Misclassification of Sex Assigned at Birth in the Behavioral Risk Factor Surveillance System and Transgender Reproductive Health

A Quantitative Bias Analysis

Diana Tordoff,[a] *Michele Andrasik,*[b,c] *and Anjum Hajat*[a]

Background: National surveys based on probability sampling methods, such as the Behavioral Risk Factor and Surveillance System (BRFSS), are crucial tools for unbiased estimates of health disparities. In 2014, the BRFSS began offering a module to capture transgender and gender nonconforming identity. Although the BRFSS provides much needed data on the this population, these respondents are vulnerable to misclassification of sex assigned at birth.

Methods: We applied quantitative bias analysis to explore the magnitude and direction of the systematic bias present as a result of this misclassification. We use multivariate Poisson regression with robust standard errors to estimate the association between gender and four sex-specific outcomes: prostate-specific antigen testing, Pap testing, hysterectomy, and pregnancy. We applied single and multiple imputation methods, and probabilistic adjustments to explore bias present in these estimates.

Results: Combined BRFSS data from 2014, 2015, and 2016 included 1078 transgender women, 701 transgender men, and 450 gender nonconforming individuals. Sex assigned at birth was misclassified among 29.6% of transgender women and 30.2% of transgender men. Transgender and gender nonconforming individuals excluded due to sex-based skip patterns are demographically distinct from those who were asked reproductive health questions, suggesting that there is noteworthy selection bias present in the data. Estimates for gender nonconforming respondents are vulnerable to small degrees of bias, while estimates for cancer screenings among transgender women and men are more robust to moderate degrees of bias.

Conclusion: Our results demonstrate that the BRFSS methodology introduces substantial uncertainty into reproductive health measures, which could bias population-based estimates. These findings emphasize the importance of implementing validated sex and gender questions in health surveillance surveys. See video abstract at, http://links.lww.com/EDE/B562.

Keywords: Gender minority; Gender nonconforming; Misclassification; Multiple imputation; Quantitative bias analysis; Reproductive health; Selection bias; Transgender

(*Epidemiology* 2019;30: 669–678)

National surveys based on probability sampling methods, such as the Behavioral Risk Factor and Surveillance System (BRFSS), are crucial tools in accurately measuring health disparities among transgender and gender nonconforming populations.[1–3] Until recently, nationally representative survey data have not included gender identity questions.[4] Consequently, the majority of existing studies on transgender health rely on non-representative community and clinical samples.[1,5] Beginning in 2014, the BRFSS began offering a module of sexual orientation and gender identity that each state could optionally include, and a growing number of studies have used these data to investigate transgender and gender nonconforming populations and health disparities.[1,6–12]

Although the BRFSS provides much-needed data on transgender populations, it is vulnerable to misclassification of sex assigned at birth among transgender respondents.[13,14] Transgender and gender nonconforming people have a gender identity that differs from the sex they were assigned at birth, and includes gender nonconforming individuals who do not identify with binary gender categories (e.g., "man" and "woman"). Validated trans-inclusive measures of sex and gender therefore recommend asking two questions to separately ascertain an individual's sex assigned at birth and current gender (the "two-step" methodology).[13,15] The use of the two-step question is endorsed by the Institute of Medicine, The Williams Institute, Fenway Health, and the "Meaningful Use" guidelines published by the Department of Health and Human Services.[16–19]

Submitted July 11, 2018; accepted May 23, 2019.
From the [a]Department of Epidemiology, University of Washington, Seattle, WA; [b]Department of Global Health, University of Washington, Seattle, WA; and [c]HIV Vaccine Trials Network, Fred Hutchinson Cancer Research Center, Seattle, WA.
The authors report no conflicts of interest.
SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).
Data are publicly available at: https://www.cdc.gov/brfss/annual_data/annual_data.htm. Analytic code will be made available on author's GitHub upon publication.
Correspondence: Diana Tordoff, University of Washington, 1959 NE Pacific Street, Health Sciences Bldg F-262, Box 357236, Seattle, WA 98195. E-mail: dtordoff@uw.edu.

ISSN: 1044-3983/19/3005-0669
DOI: 10.1097/EDE.0000000000001046

However, no health surveillance surveys to date, including the BRFSS, have adopted the recommended two-step methodology.[15] The BRFSS ascertains sex assigned at birth based on the sound of the participant's voice and through questions that conflate sex and gender (eAppendix 1A; http://links.lww.com/EDE/B542). As a result, a previous study found that the BRFSS misclassified sex assigned at birth in 30% of individuals who self-identified as transgender or gender nonconforming through the sexual orientation and gender identity module.[14]

Additionally, the BRFSS makes binary cisnormative assumptions (i.e., that all respondents are cisgender) regarding anatomy, implemented in their protocol as skip patterns based on BRFSS designated sex. Subsequently, a large portion of transgender and gender nonconforming respondents are precluded from answering questions related to their reproductive health (Figure). Current guidelines for cancer screening and primary healthcare for transgender and gender nonconforming people recommend an organ-based approach, providing care for an individual based on their anatomical structures.[20] Thus, a transgender man may still retain a uterus and cervix. If he was designated as male by the current BRFSS methodology, then as a result of skip patterns, this person would not be asked questions relevant to his health, such as Papanicolau (Pap) testing and pregnancy. Similarly, a transgender woman would be precluded from questions about prostate-specific antigen (PSA) testing if she was designated female by the BRFSS methodology. However, sex assigned at birth is an imperfect proxy for an individual's current anatomy as some, but not all, transgender individuals have a history of gender affirming surgical interventions.[21] Consequently, it is likely that the BRFSS does not adequately capture rates of screening or the prevalence of other reproductive health outcomes among transgender and gender nonconforming individuals due to conflicting and missing data. Despite these limitations, no studies examine the impact of these biases on assessing transgender reproductive health.[11,12]

Quantitative bias analysis is a method of modeling systematic, or nonrandom error, that can bias the results of epidemiologic research, including issues of misclassification, missing data, and selection bias.[22,23] This study applied multiple bias modeling methods to explore the influence of missing data that results from inappropriately excluding individuals as a result of misclassified sex assigned at birth. We quantified the magnitude and directionality of the bias on the estimated association between gender and four outcomes related to specific reproductive anatomy: PSA testing, Pap testing, pregnancy, and history of a hysterectomy.

