Mathematics & Statistics Faculty Works   Mathematics & Statistics

2-1-2022

Towards Statistical Best Practices For Gender And Sex Data

Suzanne Thornton
Swarthmore College

D. Roy

S. Parry

See next page for additional authors

Authors

Suzanne Thornton, D. Roy, S. Parry, D. Lalonde, W. Martinez, R. Ellis, and D. Corliss

Best Practices for Collecting Gender and Sex Data

Word Count: 6027

Suzanne Thornton*

Assistant Professor

Mathematics and Statistics Department, Swarthmore College

500 College Ave, Swarthmore, 19081, PA, United States

email: `sthornt1@swarthmore.edu`

Dooti Roy

Principle Methodology Statistician, Boehringer Ingelheim

Danbury, 06810, CT

email: `dootiroy@gmail.com`

Stephen Parry

Statistical Consultant, Cornell University

Ithica, 14850, NY, United States

email: `sp2332@cornell.edu`

Donna LaLonde

Director of Strategic Initiatives and Outreach, American Statistical Association

Alexandria, 22314, VA, United States

email: `DonnaL@amstat.org`

Wendy Martinez

Director, Mathematical Statistics Research, Bureau of Labor Statistics

Washington, DC, 20212, United States

email: `martinez.wendy@bls.gov`

Renee Ellis

Family Demographer, US Census Bureau

Washington, DC, 20002, United States

email: `reneerellis@gmail.com`

David Corliss

Lead data scientist, manager of the Data Science Center of Excellence, Stellantis

Auburn Hills, 48309, MI, United States.

email: `davidjcorliss@gmail.com`

March 18, 2021

∗Corresponding author. E-mail: sthornt1@swarthmore.edu

Abstract

The measurement and analysis of human sex and gender is a nuanced problem with many overlapping considerations including statistical bias, data privacy, and the ethical treatment of study subjects. Traditionally, human gender and sex have been categorized and measured with respect to an artificial binary system. The continuation of this tradition persists mainly because it is easy to replication and not, as we argue, because it produces the most valuable scientific information. Sex and gender identity data is crucial for many applications of statistical analysis and many modern scientists acknowledge the limitations of the current system. However, discrimination against sex and gender minorities poses very real privacy concerns when collecting and distributing gender and sex data. As such, extra thoughtfulness and care is essential to design safe and informative scientific studies. In this paper, we present statistically informed recommendations for the data collection and analysis of human subjects that not only respect each individual's identity and protect their privacy, but also establish standards for collecting higher quality data.

1. INTRODUCTION

When teaching introductory statistics, many instructors encounter the question: Why do we typically use a significance level of 0.05 (or 0.01 or 0.10)? The answer is simple but not very gratifying: this is a matter of convention. Helpful guidelines are well-established but there are infinite possible significance levels a researcher could decide to use. A similar conundrum occurs when collecting data on socially-defined attributes such as race or gender. The typical ways of categorizing these variables have been passed down from statisticians to medical practitioners and other researchers more often due to their ease of replication rather than their actual scientific value. However, qualities such as these represent social constructs which, by their very nature, are not fixed but are fluid concepts with evolving definitions. Merriam-Webster defines the term social construct as "an idea that has been created and accepted by the people in a society". Scientific endeavors must recognize that many qualitative human characteristics fit this definition and as researchers, we are responsible for practicing extra thoughtfulness and care when considering these traits as variables of interest in our studies. Ultimately, we cannot adequately provide scientific or medical benefits to socially

4

marginalized people if we are not collecting the right data.

The goal of this document is to present statistical practitioners with guidelines on how to collect quality data on sex and gender for adult subjects. This work is meant to serve as an introduction to data quality concerns and ethical concerns from a perspective of gender inclusivity, a perspective that has not traditionally been emphasized in the history of data science. We believe it is time to challenge convention and strive for better practices when collecting and analyzing human-centric data. This document is not a definitive source for understanding the complexities of gender identity but does provide an introduction to the nuances of gendered language from a statistical perspective and outlines a framework for researchers to adopt to better devise useful to address this specific challenge in quantifying the human experience. These recommendations are meant to be revisited and updated as our society evolves.

2. BACKGROUND INFORMATION AND TERMINOLOGY

While a thorough introduction to the language of sexual and gender identities is beyond the scope of this paper, we have listed some terms and definitions in Figure 1 which will be referenced throughout the remainder of the text. Many other reputable resources elaborate upon these terms and present more terminology than we do, e.g. Bates et al. [2019]; Bauer et al. [2019]; and Badgett and Team [2009]. For additional resources, see Appendix A.

Term	Definition
Cisgender	An adjective used to describe a person whose gender identity aligns with the gender assigned to them at birth.
Gender affirming surgery	Surgery that is performed to modify a person's body to be more aligned with that person's gender identity. These may include chest and genital surgeries, facial feminization, and/or hair removal.
Gender expression	The many ways in which someone may express their gender identity. This can include clothing, behavior, and mannerisms which may be intentionally or unintentionally ambiguous.
Gender identity	One's innermost concept of self as male, female, a blend of both or neither – how individuals perceive themselves and what they call themselves. One's gender identity can be the same or different from their sex assigned at birth. Gender identity cannot reliably be inferred from gender expression.
Gender or sex minority	A catch-all phrase that includes transgender, non-binary, other gender non-conforming, and intersex people.
Different Sexual Development (DSD)	See Intersex.
Intersex	An adjective that describes someone who has chromosomes, anatomy, or other sex characteristics that don't fit the typical definitions of female or male. This can include people with genetic predispositions to androgen sensitivity or people with hormones and genetics that indicate one sex but an anatomy that is typical of another sex.
Misgender	A verb that describes the action of referring to a person by a pronoun or other gendered term (e.g., Ms./Mr.) that incorrectly indicates that person's gender identity. This is a significant source of emotional harm for transgender individuals.
Non-binary	A term for people whose gender identity exists outside of the gender binary of male or female. Many non-binary people identify as transgender since their gender identity is likely to be different from the sex assigned to them at birth.
Sex assigned at birth	The sex (typically male or female but occasionally a neutral X) assigned to a newborn. Most often, this designation is based on the infant's anatomical characteristics.
Sexual orientation	An inherent or immutable enduring emotional, romantic or sexual attraction to other people.
Transgender	An adjective used to describe a person whose gender identity is different from the gender they were assigned at birth. Note that although intersex and transgender identities may overlap, they are not necessarily co-occuring identities.
Transition	The often lengthy process whereby a transgender person begins to acknowledge their true gender identity in their private and/or public life.

