Measurement of sex and gender identity in BRFSS

Jenine K Harris, Odalis Hernandez

3/21/2022

# Abstract

# Background

In 2021, a Gallup poll found that .6% or 60 out of every 1000 Americans, reported that they identify as transgender (Jones 2021). The percentage varied by age group, with the cohort born in 1997-2002 reporting 1.8% of survey participants identified as transgender compared to 1.2% in the previous cohort and just .3% in the cohort born before 1946. Transgender people are individuals whose biological sex, or sex recorded at birth, differs from their gender identity (Galupo, Henise, and Mercer 2016; Howerton and Harris 2021). Gender identity is how a person identifies and conceptualizes their own gender and is different from biological sex and sexual orientation (Galupo, Henise, and Mercer 2016; Charlotte Chucky Tate, Hagai, and Crosby 2020). Sex, sexual orientation, and gender identity all have implications for health, health care, and health outcomes (Legato, Johnson, and Manson 2016; Agénor et al. 2016; Gonzales, Przedworski, and Henning-Smith 2016; McClure et al. 2022).

Conducting research to understand factors that influence health, health care, and health outcomes is important for creating guidelines, policies, and programs that protect and improve health. Given the roles of sex, sexual orientation, and gender identity in health, many social science and health surveys seek to measure sex, and a growing number also collect sexual orientation and gender identity information. Until recently, measurement of gender identity was not commonly included as a survey item and measurement of sex was often done by an interviewer guessing the sex of the participant (Greta R. Bauer et al. 2017b) or asking “what is your sex?” and recording either male or female. Researchers have frequently treated gender as a synonym for sex (Krieger 2003; Johnson, Greaves, and Repta 2009; Charlotte Chuck Tate, Ledbetter, and Youssef 2013).

For decades there have been calls and attempts to improve the measurement of sex, gender identity, and sexual orientation (Krieger 2003; Charlotte Chuck Tate, Ledbetter, and Youssef 2013). In the meantime, researchers are utilizing data sets that include many different ways of measuring these concepts. One of these data sets that is commonly used is the Behavioral Risk Factor Surveillance Survey (BRFSS), which collects data on health behaviors, health conditions, and the use of health-related services from U.S residents. The Centers for Disease Control and Prevention collect BRFSS participant data each month over telephone. It is considered the world’s largest health survey system, collecting data from over 400,000 individuals yearly (Disease Control and Prevention, n.d.a).

Until 2014, BRFSS did not include questions about gender identity. In 2014, the sexual orientation and gender identity (SOGI) module was offered and 19 states included it in the survey this first year. The number of states including the SOGI module increased over time to 33 states in the most recent data release in 2020 (Disease Control and Prevention, n.d.b). The gender identity question included in the SOGI is: “Do you consider yourself to be transgender?” The follow-up question for those who answered yes was, “Do you consider yourself to be 1. Male-to-female, 2. Female-to-male, or 3. Gender nonconforming?” Prior to offering the optional SOGI module to all states, a few states included their own gender identity questions in some years (e.g., Massachusetts in 2007 to 2009 (K. J. Conron et al. 2012)).

Although a gender identity question was added, the wording of the sex question did not make any distinction between sex and gender for several years. In 2014 and 2015, the interviewer selected the sex of the participant given the prompt: “Indicate sex of respondent. Ask only if necessary.” In 2016, the prompt changed to, “Are you…” with the options of Male, Female, and Refused and a note, “Note: This may be populated from information derived from screening, household enumeration. However, interviewer should not make judgement on sex of respondent.” So, 2016 was the first year of the BRFSS that the interviewer was not permitted to guess the sex of the participant.