METHODS

Data

Data for this study are from the 2014, 2015, and 2016 BRFSS, a state-based system of telephone health surveys overseen by the Centers for Disease Control and Prevention. Eligible participants are non-institutionalized adults aged 18 or over who live in the United States. The BRFSS uses complex probability sampling so that data are collected from a representative sample within each state.[24] All analyses used pooled, weighted data from all states that participated in the sexual orientation and gender identity module (20 states in

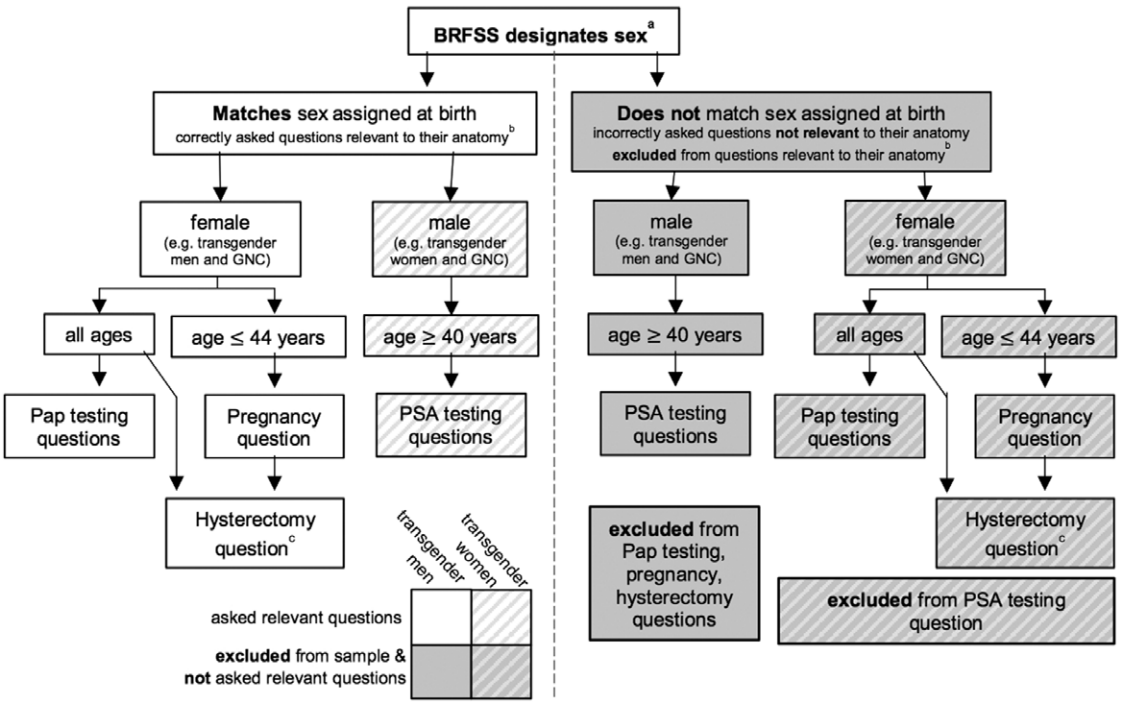


FIGURE. Schematic of sex-based skip patterns in the BRFSS.

2014, 21 in 2015, and 26 in 2016; eAppendix 1A; http://links.lww.com/EDE/B542). This research was deemed to be exempt from human subjects research status by the University of Washington Institutional Review Board.

Measures

The BRFSS protocol designates a binary sex variable (male, female), (1) based on the sound of the respondent's voice, (2) during eligibility screening and household enumeration, or (3) through demographic questions (eAppendix 1A; http://links.lww.com/EDE/B542). The interview script does not distinguish between sex assigned at birth and gender identity, for example, conflating individuals who are men with male sex assigned at birth.[14] Therefore, this measure, which we will refer to as the BRFSS designated sex, is not a valid or reliable measure of sex assigned at birth among transgender respondents. Transgender and gender nonconforming identity were ascertained through the sexual orientation and gender identity module, which asks "Do you consider yourself to be transgender?" Individuals who responded *no* are considered cisgender. Individuals who identified as *transgender, male-to-female* are further referred to as transgender women, and we assumed that their sex assigned at birth was male, regardless of how their sex was designated in the BRFSS data. Similarly, individuals who identified as *transgender, female-to-male* are further referred to as transgender men, and we assumed that their sex assigned at birth was female. We are unable to infer the sex assigned at birth of individuals who identified as gender nonconforming. Individuals who responded as *don't know/not sure* or *refused* are excluded from all analyses.

We considered four reproductive health outcomes that are subject to skip patterns based on BRFSS designated sex (e.g., sex-specific outcomes): lifetime PSA testing, lifetime Pap testing, history of hysterectomy, and currently pregnant. We compared the prevalence of never having had a PSA or Pap test, to ever having had a test, and compared the prevalence of having had a hysterectomy to not having had one, and currently being pregnant to not currently pregnant. We also considered demographic characteristics and measures of healthcare access including: age (5-year strata), race/ethnicity (White, Black, American Indian/Alaska Native, Asian, Native Hawaiian/Pacific Islander, Hispanic, multiracial), partnership status (married or living with a partner, separated/divorced/widowed, or never married), sexual orientation (straight, lesbian or gay, bisexual, other), socioeconomic status (measured by unemployment status, annual income, and educational attainment), and self-reported poor health. Healthcare access was measured by three variables: lack of health insurance, no primary healthcare provider, and inability to see a doctor due to cost within the last year (eAppendix 1B; http://links.lww.com/EDE/B542).