Figure 1: Common terms relevant to gender and sex minorities. Definitions are derived from the National LGBTQIA+ Health Education Center and the Human Rights Campaign.

It is important to note the many possible intersections of LGBTQ+ identities. Naturally, the full scope of people's identities extends beyond any one particular categorization. Human beings exist as unique intersections of multiple identities, some of which are ascribed by others and some of which are personal and may not be readily identifiable to others. Figure 2 presents a simple diagram that illustrates the possible (in)congruencies among gender and sex characteristics.
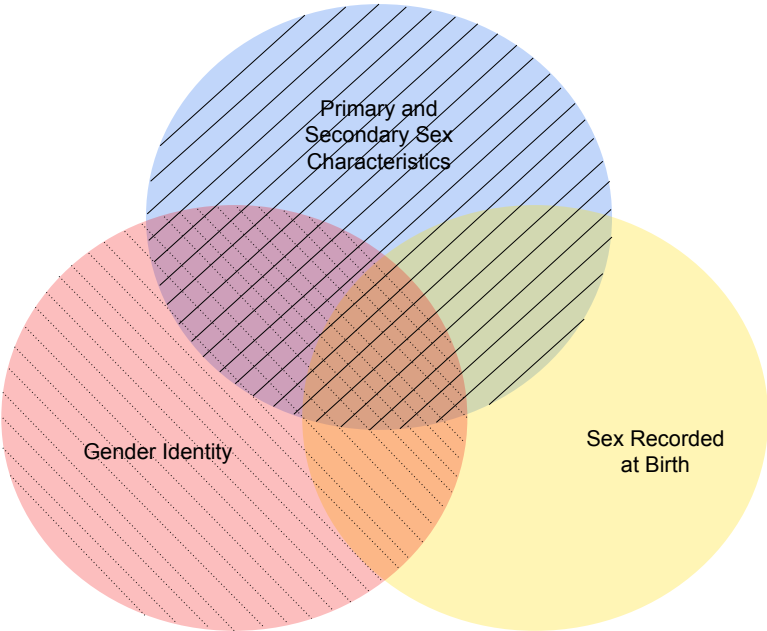

Figure 2: A Venn Diagram show three different aspects of sex and gender identity. For many people, all three characteristics align; however, for gender and sex minorities one or more of these characteristics may not align in the typical manner.

2.1 Sexuality

We are not discussing sexuality in this paper but we mention it here because sexual preference is often confused with the concepts of gender identity and sex. While sex and gender are descriptors of one's self and personal identity, sexuality is a descriptor of one's romantic attraction to others.

For example, a transgender woman may identify as straight if she is only romantically interested in men; a cisgender man may identify as gay if he is attracted to other men who may or may not be cisgender. It is important for data scientists to be aware of this distinction because potential biases

can be introduced in a data set by focusing on one identity but using general, catch-all terms like "LGBTQ+" or "queer" — neither of which specify identify gender identity or sexuality.

2.2 Gender

Though a nuanced understanding of gender may seem like a new trend to some, the modern recognition of gender as a continuum actually has roots that stretch across the globe and into the ancient past; c.f. e.g. Wiesner-Hanks [2011] and Herdt [2020]. For modern scientists to be advocating for an understanding of gender identity that extends beyond a binary classification system, is actually an example where modern science is only just now catching up to global historical and cultural wisdom.

The enornaminity of gender identity is an immutable factor of the human experience; the language with which we describe that experience however is non-constant and changes with generations. For example, although the Chicago Manual of Style renounced the formal use of "they" as a generic third-person singular pronoun in 2017, two years later the American Psychological Association endorsed the formal use of "they" to refer to a person whose gender identity is unknown or irrelevant to the context of the usage. Regardless of the social norms of the times, the purpose of data collection has always been to understand a particular population by capturing information in bite-sized, digestible pieces. Although information loss is an inevitable byproduct of this process when applied to social constructs such as gender, our goal as statistical practitioners is to design surveys and experiments to capture the most relevant bits of information as possible.

The concept of gender in modern society can be broken into two components, gender expression and gender identity. Respecting someone's gender identity is an ethical practice that honors the individual's sense of self. Gender identity is a personal understanding of oneself; it comes from within, not without. Gender expression refers to many possible expressions of gender including attributes such as clothing, hair, mannerisms, postures, etc. Gender expression is like art in that it may be formed by an artist who wants to communicate something particular but is subject to the interpretation of the viewers who may or may not understand the artist's intent. For any individual, gender identity and gender expression can change with time. For example, a transgender woman may not have transitioned until later in her life and thus in her younger years may have adopted masculine or androgynous gender expression for various reasons including her personal safety and

comfort. The process of transitioning in one's public and private life takes years and can be a non-linear process. Some transgender men for example, may identify as non-binary as they begin to question the gender identity imposed upon them from birth and then realize their identity as men years later. Gender expression can be intentionally ambiguous or it can challenge the status quo. For all of these reasons, it is best to never assume any individual's gender identity from their gender expression alone.