In 2017 the question remained, “Are you…” with responses of Male, Female, and Refused. The note for interviewers changed to all capital letters, presumably for emphasis, “INTERVIEWER NOTE: THIS QUESTION MUST BE ASKED EVEN IF INTERVIEWER HAD PREVIOUSLY ENTERED SEX IN THE SCREENING QUESTIONS. IT WILL NOT BE ASKED OF PERSONS WHO HAVE SELF-IDENTIFIED SEX IN LL HOUSEHOLD ENUMERATION.” There was also a second note, “CATI NOTE: THIS QUESTION MAY BE POPULATED BY LANDLINE HOUSEHOLD ENUMERATION ONLY. IT MAY NOT BE POPULATED BY INTERVIEWER ASSIGNMENT OF SEX DURING THE SCREENING FOR CELL PHONE OR PERSONS LIVING IN COLLEGE HOUSING.”

In 2018, the BRFSS interviewers asked, “What is your sex? or What was your sex at birth? Was it…” and gave the possible responses of Male and Female. It is unclear from the BRFSS documentation how interviewers selected one of the two questions. In 2019, the interviewer asked, “Are you male or female?”. In 2019 there was also a separate “Sex at Birth Optional Module” that included the question, “What was your sex at birth? Was it male or female?” The separate sex at birth question was asked of a subset of about 65 thousand participants and a combined sex variable was created that classified people based on their response to the sex at birth question if asked and, if sex at birth was not asked or if the reply was “Don’t know/Not Sure” or “Refused,” used the response to the sex question. In the 2020 survey administration, the 2019 sex at birth prompt was adopted as the sex question for all participants, “What was your sex at birth? Was it male or female?”

In a recent paper (Howerton and Harris 2021), Howerton and Harris found that BRFSS participants from 2015 to 2019 who reported having the same transgender gender identity sometimes had different answers for the sex question. For example, some participants who identified as female-to-male transgender identity reported their sex as male and others reported their sex as female. Likewise, among those who identified as male-to-female transgender identity, some reported their sex to be male and others reported their sex to be female. This has likely led to numerous published scientific papers examining sex differences or using sex as a covariate and potentially including participants whose sex recorded at birth was something other than their self-reported sex recorded in the BRFSS data.

The implications of how sex is measured impacts which questions participants are asked and how multiple variables are computed in the BRFSS. For example, the BRFSS variable for heavy drinking codes a male as a heavy drinker at 14 or more alcoholic drinks per week and a female as a heavy drinker at 7 or more alcoholic drinks per week. Questions related to pregnancy and pregnancy or birth-related health conditions are only asked of those who report being female and questions about human papilloma virus (HPV) screening and vaccination and birth control type and behavior differ for those who select male or female sex. Likewise, questions about mammograms, breast exams, and prostate specific antigen tests are only asked of people who report being in a specific sex category.

We sought to better understand the relationship between sex and gender identity responses in the BRFSS data from 2014 (the first year of the gender identity question) to the most recently released year of data in 2020. Specifically we answer three questions: (1) What is the relationship between sex and gender identity responses for participants over this six year period?; (2) How has the relationship between measured sex and gender identity changed as the sex question changed over this six year period?; and (3) Did the versions of the question that specified “at birth” result in different patterns of sex and gender identity reporting compared to other years?

# Methods

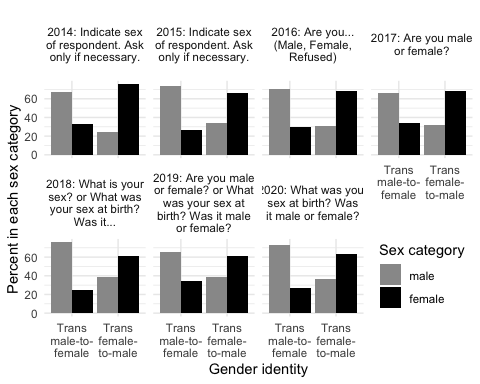
We imported the BRFSS data sets from 2014 through 2020 directly from the Centers for Disease Control and Prevention website. We used R statistical software version 4.1.1 (R Core Team 2021), the RStudio (version 1.4.1717) interactive development environment (RStudio Team 2020), and a function from version 2.4.3 of the haven package (Wickham and Miller 2021) to import the data. We used the tidyverse (Wickham 2017) software package to manage and visualize the data and the table1 package to compute descriptive statistics (Rich 2021).