Analysis

This study aimed to quantify the bias in the association between transgender and gender nonconforming identity and four outcomes vulnerable to misclassification of sex assigned at birth. We first estimated the association as observed in the data using complete cases analysis. We used weighted Poisson regression with robust standard errors and BRFSS developed sampling weights to estimate crude and adjusted prevalence ratios (PRs) for each outcome. The analytic model for never having had a PSA test compared transgender women and gender nonconforming individuals to cisgender men. The analytic models for currently being pregnant, having had a hysterectomy, and never having a Pap test compared transgender men and gender nonconforming individuals to cisgender women. Age, race/ethnicity, and partnership status were identified *a priori* as confounders, and were included in the model as covariates. Socioeconomic status, general health status, and healthcare coverage are identified *a priori* as either potential confounders or mediators; therefore, they were included as covariates in the primary analytic model. We conducted sensitivity analyses that compared the fully adjusted analytic model to a minimally adjusted analytic model that only included age, race, and partnership status (eAppendix 3; http://links.lww.com/EDE/B542). The fully adjusted model is presented as the primary analysis. We provide counts and survey weighted proportions of these health outcomes by gender, as well as demographic characteristics of respondents by BRFSS designated sex (eAppendix 2; http://links.lww.com/EDE/B542).

Bias Modeling

The above approach accounts for random error, in the form of estimated robust standard errors and confidence intervals (CIs), and also adjusts for systematic bias in that we have controlled for several confounders in the multivariate analyses. However, it fails to account for systematic bias that results from the misclassification of sex assigned at birth and any residual confounding that may be present. This is addressed through bias modeling, which assessed how nonrandom bias impacts uncertainty in the associations found in the complete-case analysis.

Misclassification and subsequent skip patterns based on BRFSS designated sex result in missing data and nonrepresentative samples of transgender and gender nonconforming individuals asked questions related to specific reproductive anatomy (e.g., the uterus, cervix, and prostate).[14] This has two consequences: First, transgender individuals are asked questions not relevant to their anatomy (cell C in Table 1), resulting in an inflation of *no, I don't know,* and *refused* responses. Second, individuals who may retain anatomical structures (e.g., a uterus) relevant to a question (e.g., hysterectomy), are excluded from the sample (cell D in Table 1). For the purposes of bias modeling, we used sex assigned at birth as an imperfect proxy for anatomical structures.[13] We assumed individuals assigned male at birth do not experience pregnancy or hysterectomy, and are not eligible for cervical Pap testing. We similarly assumed that individuals assigned female at birth do not have a prostate and are not eligible for PSA testing.

TABLE 1. Misclassification of Sex Assigned at Birth and the Subsequent Transgender and Gender Nonconforming Respondents Who Were Asked and Precluded from Reproductive Health Questions

	Asked Sex-specific Questions	Not Asked Sex-specific Questions
Sex assigned at birth *does* match BRFSS designated sex	(A) TGNC respondents are *correctly included* in the sample	(B) TGNC respondents are *correctly excluded* from the sample
Sex assigned at birth *does not* match BRFSS designated sex	(C) TGNC respondents are *incorrectly included* in the sample	(D) TGNC respondents are *incorrectly excluded* from the sample

TGNC, transgender and gender nonconforming.

We compared two approaches to multiple bias modeling (Table 2). We first considered a selection bias model. This approach benefits from its simplicity but does not incorporate the known demographic and health information of the 30% of transgender individuals who were excluded. Therefore, we considered a second modeling approach that relies on imputation methods. Both models implemented bias adjustments in two discrete steps: Record level adjustments were performed first, after which we estimated record-level adjusted $\widehat{\text{PR}}$ using the same analytic multivariate Poisson model specified above. Last, we conducted summary-level probabilistic adjustments using Monte Carlo methods and estimated bias-adjusted PRs that combine both systematic bias and random errors.[23]

Approach One

The directed acyclic graph (DAG) in Table 2 demonstrates how the reproductive health outcomes depend on BRFSS designated sex through inclusion in the subsample of respondents asked sex-specific questions. At the record-level, we excluded individuals who are incorrectly included in the sample. We were unable to infer the sex-assigned a birth for gender nonconforming individuals. Therefore, we conducted a probabilistic adjustment using Monte Carlo methods (1000 randomly sampled scenarios) to exclude a random subset of gender nonconforming individuals who responded as *no, I don't know,* or *refused,* under the assumption that they are individuals for whom the sex-specific questions were not anatomically relevant. We assumed misclassification occurs at a rate similar to what is observed among transgender men and women, that is, 30% as per Riley et al.[14]

We then performed a summary-level selection bias adjustment to account for transgender and gender nonconforming respondents incorrectly excluded from sex-specific questions. The estimated $\widehat{\text{PR}}$ is multiplied by a bias adjustment factor, the selection bias odds ratio OR_{select}. We defined the PR adjusted for systematic bias as $\text{PR}_{adj} = \widehat{\text{PR}} \times \text{OR}_{select}$.[23] In the absence of literature to inform the direction or degree of selection bias, we modeled a large range of selection bias scenarios and randomly sample 100,000 OR_{select} from a uniform distribution. We allowed OR_{select} to range between 0.10 to 10.0, where a OR_{select} of one models no selection bias (eAppendix 4; http://links.lww.com/EDE/B542). Last, we determined the maximum OR_{select} that guaranteed a PR less than one, and the minimum OR_{select} that guaranteed a PR greater than one. The choice of a uniform distribution means that each OR_{select} had an equal probability of being sampled. Although in practice a modest OR_{select} of two is more likely than an extreme OR_{select} of 10, the aim of this analysis was to determine the sensitivity of the associations to selection bias through the provision of a wide range of possible values.