A common way to communicate gender identity is to share one's pronouns, similar to sharing a surname or a title. Asking someone which pronouns they use and providing your own is a way to actively respect diverse gender identities. Pronouns can indicate information regarding how one identifies themselves along the gender spectrum, from masculine to feminine and beyond. Some people use gendered pronouns such as she/her/hers or he/him/his, other people may only use the gender neutral singular they/them/theirs, and others still may use different gender-neutral pronouns such as ze/zim/zir. In line with the American Psychological Association, we recommend the usage of the singular "they" as an appropriate and respectful way to refer to people whose gender identity is unconfirmed or non-binary.

2.3 Sex

Despite the overwhelming tendency to treat biological sex as a binary variable, the term "biological sex" is not quite so cut and dry. When sexual characteristics of individuals are a variable of interest in a study this can refer to primary or secondary sexual attributes, genetic information, gender assigned at birth, or many other sexual characteristics. The most common use of the phrase "biological sex" seems to refer either to a person's chromosomes or to their genitalia at birth, however, neither of these attributes are actually binary [Ainsworth, 2018].

A consensus statement published in 2005 endorsed the terminology disorded/different sex development (DSD) to describe individuals with congenital conditions in which the development of chromosomal, gonadal, or anatomic sex is atypical. However, the term intersex is preferred by many activists and advocates since it is not affiliated with a derogative connotation like the word "disorder" which implies something that needs to be "fixed" [Hughes et al., 2006, interACT, 2020]. In this article we will therefore refer to anyone with these congenital conditions as intersex.

Intersex people can be born with the differences mentioned above, or they may develop these

differences later in childhood. This makes pinning down the incidence rate very difficult. Most experts agree that the prevalence of intersex births is about 2% of the population which, for comparison, is similar to the prevalence of red hair; c.f. e.g. Onyiriuka et al. [2016] or Fausto-Sterling [2000]. People with ambiguous or atypical sexual characteristics are often coerced into changing their bodies, usually at a very young age. Most surgeries to change intersex traits happen in infancy and it is possible that an adult may not be aware of this procedure in their medical history as noted by Greenberg and Stam [2012].

We bring up these points to illustrate that when one is conducting a study that merely asks participants for their sex, one risks losing important information by making the assumption that all participants understand "biological sex" in the same way. This is true not just for intersex participants but also can apply to transgender patients who may have undergone gender confirming genital reconstructive surgery. When collecting data on attributes that can have different meanings, it is best for the researcher to be as explicit as possible about their working definitions to avoid misclassification bias.

3.   A MULTIFACETED ARGUMENT FOR INCLUSIVE DATA

Among the many reasons why researchers may want to be more inclusive in their data collection process, here we focus on two: the ethical treatment of human subjects and the collection of better quality data. While we do not claim to be experts in ethics, data is our specialty, and we have found that these two arguments overlap. Any ethically responsible researcher or clinician collecting data on human participants should be aware of the language needed to correctly identify different segments of the population. When options outside of the gender binary are not provided in surveys, for example, the researcher overlooks an entire group of people from their analysis. Or when individuals do not see their identities respected in the data collection process, they may experience feelings of confusion, depression, or failure as a result. On a larger scale, exclusion of gender and sex minorities means that research results may be unable to identify smaller sub-populations who may be the most vulnerable to certain illnesses.

For instance, previous research has shown that transgender individuals might be at a higher risk than the cisgender population for a variety of health issues such as stress, blood pressure,

and heart attacks [Alzahrani et al., 2019, Dubois, 2012, Irwig, 2018]. Mental health issues are also disproportionately prevalent among gender and sex minorities, and, as is widely acknowledged within the LGBTQ+ community, the main factors contributing to these issues stem from perceived and realized threats to personal safety and social inclusion; c.f. e.g. Olson and McLaughlin [2016]; Clements-Nolle et al. [2006]; or Pascoe and Richman [2009]. By collecting relevant information in a way that respects gender identities, researchers can help mitigate the crippling barriers to health care than many transgender and non-binary people face. For example, a 2018 study found that transgender people whose pronouns were respected by all or most people in their lives attempted suicide at half the rate of those who did not have their pronouns respected [Russell et al., 2018].

The fundamental goal of statistics is to articulate information about some population through the collection and analysis of data. The ethical statistician must use methodology and data that are relevant and appropriate, without favoritism or prejudice, in a manner that produces valid, interpretable, and reproducible conclusions [Association, 2018]. There is much information to be learned from gender and sex qualities of human subjects provided these qualities are researched in a way that respects the variability of identities and experience.

As is true for data collection at large, when gender and sex are attributes collected without forethought, the data can be easily and irreparably corrupted by statistical biases. A study is susceptible to non-response bias, for example, when survey participants do not see their identities represented in the questions or answers. In such situations, participants may become disgruntled and not complete the survey and therefore the conclusions from statistical analyses of the data may not apply to important subgroups of the population. Of course misclassification bias is also possible when surveys are not designed to acknowledge different identities. If a clinician assumes a participant's gender identity based on their appearance, voice, or mannerisms, the information encoded in the survey may be irrelevant, or worse, inaccurate.

With both response and non-response bias, the observed data fail to accurately represent the population of interest. Ultimately this can lead to a loss in the signal of the response variable. Although it will take time and effort to transition to more inclusive data collection practices for gender and sex, these efforts will eventually result in better quality data and more useful analyses. With more informative data, signal detection may become easier and resulting statistical models

may be useful for a wider range of the population. Example 1 considers a setting in which these biases can be hazardous to the health of transgender patients in particular.

Example 1

Setting: Healthcare providers often hesitate to prescribe drugs citing that there is no information about the safety of the drug on the specific population on the drug label. If there are no data on a specific group of people and the healthcare providers suspect that the subgroup population may react differently than the rest, they may be reluctant to prescribe the drug. A number of studies have highlighted the increased risk of mental illnesses (e.g. depression, anxiety, and suicidality) in gender and sex minorities, e.g. Su et al. [2016]. For clinical trials that are focused on treatments for mental illnesses, it is therefore crucial to identify individuals who may have a higher risk of certain adverse events. Furthermore, transgender and non-binary patients may take large doses of hormones as part of their transition and hormone therapy may significantly influence and impact the safety or efficacy of a particular drug. For these reasons, it is especially relevant for these trials to be able to identify transgender patients to ensure that a drug is safe and effective across all populations.