# Results

## What is the relationship between sex and gender identity responses for participants over this six year period?

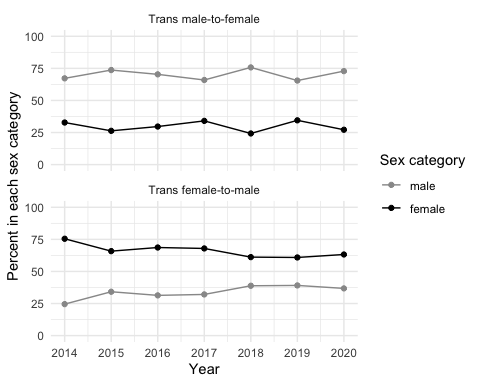
The majority of those who reported being male-to-female transgender also reported their sex as male each year (Figure 1). Likewise, the majority of those who reported being female-to-male transgender reported their sex as female. With the exception of 2017, there was a higher percentage of male-to-female transgender participants reporting themselves to be male sex compared to the percentage of female-to-male transgender participants reporting themselves to be female.

*Figure 1. Distribution of reported sex among BRFSS participants who reported being transgender (2014-2020).*



## How has the relationship between measured sex and gender identity changed as the sex question changed over this six year period?

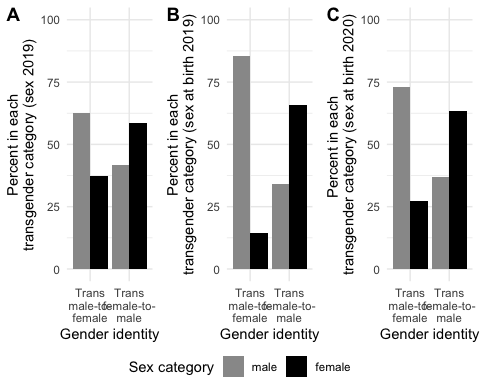
Although the sex question asked of participants changed multiple times from 2014 to 2020, the percentage in each transgender group choosing male and female remained relatively stable for both groups (Figure 2).



## Did the versions of the question that specified “at birth” result in different patterns of sex and gender identity reporting compared to other years?

The transgender male-to-female group ranged from 67% to 74% male sex with no apparent pattern given the question wording changes. The transgender female-to-male group ranged from 66% to 75% female sex with a pattern of decline in responding “female” to the sex question from 2014 to 2020.

The structure of the 2019 survey provides some additional information about differences in response patterns when including or not including “at birth” in a sex question. In 2019, there was more variability in responses among participants who were asked their sex (Figure 3A) compared to participants who were asked their sex at birth (Figure 3B). However, when the sex at birth version of the question was adopted for everyone in 2020 (Figure 3C), the distribution of responses was somewhere between the responses from those who responded to the sex question in 2019 and those who responded to the sex at birth question in 2019.



# Discussion

We found that about 70% of participants in the 2014 to 2020 BRFSS who identify as transgender male-to-female report male sex while about 30% report female sex. During the same time frame, participants who identify as transgender female-to-male report female sex 65% of the time and report male sex 35% of the time. These values were relatively consistent over the seven administrations of the survey; there were some minor shifts year-to-year in the sex distribution in the male-to-female group and a relatively steady increase in reporting male sex in the female-to-male group. One exception was for the participants who answered the optional module on gender identity in 2019. When asked their sex at birth, 85% of the male-to-female transgender group chose male. The value of the sex variable is used in computing some health related variables (e.g., heavy drinking) and in routing participants to questions about birth control, HPV vaccines, cancer screenings, and other preventive health behaviors.