Approach Two

We model the systematic bias as a missing data problem, whereby transgender and gender nonconforming individuals with misclassified sex assigned at birth are missing data on certain reproductive health questions. First, in order to consider the full range of possible estimates that are observable in the data, we considered two single-imputation models of these extreme scenarios in which data can be missing. For all transgender respondents with missing data, we first assumed that all excluded individuals had the outcome and imputed a value of one for each relevant outcome. Then, we assumed that all excluded transgender individuals did not have the outcome, and imputed a value of 0. For each scenario, we estimated $\widehat{\text{PR}}$, using the same analytic model.

Multiple imputation by chained equations is a statistical technique for handling missing data, whereby the observed data are used to jointly estimate plausible values for missing observations.[25–27] We used this method to impute the missing outcome variables due to the method's flexibility to accommodate categorical variables and skip patterns. For each sex-specific outcome we created 30 imputed datasets, assessed trace plots for evidence of convergence and good fit, and estimated record-level adjusted $\widehat{\text{PR}}$ on the imputed datasets.[28] As in approach one, we excluded data based on inferred sex assigned at birth. For gender nonconforming individuals, we were unable to condition the imputation model on an inferred sex assigned at birth. Imputing sex-specific outcomes for all gender nonconforming individuals would inappropriately rely on information from individuals without relevant anatomy, yielding unreliable estimates. Therefore, multiple imputation by chained equations was only applied to model bias among transgender men and women, and cisgender respondents.

Multiple imputation by chained equations relies on the assumption that data are missing at random (MAR).[25,26] However, there likely is an underlying non-ignorable process driving misclassification of sex assigned at birth and subsequent missingness.[13,14,25,26] Therefore, we adjusted the multiple imputation $\widehat{\text{PR}}$ for unknown confounding. We defined the risk ratio due to confounding as RR_{conf}, and by convention, the PR adjusted for confounding is $\text{PR}_{adj} = \frac{\widehat{\text{PR}}}{\text{RR}_{conf}}$.[23] As in approach one, in the absence of literature to inform the direction or degree of confounding, we assigned a uniform distribution and allow RR_{conf} to range between 0.10 and 10.0, where a RR_{conf} of one models no confounding.[29] We randomly sample 100,000 confounding scenarios and estimated bias-adjusted PR that combine both systematic bias and random errors (eAppendix 5; http://links.lww.com/EDE/B542). As before, we determined the maximum OR_{select} that guaranteed a PR greater than one, and the minimum OR_{select} that guaranteed a PR less than one.

We conducted statistical analyses and multiple imputation in Stata version 15.1 (StataCorp, College Station, Texas) and Monte Carlo simulations in R version 3.4.2 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Combining BRFSS data from 2014, 2015, and 2016, a total of 518,982 participants responded to the gender identity question and were included in our sample. The data include 1078 transgender women, 701 transgender men, and 450 gender nonconforming individuals; 2970 participants (0.6%) responded *don't know/not sure* and 4286 respondents (0.8%) refused to answer the gender identity question and were excluded from the analysis.

Congruent with past studies based only on 2014 data, sex assigned at birth is presumed to be misclassified among 29.6% (319/1078) of transgender women and 30.2% (212/701) of transgender men. eAppendix 2; http://links.lww.com/EDE/B542 presents demographic characteristics and measures of healthcare access, comparing individuals who were asked and precluded from sex-specific questions. These results suggest that individuals who were excluded from the sample as a result of misclassification of sex assigned at birth are demographically distinct from those who were included in the sample. Responses to the four sex-specific questions by gender identity are reported in Table 3. These proportions, however, are subject to selection bias and may not be representative of the transgender and gender nonconforming population.

TABLE 2. Overview of Multiple-bias Modeling Approaches

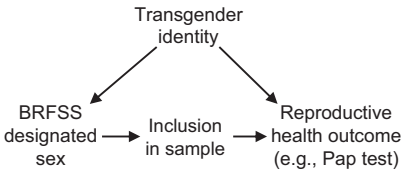
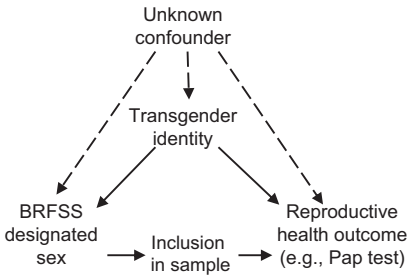
	Approach One: Selection Bias	**Approach Two: Missing Data and Unknown Confounding**
Directed acyclic graph (DAG)		
Analytic step 1: record-level adjustments	Remove transgender and gender nonconforming individuals incorrectly included in the sample (cell C in Table 1): For transgender men and women, exclude individuals based on their inferred sex assigned at birth. For gender nonconforming individuals, use Monte Carlo sampling methods to remove 30% of the inflated *no, I don't know,* and *refused* responses.	First, use single-imputation to model extreme scenarios in which transgender and gender nonconforming individuals missing data either all had the outcome, or all did not have the outcome. Second, use multiple imputation by chained equations to impute missing outcomes for individuals who were incorrectly excluded from the sample (cell D in Table 1). Remove transgender and gender nonconforming individuals incorrectly included in the sample using similar methods as approach one.
Analytic step 2: summary-level adjustments	Conduct a probabilistic adjustment for selection bias that results from individuals being incorrectly excluded from the sample (cell D in Table 1). Using Monte Carlo sampling methods, multiply the estimated $\widehat{\text{PR}}$, by the selection bias odds ratio, OR_{select}, to obtain a measure of association that accounts for systematic bias: $\text{PR}_{adj} = \widehat{\text{PR}} \times \text{OR}_{select}$	Conduct a probabilistic adjustment to multiple imputation based estimates to further account for the unknown confounding that underlies the missingness (i.e. because sex-specific outcomes have non-ignorable missingness). Using Monte Carlo sampling methods, divide the estimated $\widehat{\text{PR}}$, by the risk ratio due to confounding, RR_{conf}, to obtain a measure of association that accounts for systematic bias: $\text{PR}_{adj} = \frac{\widehat{\text{PR}}}{\text{RR}_{conf}}$