Problem: Most modern clinical trials are not designed to identifying transgender patients, even if the patient voluntarily self-identifies. Also, within a clinical trial setting, researchers are required to justify and explain why certain information is collected and needed for the study.

Proposed solution: Collect both "sex assigned at birth" and "gender identity" as the baseline characteristics during the conduct of the clinical trial. If these two fields do not align, this indicates a patient who is potentially transgender. Researchers can then examine these sub-populations to determine if there is an adequate representation within the sample and to see if there may be reason to suspect that the efficacy and safety of the test drug is different for this sub-population. This solution can help clinicians determine if transgender subjects may have a different response to the test drug.

Another consideration when incorporating gender or sex information into a statistical model is to be aware of the potential biases from imputing missing data. There are many programs that attempt to infer a person's gender using data mining techniques. The possible biases resulting from these methods depend on the basis with which these packages determine gender and on whether gender identity or sex at birth (or something else) is the information being imputed. The gender

R package for example, categorizes a person as "male" or "female" based historical data of gendered names [Mullen, 2018]. Therefore, as acknowledged in the package documentation, in no way can this package determine a person's gender identity or can it determine any specific details about a person's sexual characteristics. Furthermore, it can only partially identify subjects from historically marginalized gender roles as it does not measure transgender or other gender-non-conforming individuals. Another tactic that is often used to impute subjects' missing gender or sex classifications is to use crowdsourcing techniques. Most often, these methods rely on people who are paid to look at names and sometimes pictures of the subjects and then make a personal judgment to classify the subject as male or female. Hence, this method is subject to both personal and cultural bias and prone to misclassification errors.

Collecting data in a way that honors different gender identities will contribute to an increase in the general knowledge of and respect for various gender identities in the greater scientific community. The results of these efforts could have an enormous effect on the quality of life of gender and sex minorities both by setting better scientific standards for the ways in which human subjects are treated with respect and by providing better targeted health care through more representative data analyses. Since knowledge of gender and sex demographics can better inform a clinician in their own practice, the statistical argument for more informative data and the usefulness of the resulting statistical models is highly connected to the ethical argument for inclusive language and data collection procedures.

4.   STATISTICAL CONSIDERATIONS AND PRACTICAL GUIDELINES

Here we present a three-step approach for statistical practitioners to consider whenever collecting data on human sex and gender. Broadly speaking, the steps we encourage each researcher to take are to 1) identify the information relevant to the study; 2) respect the study participants' identities with inclusive language; 3) protect the data and participants' identities. We discuss statistical considerations that arise within these steps and offer some solutions for common issues that may arise, and apply these suggestions in two real examples.

4.1 Identify relevant information

Gender and sex information does not necessarily need to be included in every study. Before researchers begin a project, it is important to determine whether it possible that gender or sex may affect the efficacy of the treatment. If researchers do not believe that gender or sex can moderate an effect, then they may not need to ask for this sensitive information. However, in many clinical settings in particular, the researcher may expect gender identity or sex to significantly affect the efficacy of a treatment. In such cases, it is important for the researcher to explicitly state the underlying mechanisms. For instance, if sex may affect the efficacy, which sexual characteristics specifically are hypothesized to affect the treatment? If someone is receiving hormone therapy treatment, is this an important factor in the efficacy of the treatment under study? If someone has undergone gender-affirming surgery, is this an important factor in the efficacy of the treatment? Ultimately, asking more specific questions about bodily attributes rather than using gender or sex as a catch-all variable will lead to more reliable and inclusive data. For example, cisgender women, transgender women, and transgender men may all need breast cancer screening. Therefore, a survey that asks patients for their gender identity only will exclude transgender men and conversely, a survey that asks patients for their sex assigned at birth only, will likely exclude transgender women.

Often, researches use the words "sex" and "gender" as interchangeable characteristics. This is a dangerous practice in that it will most certainly lead to some type of misclassification bias. If gender or sex are relevant features of a statistical model, the researcher must dig deeper and ask is "What is relevant to my study, gender or sex or both?" This is merely an example of the more general statistical best practice, when collecting data from which to construct a statistical model, the researcher must carefully consider what are the variables of interest. Although there are many statistical techniques for variable selection, there is no equal substitute for domain-specific knowledge when it comes to variable selection. Example 2 provides an example of a clinical study where gender and sex information were conflated, influencing the conclusions of the data analysis.

Example 2

Setting: *One of the authors (ST) served as a methodologist for a research project that developed a predictive model for a certain type of drug-resistant epilepsy. The researchers were interested in determining the efficacy of an anti-epileptic drug on people with different types of epilepsy [Choi*

et al., 2020]. To predict whether or not this drug will reduce seizures in various types of patients, the researchers conducted a retrospective study and developed a predictive model based on data that had been collected many years prior across two different clinical groups. Useful predictive variables included age; different seizure types; whether or not there was a change in seizure frequency with a patient's menstrual cycle (called Catamenial epilepsy); and the sex of the patient (which was stated without explanation and only marked as male or female).

Problem: *The most important predictive variables are heavily related to one another. In particular, Catamenial epilepsy and sex contain overlapping information. The solution proposed by the statistician was to develop a new, three-level categorical predictor variable that indicated both sex and the occurrence of Catamenial epilepsy rather than to use each of these variables separately. Upon further reflection, the statistician started to realize the limitations of assessing signals from ambiguous data collection procedures regarding sex and gender and realized that the data collection procedure regarding the sex of the patients in this study was insufficient for capturing what appears to be the most relevant information about this variable.*

Proposed solution: *For future studies, at the data collection stage, rather than asking for a patient's sex, more informative data could be easily obtained by instead asking patients both for their gender identity and whether or not they regularly experienced menses. This simple adjustment to the data collection stage would prevent the exclusion or missclassification of transgender men (who may still experience menses) and cisgender women who may not experience menses due to age or medical procedures (such as a hysterectomy) or having non-ovary producing internal genitalia.*

In a clinical trial setting where a group of many typical characteristics of human sex development is of interest, an inclusive question may ask participants for their sex assigned at birth. This will typically be either assigned male at birth (AMAB) or assigned female at birth (AFAB). However, it is worth considering how important information about intersex people may be lost in this classification system or how important information about transgender people may be lost if gender identity is not also specified. Furthermore, in the US and in other countries, gender neutral birth markers (X) are an increasingly popular option for new parents.

