI data, investigators and their study personnel should be inclusive in their interactions with potential participants and their caregivers. Asking personal pronouns as a matter of course, for instance, demonstrates respect and acceptance of all eligible participants.

U.S. Government Agency Updates

The National Institutes of Health (NIH) Sexual and Gender Minority Research Office (SGMRO) coordinates with other NIH offices on initiatives related to the health of diverse sexual and gender populations. The NIH SGMRO commissioned the NASEM as an independent body to analyze and recommend approaches to data collection on sex, gender, and sexual orientation across all U.S. research agencies. Ideally, President Biden and the U.S. Office of Management and Budget would issue an executive order to mandate the collection of these data. Without an executive order, agencies are unlikely to collect these data, and without data it is challenging for the NIH SGMRO to push for new initiatives. Looking ahead, the President’s Budget for fiscal

---

year 2023 includes funding research on sex and gender identity, including the development and validation of data collection methods for survey instruments.  

Several reports and guidance documents demonstrate the U.S. Food and Drug Administration’s (FDA’s) interest in the roles of sex and gender in clinical research. In 1992, the U.S. Government Accountability Office issued a report that recommended FDA encourage the inclusion of women in clinical trials, and in 1993, FDA issued a guidance document that called for data analysis by sex. FDA implemented a rule in 1998 that called for safety and efficacy data to be presented across sex, age, and racial subgroups. More recently, FDA issued guidance in 2020 on enhancing the diversity of clinical trial populations and released draft guidance in April 2022 to improve the enrollment of underrepresented populations in research. While the draft guidance focused on racial and ethnic populations, it did reference the importance of sex and gender identity.

Conclusions

One area for future work is identifying whether, when, and how SOGI data should be collected in clinical research. While it is reasonable to assume that sex, gender, and sexual orientation – or rather, the variables they serve as proxy for – are scientifically relevant in at least some circumstances, background epidemiological data are necessary to understand their relevance and impact. These epidemiological data, however, will only be available if they are collected. To begin this process, SOGI data collection and aggregation could borrow lessons from other domains, such as pregnancy exposure registries and rare disease research that collect observational data. Discussion at the Bioethics Collaborative suggested that TGD be considered separately from sexual orientation. Ethical principles should be applied to develop best practices to manage SOGI data, including protecting individual privacy and confidentiality.

---

Existing best practice standards for handling sensitive information provide a foundation on which to build.

A second area of future work might be to evaluate each step of the clinical trial process and develop guidance on how to respect and include diverse individuals in research. Possible considerations include designing protocols and recruitment materials that employ inclusive language, training research staff to ask for participants’ pronouns, siting research at locations with gender-neutral facilities, and challenging sponsors and researchers to design inclusive eligibility criteria and IRBs to evaluate their scientific and ethical validity. Community partners and other stakeholders are important for this work, which should be inclusive by design.