TABLE 3. Responses to Reproductive Health Questions by Gender Identity, BRFSS 2014–2016

	Transgender Women	Transgender Men	Gender Nonconforming	Cisgender Women	Cisgender Men
Total (N)	1078	701	450	298,391	218,362
Lifetime PSA test, N (%)[a]					
Yes	201 (44)		48 (35)		68,677 (49)
No	178 (52)		41 (59)		45,506 (44)
Don't know	14 (1)		2 (1)		5058 (4)
Refused	2 (3)		1 (5)		573 (3)
Missing[b]	193		51		51,306
Excluded[c]	248		146		0
Lifetime Pap test, N (%)[d]					
Yes		290 (76)	125 (53)	195,295 (87)	
No		45 (22)	31 (45)	12,289 (10)	
Don't know		2 (1)	1 (1)	846 (0)	
Refused		2 (1)	2 (1)	786 (3)	
Missing[b]		150	69	89,175	
Excluded[e]		212	222	0	
Hysterectomy, N (%)[f]					
Yes		91 (16)	40 (22)	54,947 (21)	
No		226 (84)	95 (78)	132,527 (76)	
Don't know		1 (0)	0 (0)	232 (0)	
Refused		1 (0)	2 (1)	809 (4)	
Missing[b]		170	91	109,874	
Excluded[e]		212	222	0	
Currently pregnant, N (%)[g]					
Yes		5 (2)	5 (3)	2638 (4)	
No		131 (98)	92 (97)	66,639 (93)	
Don't know		0 (0)	1 (0)	202 (0)	
Refused		1 (0)	0 (0)	309 (3)	
Missing[b]		0	0	1	
Excluded[e]		72	89	0	

Unweighted counts and survey weighted percentages from 2014, 2015, and 2016 BRFSS states participating in the Sexual Orientation and Gender Identity module. We do not report responses of transgender men whose BRFSS designated sex is male, nor the responses of transgender women whose BRFSS designated sex is female.

[a]These data are restricted to respondents 40 years or older and whose BRFSS designated sex is male.

[b]Missing indicates respondents were not asked the survey item due to some state's reliance on questionnaire versions that *a priori* include a certain subset of questions, as well as partially due to survey fatigue.

[c]These respondents were excluded from being asked this question because their BRFSS designated sex was female.

[d]These data are restricted to respondents whose BRFSS designated sex is female.

[e]These respondents were excluded from being asked this question because their BRFSS designated sex was male.

[f]These data are restricted to non-pregnant respondents whose BRFSS designated sex is female.

[g]These data are restricted to respondents 44 years or younger and whose BRFSS designated sex is female.

Table 4 reports the result of the crude and covariate adjusted results from a complete case analysis. Complete case analyses do not account for systematic biases, but suggest that after adjusting for demographic and socioeconomic factors, transgender and gender nonconforming respondents are more likely to have never had a Pap test (transgender men PR 1.72, 95% CI = 0.98, 3.03; gender nonconforming PR 2.71, 95% CI = 07, 3.55) and transgender men are more likely to have had hysterectomy (PR 1.26, 95% CI = 0.99, 1.61) and less likely to be pregnant (PR 0.38, 95% CI = 0.13, 1.15), compared to cisgender women. Transgender women and gender nonconforming respondents have similar prevalence of PSA testing compared to cisgender men.

Table 4 also reports the results of the imputed regression analyses. The multiple imputation by chain equation estimates are similar but move towards the null, compared to complete case analyses. This suggests that, before modeling the impact of an unknown confounder, there is an attenuated association between gender and reproductive health outcomes based solely on differences in demographics and healthcare access. That is, misclassification of sex assigned at birth results in a sample with lower rates of PSA and Pap testing, lower prevalence of pregnancy, and higher prevalence of hysterectomy.

The single imputation results explore the possible upper and lower bounds of the association between gender identity and reproductive health outcomes given our assumptions

TABLE 4. Complete-case Analysis, Single Imputation and Multiple Imputation Regression Results

			Single Imputation Models		
	Crude PR	Adjusted PR[a]	Lower Bound PR[b]	Upper Bound PR[c]	MICE Model PR, PR (95% CI)
	PR (95% CI)	PR (95% CI)	PR (95% CI)	PR (95% CI)	
Never lifetime PSA test					
Transgender women	1.17 (0.89, 1.55)	1.09 (0.81, 1.48)	0.57 (0.38, 0.84)	1.55 (1.37, 1.75)	1.03 (0.82, 1.31)
Gender nonconforming	1.35 (0.87, 2.10)	0.95 (0.70, 1.31)	0.49 (0.28, 0.86)	1.72 (1.37, 2.17)	—
Never lifetime Pap test					
Transgender men	2.31 (1.31, 4.08)	1.72 (0.98, 3.03)	0.98 (0.53, 1.80)	4.32 (3.36, 5.54)	1.26 (0.77, 2.06)
Gender nonconforming	4.63 (3.17, 6.76)	2.71 (2.07, 3.55)	1.04 (0.61, 1.78)	6.38 (5.15, 7.91)	—
Had hysterectomy					
Transgender men	0.97 (0.76, 1.23)	1.26 (0.99, 1.61)	0.66 (0.49, 0.89)	1.50 (1.33, 1.70)	1.15 (0.95, 1.41)
Gender nonconforming	0.63 (0.43, 0.91)	0.86 (0.68, 1.08)	0.25 (0.17, 0.38)	1.50 (1.34, 1.68)	—
Currently pregnant					
Transgender men	0.47 (0.14, 1.53)	0.38 (0.13, 1.15)	0.27 (0.09, 0.78)	9.76 (5.98, 15.93)	0.59 (0.17, 1.99)
Gender nonconforming	0.65 (0.21, 1.94)	0.76 (0.23, 2.58)	0.33 (0.10, 1.07)	15.17 (10.55, 21.82)	—