Research has shown that individuals from the LGBTQ+ population tend to experience higher rates of depression and suicidal ideation as well as perceived and experienced threats to personal

safety [Almeida et al., 2009]. If mental health could possibly moderate the effect of a treatment, then sexual orientation and gender identity might be important variables to collect.

Researchers may be interested in gender identity or sexual characteristics of participants, but must ensure binary sex is not being relied upon as a proxy for another more precise variable such as the presence of breast tissue or type of hormonal balances. Imagine, for instance, designing a study to identify determinants of prostate cancer. Rather than ask patients for their sex, one may instead simply ask patients if they have (or have ever had) a prostate. This type of gender-neutral re-wording of survey questions is much more specific and can collect more relevant information in a way that respects individuals' identities. These measures can help avoid the confusion of dealing with transient variables such as sex or gender, both of which may change with time depending on how they are defined. Furthermore, this type of gender-neutral questioning avoids causing any unnecessary mental trauma to transgender and gender-non-conforming individuals who may feel outted or experience intense gender dysphoria as a result of being forced to identify with a binary gender that does not match their own identity. Figure 3 outlines general steps for evaluating what information is relevant for differently motivated studies.

4.2 Center inclusivity and respect

The purpose of institutional review boards is to protect the rights and welfare of human subjects. Human welfare is intimately connected with perceived respect, especially for those that identify with various socially and politically marginalized groups [Durwood et al., 2017]. The discretization of social constructs for the purpose of statistical analysis is not new and neither, unfortunately, are data collection processes that disregard individual identities. For example, it was not until 1960 that individuals could select their own racial identity in the US Census [Brown, 2020]. As a consequence of the National Research Act being signed into law in 1974, the Belmont Report identified "respect for persons" as a basic ethical principle of research. The report defines "respect for persons" based on two fundamental ethical convictions: first, that individuals should be treated as autonomous agents, and second, that persons with diminished autonomy are entitled to protection [HEW, 1979]. It is critical for the researcher to take extra measures to ensure the respectful treatment of human subjects is an intrinsic feature of any human-centric scientific endeavor. Unfortunately, the considerations necessary for the respectful treatment of gender and sex minorities are too often overlooked.
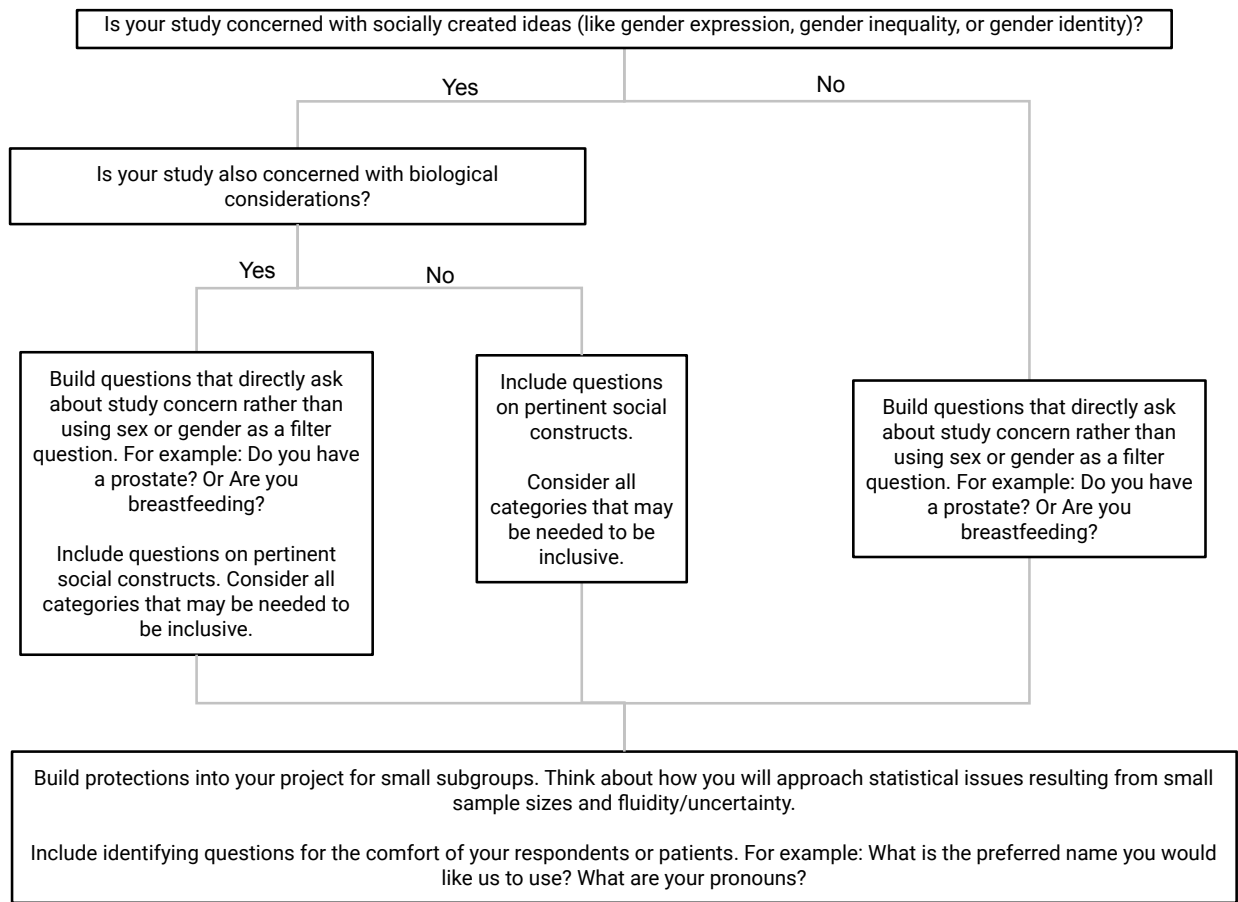
Figure 3: A flow-chart outlining how to identify the relevant gender and sex information to reference in a particular study.