The lack of *accurate* population-level epidemiological data for transgender people is problematic (Dente et al. 2019). For example, population-level research is lacking on breast cancer among transgender people, even though many take hormone replacement therapy which could influence breast cancer risk (Dente et al. 2019). In another example (Tabaac et al. 2018), three years of BRFSS data were used to determine colonoscopy and mammography screenings among transgender people but neglected to discuss the possible mismatch between the participant response to the sex variable and the types of screenings recommended. Likewise, a study using BRFSS data relied on the *Are you transgender?* question and the sex question to classify people as transgender men and transgender women (Tabaac et al. 2018) before analysis of the associations between a five category variable (cisgender men, cisgender women, transgender men, transgender women, and gender non-conforming) and several cancer screening tests including pap and PSA tests. While both transgender men and transgender women were asked about mammogram, transgender women were not asked about pap tests and transgender men were not asked bout PSA. In this case, with the way these groups were classified by the authors, those who reported male sex but transgender status female-to-male, and therefore may have a cervix and not a prostate (n = 212), would have been asked the PSA screening question and not the pap screening question. Likewise, those who reported female sex but male-to-female transgender status, and therefore may have a prostate and not a cervix (n = 319), would have been asked the pap screening question but not the PSA. So, out of the 1,779 transgender male-to-female or female-to-male participants, this strategy likely missed 212 people who likely had a cervix (Gatos 2018) but were not asked about a pap test and 319 of people who potentially had a prostate but were not asked about a PSA test.

Similar inadequate data storage in medical records is exacerbating the problem. For example, although most transgender men still have a cervix, they are less likely to be screened for cervical cancer compared to their cisgender counterparts (Gatos 2018) with one of the reasons being that some health systems base screening invitations on the sex listed health records (Connolly, Hughes, and Berner 2020). Lack of screening for transgender people can lead to later diagnoses and more severe disease Yang et al. (2018). It is clear that our existing tools are inadequate for collecting, storing, and using data effectively for transgender people.

In addition ineffective direct measurement of sex, the BRFSS includes computed variables that rely on the value of the sex variable to classify participants. Two of these variables are are *heavy drinking* and *binge drinking*. These measures classify participants as heavy drinkers or binge drinkers based on the number of drinks they report having in a week and their sex. They do not consider gender identity in classification; the threshold for heavy drinking is quite different for males (14 drinks per week) compared to females (7 drinks per week), while the binge drinking definition differs by a single drink per episode (5 for males, 4 for females). Given participant responses for sex vary within a single gender identity, these thresholds may not be accurately capturing heavy or binge drinking. In one example of how this impacts research, Azagba et al. (Azagba, Latham, and Shan 2019) did not include sex as a covariate in logistic regression models computing the odds of binge drinking for different transgender groups compared to cisgender groups. Perhaps not surprising, the findings suggested that female-to-male transgender participants have a lower odds of binge drinking compared to cisgender males and male-to-female transgender participants had higher odds of binge drinking compared to cisgender females (Azagba, Latham, and Shan 2019). Given the importance of understanding alcohol consumption in the transgender population (Gilbert et al. 2018), a way of classifying hazardous consumption that takes into consideration physical attributes rather than relying solely on sex could improve accuracy of this measure and thus improve the quality of research on alcohol use among transgender people.

There are at least two possible ways to reduce the current limitations on accurate population level data for transgender people. First, the BRFSS (and other data collection and storage tools) could build in the flexibility to ask all of the questions currently asked of only male sex or only female sex for any person who identifies as transgender. In this way we can ensure more accurate information about important cancer screenings, health conditions, and health behaviors among transgender people. Second, researchers including the sex variable in their work should pay attention to measures of gender identity if they are available and, if gender identity measures are not available, might report on this limitation given that data on reported sex for transgender people is likely a mixture of sex recorded at birth and gender identity.