Referent group for PSA testing is cisgender men; referent group for Pap testing, hysterectomy and pregnancy is cisgender women.
[a]Adjusted for race/ethnicity, age, unemployment, income, education, partnership status, poor health, and no insurance coverage.
[b]Single-imputation model assuming all TGNC respondents missing outcome data did not have the outcome.
[c]Single-imputation model assuming all TGNC respondents missing outcome data did have the outcome.

regarding the impact misclassification, and the subsequent sample exclusions may have on the observed PRs. The range of possible PR for lifetime PSA testing and hysterectomy was modest compared to the possible range for lifetime Pap testing and pregnancy. For example, the PR for never lifetime PSA testing is at most 1.55 (95% CI = 1.37,1.75) and 1.72 (95% CI = 1.37, 2.17) for transgender women and gender nonconforming individuals, respectively. The lower bounds are similarly informative, suggesting that given the observed data and our modeling assumptions, pregnancy is at minimum 1/10th as prevalent among transgender men and gender nonconforming individual (transgender men PR 0.27, 95% CI = 0.09, 0.78; gender nonconforming PR 0.33, 95% CI = 0.10, 1.07). Pap testing was similar among transgender men and gender nonconforming individuals and cisgender women only under the extreme assumption that all transgender men and gender nonconforming respondents with missing data have had a Pap test in their lifetime (transgender men PR 0.98, 95% CI = 0.53, 1.80; gender nonconforming PR 1.04, 95% CI = 0.61, 1.78). The possible upper bounds suggest that Pap testing was, at most 4.32 (95% CI = 3.36, 5.54) and 6.38 (95% CI = 5.15, 7.91) times less prevalent among transgender men and gender nonconforming individuals, respectively.

The results of both bias modeling approaches are reported in Table 5. We report the minimum and maximum adjustment factors guaranteeing that the bias-adjusted PR was either above or below one. The closer the minima and maxima are to one, the more robust the estimates are to bias. Conversely, the larger the range in which the resulting PR can be either above or below one, the more vulnerable an estimate was to selection bias and confounding. Both bias modeling approaches suggest that the PR estimates for hysterectomy among transgender men and gender nonconforming individuals, lifetime PSA testing among transgender women and Pap testing among transgender men were robust to small degrees of systematic bias. For example, if the selection bias odds ratio (OR) was 1.49 or higher, then the bias adjusted PR for never PSA testing is greater than one among transgender women. Among transgender men, selection bias ORs above 2.05 for Pap testing and above 1.09 for hysterectomy, guaranteed a bias adjusted PR greater than one. In contrast, the PR estimates for pregnancy and all outcomes among gender nonconforming individuals were highly sensitive to bias. We observed a large range of selection bias scenarios that would result in an association either above or below 1 for PSA testing (OR_{select} = 0.45–9.89), Pap testing (OR_{select} = 0.23–5.00), and pregnancy (OR_{select} = 0.18–9.38) among gender nonconforming individuals. We similarly observed that pregnancy among transgender men was highly sensitive to selection bias (OR_{select} = 0.53–9.99). Approach two mirrors these results and reinforces our findings from approach one for transgender men and women.

DISCUSSION

We applied quantitative bias methods to explore the magnitude and direction of systematic bias in the associations between gender identity and reproductive health outcomes in the BRFSS. Our results demonstrated that there was a large degree of uncertainty introduced by the BRFSS methodology for ascertaining sex assigned at birth and its reliance on skip patterns based on sex. These findings highlight the importance of attention to survey methodologies and implementation of validated sex and gender identity question for health surveys,

TABLE 5. Magnitude of Bias Adjustment Factors Required to Change the Direction of Association of the Systematic Bias and Random Error Adjusted PRs

	Approach One: Selection Bias		Approach Two: Missing Data and Confounding	
	Maximum Selection Bias OR Guaranteeing, PR < 1	Minimum Selection Bias OR Guaranteeing, PR > 1	Minimum RR Due to Confounding Guaranteeing, PR < 1	Maximum RR Due to Confounding Guaranteeing, PR > 1
Never lifetime PSA test				
Transgender women	0.54	1.49	1.55	0.73
Gender nonconforming	0.45	9.89	—	—
Never lifetime Pap test				
Transgender men	0.27	2.05	3.56	0.59
Gender nonconforming	0.23	5.00	—	—
Had hysterectomy				
Transgender men	0.53	1.09	1.57	0.92
Gender nonconforming	0.73	1.38	—	—
Currently pregnant				
Transgender men	0.53	9.99	5.90	0.12
Gender nonconforming	0.18	9.38	—	—

The bias-adjusted PRs are calculated as $PR_{adj} = \widehat{PR} \times OR_{select}$ or as $PR_{adj} = \frac{\widehat{PR}}{RR_{conf}}$. This table presents minimum and maximum adjustment factors that guaranteed a bias-adjusted PR either above or below one. The closer the minima and maxima are to one, the more robust the estimates are to bias.