One important measure that expresses respect to participants is the explicit communication of the purpose behind collecting sensitive data. Participants have the right to understand why their information is being collected and how it is going to be used and their identities protected. As discussed in the previous section, another important consideration is to recognize that questions about sex and gender may cause an undue amount of stress or confusion for transgender and gender-non-conforming respondents. If it is necessary to question a study participant about their sex, the researcher should clearly define what is meant by "sex" in a way that does not coerce the participant to adopt an inappropriate gender identity but instead emphasized specific sexual characteristics without gendered language. For example, in many non-clinical studies, gender is recorded because the variable of interest is whether or not an individual belongs to a historically marginalized group. In such cases, it would be misleading to discount the marginalization of transgender men, say, by grouping them together with cisgender men. In this case, the question about gender could be reworded to instead ask participants if they are cisgender men, cisgender women, transgender (men or women) or otherwise gender-non-conforming. As another example, if the outer genitalia a person currently has is a key feature of interest, it's best to ask this explicitly rather than ask the participant if they are male or female. On the other hand, if the genitalia a person was born with is the key factor, then the question can be easily adapted to adjust for this without challenging the participant's gender identity.

In situations where the data analyst has no control over the data collection procedure, such as in retroactive studies, the analyst is responsible for noting any ambiguities or possible sources of bias in the data at hand. If gender information was collected but there were no options for people who did not identify as a man or a woman, this is important to note because the study is limited by not applying to gender-non-comforming populations. If the retroactive study collected information on the sex of subjects, the data analyst should strive to understand this variable in the context of the original study. In the resulting analysis, the researcher can still demonstrate respect for the human subjects by including a discussion on how sex was coded and, following the suggestions in the previous section, by discussing what particular sex characteristics were probably of interest using gender-neutral language. The importance of this careful and thoughtful data collection is only growing as many states (and countries) begin to offer gender-neutral identifiers on official

18

documentation such as birth certificates and drivers' licenses [Fitzsimons, 2019].

Even if the data analyst can control the manner in which the data is collected, creating additional categories such as "non-binary" for gender or "intersex" for sex can create difficulties in the statistical analysis if small sample size considerations are not taken into account. Whether or not a particular category is a valid option for a variable in a statistical model often comes down to an issue of sample size. To avoid this small sample size problem that often results in wasted data, the researcher must take time before collecting the data to plan the study with carefully developed sampling strategies and/or with explicitly defined questions that take into account the topics discussed in the previous section. Obviously these measures are important for studies that focus on the LGBTQ+ community. These considerations are especially important for gender and sex minorities, c.f. e.g. Lau et al. [2020]; Goldstein et al. [2017]; and Gupta et al. [2016].

Some additional considerations for surveys that take into account the points discussed above. When asking about participants' sex, consider including intersex as an option or if possible, consider rewording the question to ask about the specific sexual characteristics of interest. When asking about gender identity for classification purposes, include at least a third option for non-binary or gender non-conforming individuals. Similarly, if asking for participant pronouns be sure to include a gender neutral option such as the singular "they" in addition to the typical binary options of she/her/hers and he/him/is. When it is important to identify transgender and other gender and sex minority groups, we recommend using the two-step approach, asking participants both for their sex assigned at birth and for their gender identity [Badgett and Team, 2009].

4.3 Protect the participant and the data

In our day-to-day activities, we regularly find ourselves being asked to provide our sex for events such as doctor's appointments, filling out the Census, registering a profile on a social media page, or applying for a job. For cisgender people, these questions may not cause a second through but for transgender and gender-non-conforming people, these questions can elicit not just mild discomfort but feelings of dread and fear for personal safety. There are many reasons why a person may not want to publicly reveal that their sex assigned at birth does not align with their gender identity; c.f. e.g., Safer et al. [2016] and Wirtz et al. [2020]. From a statistical perspective, is important to recognize these concerns from the perspective of gender and sex minorities. If data protection and

personal privacy cannot be guaranteed to these groups where the degree of potential harm from data misuse is great, there may be a severe non-response bias. More importantly, an individual's safety may be under threat. Of course from an ethical perspective, scientists recognize that data collection should not cause undue levels of stress to participants [Bennouna et al., 2017]. There are two levels of data security to consider when collecting information on a person's gender identity or sex. The security of this information is crucial when collecting the data and when analyzing or sharing the data.

In practice, data privacy is complicated by a lack of general or consistent standards. Differing approached to privacy can be illustrated by considering the parts of speech used. The usual method for protecting data privacy is through nouns: creating a list of data objects that qualify for protection as Personally Identifying information (PII). This results in conflicting standards as each regulating organization may have a different list. For example, the US Department of Homeland Security has a list of nineteen items, six of which are considered PII alone and nine are considered PII if paired with another item on the list [DHS, 2012]. However, the US Department of Health and Human Services has a single tier of 18 items. Only thirteen items are on both lists. Neither protects gender; sexual orientation qualifies as Sensitive PII for Homeland Security but is left unprotected by HHS. Modern analytic algorithms pose an additional problem: identifying persons through a multi-point match of data objects appearing on no list. A different approach to data privacy is suggested Australian Data Privacy act, established in 2008 and updated it in 2014 and 2019 [Morris and Cravigan, 2020]. This act takes a verb-based approach to privacy, placing restrictions where an individual "reasonably identifiable". Taking this notion a step further, we can better understand data privacy as an adverb: the manner in which things are done. Assuring data, personal identify, and privacy are protected with adverbs - securely, transparently, ethically - achieves the desired intent.