More flexible and accurate ways of measurement of sex and gender identity are also necessary to protect and improve the health of transgender people (Beischel et al. 2022; Greta R. Bauer et al. 2017a). Current ways of asking gender identity include single questions with a transgender option and two-step questions typically asking about sex at birth and gender identity separately (Becker 2022). While the two-step questions provide more detailed information and has been suggested by some groups as a best practice (K. Conron, Lombardi, and Reisner 2014), they are not perfect as. Numerous studies suggest drawbacks including a small number of people who would fit the definition of transgender but who do not identify as transgender and a small number of cisgender participants responding that they are transgender because they do not understand the terminology (Becker 2022). One study also found evidence that some transgender people may misunderstand trangender terminology, with a small number of transgender women self-identifying as transgender men (Costa, Rosa, and Fontanari 2022). In our study, the two-step question using the “at birth” terminology in 2020 did not result in notable changes in the reporting of sex and gender identity from prior years.

While there does not seem to be consensus yet on the best way to measure sex and gender, scientists from different fields continue to examine the data and develop guidelines. For example, a 2017 study (Greta R. Bauer et al. 2017a) examined different ways of measuring sex and gender and consulted with experts to identify a set of three questions, “What sex were you assigned at birth, meaning on your original birth certificate?”; “Which best describes your current gender identity?”; and, for those whose gender identity differs from birth sex, “What gender do you currently live as in your day-to-day life?” A 2021 paper suggests gender identity be measured by two questions, “Which of the following best describes your gender identity? Woman, Man, Non-binary, genderfluid, or gender non-conforming, or Other (please specify)” and “Do you identify as transgender? Yes, No, Other (please specify)” (Kosciesza 2022). Likewise, the National Academies of Sciences, Engineering, and Medicine recommended a two question set for sex and gender identity, “What sex were you assigned at birth, on your original birth certificate? Female, Male, Don’t know, Prefer not to answer” and “What is your current gender? Female, Male, Transgender, (if respondent is American Indian / Alaskan Native) Two-Spirit, I use a different term (free test), Don’t know, Prefer not to answer” (Becker 2022). In contrast, Thornton et al. (Thornton et al. 2022) argue that, as a step toward best practice for gender and sex data collection, ethical measurement could skip the sex question as a filter and use questions that are relevant to a study topic like “Do you have a prostate?” or “Are you breastfeeding?”

Whether one of these options is the key to more effective and accurate data collection remains to be seen. What is clear, though, is that ineffective data collection, storage, and use is a threat to the health of transgender people and researchers using these data should be aware of its limitations. More accurate, flexible, and inclusive measurement along with data systems that are more flexible for data collection, storage, and use are strategies that could improve the health and well-being of transgender people.

# References

Agénor, Madina, S Bryn Austin, Daniel Kort, Erika L Austin, and Christina A Muzny. 2016. “Sexual Orientation and Sexual and Reproductive Health Among African American Sexual Minority Women in the US South.” *Women’s Health Issues* 26 (6): 612–21.

Azagba, Sunday, Keely Latham, and Lingpeng Shan. 2019. “Cigarette, Smokeless Tobacco, and Alcohol Use Among Transgender Adults in the United States.” *International Journal of Drug Policy* 73 (November): 163–69. <https://doi.org/10.1016/j.drugpo.2019.07.024>.

Bauer, Greta R., Jessica Braimoh, Ayden I. Scheim, and Christoffer Dharma. 2017a. “Transgender-Inclusive Measures of Sex/Gender for Population Surveys: Mixed-Methods Evaluation and Recommendations.” Edited by Andrew R. Dalby. *PLOS ONE* 12 (5): e0178043. <https://doi.org/10.1371/journal.pone.0178043>.

Bauer, Greta R, Jessica Braimoh, Ayden I Scheim, and Christoffer Dharma. 2017b. “Transgender-Inclusive Measures of Sex/Gender for Population Surveys: Mixed-Methods Evaluation and Recommendations.” *PloS One* 12 (5): e0178043.