specifically those that are nationally representative. Surveys that intend to collect unbiased data on transgender health should employ comprehensive and inclusive questions about gender, ask all participants about their sex assigned at birth rather than determining this based on the sound of their voice, and avoid cisnormative assumptions about sex and anatomy, including reliance on sex-based skip patterns.[13]

We found that estimates among gender-nonconforming respondents were the most vulnerable to small degrees of bias, whereas estimates of lifetime PSA testing among transgender women, and lifetime Pap testing and hysterectomy among transgender men were robust to small degrees of bias. All associations were sensitive to large degrees of bias. Further, single-imputation models enabled us to estimate potential upper and lower bounds for each PR given our assumptions about the mechanisms for selection bias and missing data. We found that the range of possible associations between gender, lifetime PSA testing, and hysterectomy was modest compared to the large possible range for lifetime Pap testing and pregnancy. Nonetheless, these potential upper and lower bounds placed on the estimated associations were informative given the current sparsity of literature on reproductive health outcomes in transgender and gender nonconforming populations.

Descriptive results suggested that the 29.6% of transgender women and 30.2% of transgender men excluded from answering sex-specific questions are demographically distinct from those who were included in the sample. This supports our hypothesis that there is substantial selection bias present in the BRFSS measures of reproductive health among transgender and gender nonconforming respondents. The survey-weighted responses, taken in context of our bias analysis, suggest it is likely that Pap testing is less prevalent among transgender men and gender nonconforming individuals compared to cisgender women; that PSA testing among transgender women and gender nonconforming individuals may not be dissimilar from cisgender men; and the prevalence of hysterectomy may not be dissimilar among transgender men, gender nonconforming individuals, and cisgender women. It also suggests that 1.9% of transgender men and 2.6% of gender nonconforming individuals were pregnant at the time of interview, compared to 4.0% of cisgender women. Imputation analyses suggest that the pregnancy may at minimum be 1/10th as prevalent among transgender men and gender nonconforming respondents when compared to cisgender women.

There is limited literature in which to contextualize these estimates. A systematic review of clinical and community samples suggests that the proportion of transgender men current on their Pap testing is between 5.0% and 9.2% lower than cisgender women.[30] Estimates of the prevalence of hysterectomy among transgender men range from 5.5% to 14%.[21,31–33] Despite documented cases of prostate cancer among transgender women, there are no studies that examine rates of PSA testing.[34,35] Few studies characterize the health of gender nonconforming populations.[5,11] Lastly, several studies document transgender men's experience with pregnancy, but their sampling methods preclude an estimate of the prevalence of pregnancy.[36–38] Therefore, unbiased estimates of the prevalence of pregnancy among transgender men and gender

nonconforming individuals have important implications about the availability of trans-competent and inclusive family planning and antenatal care for transgender individuals experiencing pregnancy.

A strength of this study is the comparison of two approaches of modeling the systematic bias that results from the misclassification of sex assigned at birth and the subsequent missing data. The fact that the results of these two approaches reinforce each other lends credence to our findings. Imputation methods proved to be particularly useful in estimating the potential bounds of bias and brings in known information about respondents who were missing data. This is also the first study, to our knowledge, to undertake rigorous bias analysis to examine health disparities in the transgender population.

A limitation of our approach is differential handling of gender nonconforming respondents, for whom we are unable to infer their sex assigned at birth. For approach one, we observe significant variation in the estimated PR that is most likely due to the additional uncertainty incorporated into the record-level bias adjustment. Further, we were unable to apply the secondary approach based on multiple implication by chain equations among gender-nonconforming respondents, as imputation would inappropriately rely on information from individuals without relevant anatomy and yield unreliable estimates.

Another limitation is our reliance of sex assigned at birth as a proxy for reproductive anatomy. These assumptions do not account for intersex individuals or individuals who pursue gender affirming procedures. Therefore, our denominator of who is at risk may be inflated. We also were unable to model bias in questions about breast cancer screening. According to the American Cancer Society, all individuals with breasts over the age of 40 should have the option to begin annual screening.[39] However, the presence or absence of breast tissue depends on a wide range of individual gender-affirming medical choices a transgender person decides to pursue.[21]

As with all health surveys, the BRFSS relies on self-report. Therefore, transgender identity is more accurately a measure of willingness to report transgender and gender nonconforming identities. We hypothesize that willingness to report is associated with regional and personal factors, including the sociopolitical context where a person lives, geography, race, and age. Additionally, the BRFSS only samples non-incarcerated, non-institutionalized individuals, explicitly excluding group homes and shelters. Because of high rates of incarceration among transgender women of color, and high rates of homelessness among transgender and gender nonconforming individuals, the BRFSS does not capture the health of these vulnerable populations.[21]

Due to the lack of external data to inform the degree of selection bias or confounding, we considered a wide range of bias scenarios. Our models assume a uniform distribution, i.e., although unlikely, we assume all bias parameters are of equal probability. Although a modest bias adjustment parameter of two is more likely than an extreme value 10, the aim of this analysis was to determine the sensitivity of the associations to bias through the provision of a wide range of possible values. Nevertheless, this is a limitation of the analysis. Future validation studies that allow us to estimate the true selection probabilities among transgender and gender nonconforming respondents would allow more accurate estimates of the selection bias odds ratio.[22] Similarly, additional data could inform bounds on the risk ratio due to confounding, per Flanders and Khoury.[29]

This study provides evidence that BRFSS data provide non-representative estimates of transgender and gender nonconforming individuals who are asked questions about their reproductive health, and that these outcomes are vulnerable to bias. Therefore, when using BRFSS it should be emphasized that these data do not produce unbiased population-based estimates. Ideally, analyses should be accompanied by quantitative bias analysis that acknowledges the impact of the systematic misclassification of sex assigned at birth, and the subsequent selection bias and missing data issues that arise.