Socially typical safety standards may not be adequate especially for transgender individuals [Roffee and Waling, 2017]. As noted, the US HHS standard conflicts with others and affords no protection to sexual identity data. The unfortunate reality is that many gender-non-conforming children may not feel safe to reveal their gender identities if a parent or legal guardian is present; c.f. e.g. Riggs and Bartholomaeus [2018] and Catalpa and McGuire [2018]. Although we are limiting our

discussion in this article to data collection for adults, note that extra considerations are needed if the participants of a study are minors. These considerations for safety of the subject participants are relevant for adults as well. If a person is verbally asked to identify their sex and/or gender identity in say a doctor's waiting room, a transgender patient may feel vulnerable to harassment to others in the room. Thus, it is important that security and privacy measures are taken into account at the data collection stage by minimizing the opportunities for discrimination and harassment of the participants.

The privacy of personal information, especially in small studies, can leave minorities vulnerable to discovery if the information collected is too specific. For example, a transgender patient's identity may be identifiable from the data if it includes particular hospital information or information about a specific type of treatment. It is possible that gender and sex minorities may be identified even if a study adheres to HIPPA requirements [Corliss, 2015]. While a full discussion of data identifiability concerns is beyond the scope of this article, a tried-and-true statistical recommendation is to plan studies to avoid small sample size problems. This may require budgeting time or money to implement carefully designed sampling procedures. According to a 2018 Gallup poll, approximately 5% of the US adult population identify as LGBTQ+. Because gender and sex minorities represent a small proportion of the overall population, any survey collecting gender or sex data is likely to face the issue of unavoidably small sample sizes. In such instances, we recommend analyzing aggregated information rather than rely on a single survey for analysis. There are many statistical methods for conducting meta-analyses on aggregated data and the benefits of such methods may include higher statistical power in addition to stronger data privacy [Saha-Chaudhuri and Weinberg, 2017, Gatewood et al., 2001, Abowd, 2018].

5. DISCUSSION

LGBTQ+ individuals have always been part of society, even if these identities have not been encoded in previous data collection efforts. Much research concerns response variables that can be moderated by gender, sexual characteristics, or both. To prevent the continued erasure of minority populations from scientific research, it is time for the standards of data collection procedures to include options outside of the traditional gender and sex binary. Another important result of refining how gender

and sex data is collected will be a decrease in measurement error and bias in the results of statistical analyses.

A.   APPENDIX

For the interested reader, we have compiled some additional material that may be of use for implementing more inclusive data collection procedures. Example of inclusive demographic survey questions can be found on many survey sites such as Alchemer (https://www.alchemer.com/resources/blog/how-to-write-better-demographic-questions). There are also many consulting and training resources offered by companies such as Speaking of Transgender (https://speakingoftransgender.com/transgender-training-and-course) and Sylveon Consulting (https://www.sylveonconsulting.com). Writing tools and guidance are provided by many freelance professionals and professional groups such as transgender author Julia Serano (https://www.patreon.com/juliaserano/overview), The Society of Professional Journalists (https://www.spj.org/race-gender-hotline.asp), Radical Copy Editor

(https://radicalcopyeditor.com/2017/08/31/transgender-style-guide)

A wealth of additional educational material can be found on the websites for the Federal Committee on Statistical Methodology (https://nces.ed.gov/FCSM/SOGI.asp), Planned Parenthood (https://www.plannedparenthood.org/learn/gender-identity), The LGBT Foundation (https://lgbt.foundation/research-guidance), and The National LGBTQIA+ Health Education Center (https://www.lgbthealtheducation.org/resources/type/publication).

REFERENCES

J. M. Abowd. The us census bureau adopts differential privacy. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2867–2867, 2018.

C. Ainsworth. Sex redefined: The idea of 2 sexes is overly simplistic. *Scientific American*, 22, 2018.

J. Almeida, R. M. Johnson, H. L. Corliss, B. E. Molnar, and D. Azrael. Emotional distress among lgbt youth: The influence of perceived discrimination based on sexual orientation. *Journal of youth and adolescence*, 38:1001–1014, 2009.

T. Alzahrani, T. Nguyen, A. Ryan, A. Dwairy, J. McCaffrey, R. Yunus, J. Forgione, J. Krepp, C. Nagy, R. Mazhari, and J. Reiner. Cardiovascular disease risk factors and myocardial infarction in the transgender population. *Circulation: Cardiovascular Quality and Outcomes*, 12(4), 2019.

American Statistical Association. Ethical guidelines for statistical practice. Available at https://www.amstat.org/asa/files/pdfs/EthicalGuidelines.pdf, 2018. Committee on Professional Ethics.

M. Badgett and Sexual Minority Assessment Research Team. Best practices for asking questions about sexual orientation on surveys. *UCLA: The Williams Institute*, 2009.

N. Bates, Y. A. García Trejo, and M. Vines. Are sexual minorities hard-to-survey? insights from the 2020 census barriers, attitudes, and motivators study (cbams) survey. *Journal of Official Statistics*, 35(4):709–729, 2019.

G. Bauer, A. Devor, M. Heinz, Z. Marshall, A. Pullen Sansfaçon, and J. Pyne. Cpath ethical guidelines for research involving transgender people & communities. *Canadian Professional Association for Transgender Health*, 2019.

C. Bennouna, H. Mansourian, and L. Stark. Ethical considerations for children's participation in data collection activities during humanitarian emergencies: A delphi review. *Conflict and Health*, 11(1):1–15, 2017.

A. Brown. The changing categories the u.s. census has used to measure race. Pew Research Center, 2020.

J. M. Catalpa and J. K. McGuire. Family boundary ambiguity among transgender youth. *Family Relations*, 67(1):88–103, 2018.