Becker, Tara, ed. 2022. *Measuring Sex, Gender Identity, and Sexual Orientation*. National Academies Press. <https://doi.org/10.17226/26424>.

Beischel, Will J., Zach C. Schudson, Rhea Ashley Hoskin, and Sari M. van Anders. 2022. “The Gender/Sex 3×3: Measuring and Categorizing Gender/Sex Beyond Binaries.” *Psychology of Sexual Orientation and Gender Diversity*, February. <https://doi.org/10.1037/sgd0000558>.

Bertoncelli Tanaka, Mariana, Kirpal Sahota, James Burn, Alison Falconer, Mathias Winkler, Hashim U. Ahmed, and Tina G. Rashid. 2021. “Prostate Cancer in Transgender Women: What Does a Urologist Need to Know?” *BJU International* 129 (1): 113–22. <https://doi.org/10.1111/bju.15521>.

Connolly, Dean, Xan Hughes, and Alison Berner. 2020. “Barriers and Facilitators to Cervical Cancer Screening Among Transgender Men and Non-Binary People with a Cervix: A Systematic Narrative Review.” *Preventive Medicine* 135 (June): 106071. <https://doi.org/10.1016/j.ypmed.2020.106071>.

Conron, Keith, Emilia Lombardi, and SL Reisner. 2014. “Identifying Transgender and Other Gender Minority Respondents on Population-Based Surveys: Approaches.” *Best Practices for Asking Questions to Identify Transgender and Other Gender Minority Respondents on Population-Based Surveys. Los Angeles, Ca2014*, 9–18.

Conron, Kerith J., Gunner Scott, Grace Sterling Stowell, and Stewart J. Landers. 2012. “Transgender Health in Massachusetts: Results From a Household Probability Sample of Adults.” *American Journal of Public Health* 102 (1): 118–22. <https://doi.org/10.2105/ajph.2011.300315>.

Costa, Angelo Brandelli, Letícia de Oliveira Rosa, and Anna Martha Vaitses Fontanari. 2022. “Sampling Bias in Brazilian Studies on Transgender and Gender Diverse Populations: The Two-Step Measure for Assessing Gender Identity in Surveys.” *Cadernos de Saúde Pública* 38 (1). <https://doi.org/10.1590/0102-311x00180420>.

Dente, Elizabeth, Richard Farneth, Jennifer Purks, and Stefan Torelli. 2019. “Evaluating Risks, Reported Cases and Screening Recommendations for Breast Cancer in Transgender Patients.” *Georgetown Medical Review*, May. <https://doi.org/10.52504/001c.7774>.

Disease Control, Centers for, and Prevention. n.d.a. “"Behavioral Risk Factor Surveillance System".” <https://www.cdc.gov/brfss/>.

———. n.d.b. “"CDC – BRFSS – 2020 BRFSS Modules Used by Category".” <https://www.cdc.gov/brfss/questionnaires/modules/category2020.htm>.

Galupo, M Paz, Shane B Henise, and Nicholas L Mercer. 2016. “‘The Labels Don’t Work Very Well’: Transgender Individuals’ Conceptualizations of Sexual Orientation and Sexual Identity.” *International Journal of Transgenderism* 17 (2): 93–104.

Gatos, Kayla C. 2018. “A Literature Review of Cervical Cancer Screening in Transgender Men.” *Nursing for Women’s Health* 22 (1): 52–62. <https://doi.org/10.1016/j.nwh.2017.12.008>.

Gilbert, Paul A., Lauren E. Pass, Alex S. Keuroghlian, Tom K. Greenfield, and Sari L. Reisner. 2018. “Alcohol Research with Transgender Populations: A Systematic Review and Recommendations to Strengthen Future Studies.” *Drug and Alcohol Dependence* 186 (May): 138–46. <https://doi.org/10.1016/j.drugalcdep.2018.01.016>.

Gonzales, Gilbert, Julia Przedworski, and Carrie Henning-Smith. 2016. “Comparison of Health and Health Risk Factors Between Lesbian, Gay, and Bisexual Adults and Heterosexual Adults in the United States: Results from the National Health Interview Survey.” *JAMA Internal Medicine* 176 (9): 1344–51.

Howerton, Isabelle, and Jenine K. Harris. 2021. “Transgender Identity and Cardiovascular Disease.” *Transgender Health*, August. <https://doi.org/10.1089/trgh.2020.0188>.

Johnson, Joy L, Lorraine Greaves, and Robin Repta. 2009. “Better Science with Sex and Gender: Facilitating the Use of a Sex and Gender-Based Analysis in Health Research.” *International Journal for Equity in Health* 8 (1): 1–11.

Jones, Jeffery. 2021. “"LGBT Identification Rises to 5.6.” *Gallup*. <https://news.gallup.com/poll/329708/lgbt-identification-rises-latest-estimate.aspx>.

Kosciesza, Aiden James. 2022. “Intersectional Gender Measurement: Proposing a New Metric for Gender Identity and Gender Experience.” *Feminist Media Studies*, January, 1–16. <https://doi.org/10.1080/14680777.2021.2018008>.

Krieger, Nancy. 2003. “Genders, Sexes, and Health: What Are the Connections—and Why Does It Matter?” *International Journal of Epidemiology* 32 (4): 652–57.

Legato, Marianne J, Paula A Johnson, and JoAnn E Manson. 2016. “Consideration of Sex Differences in Medicine to Improve Health Care and Patient Outcomes.” *Jama* 316 (18): 1865–66.

McClure, Robert C, Caroline L Macumber, Clair Kronk, Chris Grasso, Robert J Horn, Roz Queen, Steven Posnack, and Kelly Davison. 2022. “Gender Harmony: Improved Standards to Support Affirmative Care of Gender-Marginalized People Through Inclusive Gender and Sex Representation.” *Journal of the American Medical Informatics Association* 29 (2): 354–63.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.

Rich, Benjamin. 2021. “Table1: Tables of Descriptive Statistics in HTML.” <https://CRAN.R-project.org/package=table1>.

RStudio Team. 2020. *RStudio: Integrated Development Environment for r*. Boston, MA: RStudio, PBC. <http://www.rstudio.com/>.

Tabaac, Ariella R., Megan E. Sutter, Catherine S. J. Wall, and Kellan E. Baker. 2018. “Gender Identity Disparities in Cancer Screening Behaviors.” *American Journal of Preventive Medicine* 54 (3): 385–93. <https://doi.org/10.1016/j.amepre.2017.11.009>.

Tate, Charlotte Chuck, Jay N Ledbetter, and Cris P Youssef. 2013. “A Two-Question Method for Assessing Gender Categories in the Social and Medical Sciences.” *Journal of Sex Research* 50 (8): 767–76.

Tate, Charlotte Chucky, Ella Ben Hagai, and Faye J Crosby. 2020. *Undoing the Gender Binary*. Cambridge University Press.

Thornton, Suzanne, Dooti Roy, Stephen Parry, Donna LaLonde, Wendy Martinez, Renee Ellis, and David Corliss. 2022. “Towards Statistical Best Practices for Gender and Sex Data.” *Significance* 19 (1): 40–45. <https://doi.org/10.1111/1740-9713.01614>.

Wickham, Hadley. 2017. “The Tidyverse.” *R Package Ver* 1 (1).

Wickham, Hadley, and Evan Miller. 2021. *Haven: Import and Export ’SPSS’, ’Stata’ and ’SAS’ Files*. <https://CRAN.R-project.org/package=haven>.

Yang, Daniel X., Pamela R. Soulos, Brigette Davis, Cary P. Gross, and James B. Yu. 2018. “Impact of Widespread Cervical Cancer Screening.” *American Journal of Clinical Oncology* 41 (3): 289–94. <https://doi.org/10.1097/coc.0000000000000264>.