H. Choi, K. Detyniecki, C. Bazil, S. Thornton, P. Crosta, H. Tolba, and M. Muneeb et al. Development and validation of a predictive model of drug-resistant genetic generalized epilepsy. *Neurology*, 95(15):e2150–e2160, 2020.

K. Clements-Nolle, R. Marx, and M. Katz. Attempted suicide among transgender persons: The influence of gender-based discrimination and victimization. *Journal of Homosexuality*, 51(3): 53–69, 2006.

D. J. Corliss. Analytics and protected healthcare information. Presentation for Michigan Association of Community Mental Health Boards https://www.peace-work.org/peace-work-studies, 2015.

US DHS. Handbook for safeguarding sensitive personally identifiable information, 2012.

L. Z. Dubois. Associations between transition-specific stress experience, nocturnal decline in ambulatory blood pressure, and c-reactive protein levels among transgender men. *American Journal of Human Biology*, 24(1):52–61, 2012.

L. Durwood, K. A. McLaughlin, and K. R. Olson. Mental health and self-worth in socially transitioned transgender youth. *Journal of the American Academy of Child and Adolescent Psychiatry*, 56(2):116–123, 2017.

A. Fausto-Sterling. *Sexing the body: Gender politics and the construction of sexuality*. Basic Books, 2000.

T. Fitzsimons. N.j. to become fourth state with gender-neutral birth certificate option. NBC News, 2019.

George Gatewood et al. A monograph on confidentiality and privacy in the us census. *US Bureau of the Census, Washington*, 2001.

Z. Goldstein, T. A. Corneil, and D. N. Greene. When gender identity doesn't equal sex recorded at birth: the role of the laboratory in providing effective healthcare to the transgender community. *Clinical chemistry*, 63(8):1342–1352, 2017.

J. A. Greenberg and R. Stam. *Intersexuality and the Law : Why Sex Matters*. 2012.

S. Gupta, K. L. Imborek, and M. D. Krasowski. Challenges in transgender healthcare: the pathology perspective. *Laboratory medicine*, 47(3):180–188, 2016.

G. Herdt. *Third sex, third gender: Beyond sexual dimorphism in culture and history*. Princeton University Press, 2020.

HEW. The belmont report: Ethical principles and guidelines for the protection of human subjects of research. Available at http://www.nuhresearch.nhg.com.sg/obr/Belmont%20Report.pdf, 1979.

I. A. Hughes, C. Houk, S. F. Ahmed, P. A. Lee, and Lawson Wilkins Pediatric Endocrine Society/European Society for Paediatric Endocrinology. Consensus statement on management of intersex disorders. *Journal of pediatric urology*, 2(3):148–162, 2006.

interACT. Statement on intersex terminology. Available at https://interactadvocates.org/interact-statement-on-intersex-terminology/, 2020.

M. S. Irwig. Cardiovascular health in transgender people. *Reviews in Endocrine and Metabolic Disorders*, 19(3):243–251, 2018.

F. Lau, M. Antonio, K. Davison, R. Queen, and A. Devor. A rapid review of gender, sex, and sexual orientation documentation in electronic health records. *Journal of the American Medical Informatics Association*, 27(11):1174–1783, 2020.

M. Morris and E. Cravigan. The privacy, data protection and cybersecurity law review: Australia. https://thelawreviews.co.uk/title/the-privacy-data-protection-and-cybersecurity-law-review/australia, 2020.

L. Mullen. gender: Predict gender from names using historical data. Available at https://github.com/ropensci/gender, 2018. R package version 0.5.2.

K. R. Olson and K. A. McLaughlin. Mental health of transgender children who are supported in their identities. *Pediatrics*, 137(3), 2016.

A. N. Onyiriuka, U. Kuhnle-Krahl, E. Sadoh, and J. B. E. Elusiyan. External genital anomalies in newborns in two nigerian hospitals: A pilot study of the birth prevalence. *International Journal of Child and Adolescent Health*, 9(2):187–193, 2016.

E. A. Pascoe and L. Smart Richman. Perceived discrimination and health: A meta-analytic review. *Psychological Bulletin*, 135(4):531–554, 2009.

D. W. Riggs and C. Bartholomaeus. Gaslighting in the context of clinical interactions with parents of transgender children. *Sexual and Relationship Therapy*, 33(4):382–394, 2018.

J. A. Roffee and A. Waling. Resolving ethical challenges when researching with minority and vulnerable populations: Lgbtiq victims of violence, harassment and bullying. *Research Ethics*, 13 (1):4–22, 2017.

S. T. Russell, A. M. Pollitt, G. Li, and A. H. Grossman. Chosen name use is linked to reduced depressive symptoms, suicidal ideation, and suicidal behavior among transgender youth. *The Journal of adolescent health : official publication of the Society for Adolescent Medicine*, 63(4): 503–505, 2018.

J. D. Safer, E. Coleman, J. Feldman, R. Garofalo, W. Hembree, A. Radix, and J. Sevelius. Barriers to health care for transgender individuals. *Current opinion in endocrinology, diabetes, and obesity*, 23(2):168–168, 2016.

P. Saha-Chaudhuri and C. R. Weinberg. Addressing data privacy in matched studies via virtual pooling. *BMC Medical Research Methodology*, 17(1), 2017.

D. Su, J. A. Irwin, C. Fisher, A. Ramos, M. Kelley, D. A. Rogel Mendoza, and J. D. Coleman. Mental health disparities within the lgbt population: A comparison between transgender and nontransgender individuals. *Transgender Health*, 1(1):12–20, 2016.

M. E. Wiesner-Hanks. *Gender in history: Global perspectives.* John Wiley & Sons, 2011.

A. L. Wirtz, T. C. Proteat, M. Malik, and N. Glass. Gender-based violence against transgender people in the united states: A call for research and programming. *Trauma, Violence, & Abuse*, 21(2):227–241, 2020.