REFERENCES

1. Meyer IH, Brown TN, Herman JL, Reisner SL, Bockting WO. Demographic characteristics and health status of transgender adults in select US regions: behavioral risk factor surveillance system, 2014. *Am J Public Heal*. 2017;107:582–589.
2. Edmiston EK, Donald CA, Sattler AR, Peebles JK, Ehrenfeld JM, Eckstrand KL. Opportunities and gaps in primary care preventative health services for transgender patients: a systemic review. *Transgend Health*. 2016;1:216–230.
3. Reisner SL, Deutsch MB, Bhasin S, et al. Advancing methods for US transgender health research. *Curr Opin Endocrinol Diabetes Obes*. 2016;23:198–207.
4. Meerwijk EL, Sevelius JM. Transgender population size in the United States: a meta-regression of population-based probability samples. *Am J Public Health*. 2017;107:e1–e8.
5. Reisner SL, Poteat T, Keatley J, et al. Global health burden and needs of transgender populations: a review. *Lancet*. 2016;388:412–436.
6. Dai H, Hao J. Sleep deprivation and chronic health conditions among sexual minority adults. *Behav Sleep Med*. 2019;17:254–268.
7. Crissman HP, Berger MB, Graham LF, Dalton VK. Transgender demographics: a household probability sample of US adults, 2014. *Am J Public Health*. 2017;107:213–215.
8. Cloyes K, Yang R, Latimer S. Refusal to respond to BRFSS sexual orientation and transgender survey items among adults age 60+. *Gerontologist*. 2017;56(suppl 3):438.
9. Motwani A, Fatehchehr S. Quality of life and health care access in transgender population: findings from 21 U.S. states in the Behavioral Risk Factor Surveillance System (BRFSS) Survey. *J Minim Invasive Gynecol*. 2017;24:S59–S59.
10. Streed CG Jr, McCarthy EP, Haas JS. Association between gender minority status and self-reported physical and mental health in the United States. *JAMA Intern Med*. 2017;177:1210–1212.
11. Gonzales G, Henning-Smith C. Barriers to care among transgender and gender nonconforming adults. *Milbank Q*. 2017;95:726–748.
12. Tabaac AR, Sutter ME, Wall CSJ, Baker KE. Gender identity disparities in cancer screening behaviors. *Am J Prev Med*. 2018;54:385–393.
13. Bauer GR, Braimoh J, Scheim AI, Dharma C. Transgender-inclusive measures of sex/gender for population surveys: Mixed-methods evaluation and recommendations. Dalby AR, ed. *PLoS One*. 2017;12:e0178043.
14. Riley NC, Blosnich JR, Bear TM, Reisner SL. Vocal timbre and the classification of respondent sex in US phone-based surveys. *Am J Public Health*. 2017;107:1290–1294.
15. Patterson JG, Jabson JM, Bowen DJ. Measuring sexual and gender minority populations in health surveillance. *LGBT Health*. 2017;4:82–105.

16. Reisner SL, Radix A, Deutsch MB. Integrated and gender-affirming transgender clinical care and research. *J Acquir Immune Defic Syndr*. 2016;72(suppl 3):S235–S242.
17. Cahill SR, Baker K, Deutsch MB, Keatley J, Makadon HJ. Inclusion of sexual orientation and gender identity in stage 3 meaningful use guidelines: a huge step forward for LGBT health. *LGBT Heal*. 2016;3:100–102.
18. The GenIUSS Group. *Best Practices for Asking Questions to Idenfity Transgender and Other Gender Minority Respondents on Population-Based Surveys*. Los Angeles, CA: The Williams Institute, UCLA School of Law; 2014. Available at: https://williamsinstitute.law.ucla.edu/wp-content/uploads/geniuss-report-sep-2014.pdf. Accessed July 24, 2017.
19. Institute of Medicine (US) Committee on Lesbian, Gay, Bisexual, and Transgender Health Issues and Research Gaps and Opportunities. *The Health of Lesbian, Gay, Bisexual, and Transgender People: Building a Foundation for Better Understanding*. Washington, DC: National Academies Press; 2011. Available at: https://www.ncbi.nlm.nih.gov/books/NBK64795/. Accessed July 12, 2018.
20. Deutsch MB. Guidelines for the Primary and Gender-Affirming Care of Transgender and Gender Nonbinary People. 2016. Available at: http://transhealth.ucsf.edu/protocols. Accessed November 26, 2018.
21. James SE, Herman JL, Rankin S, Keisling M, Mottet L, Anafi M. *The Report of the 2015 U.S. Transgender Survey*. Washington, D.C.: National Center for Transgender Equality; 2016. Available at: http://www.transequality.org/sites/default/files/docs/usts/USTS Full Report - FINAL 1.6.17.pdf.
22. Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. *Int J Epidemiol*. 2014;43:1969–1985.
23. Lash TL, Fox MP, Fink AK. *Applying Quantitative Bias Analysis to Epidemiologic Data*. New York, NY: Springer; 2009.
24. Mokdad AH. The Behavioral risk factors surveillance system: past, present, and future. *Annu Rev Public Health*. 2009;30:43–54.
25. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med*. 2011;30:377–399.
26. Graham JW. Missing data analysis: making it work in the real world. *Annu Rev Psychol*. 2009;60:549–576.
27. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. Hoboken, NJ: John Wiley & Sons, Inc.; 1987.
28. Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prev Sci*. 2007;8:206–213.
29. Flanders WD, Khoury MJ. Indirect assessment of confounding: graphic description and limits on effect of adjusting for covariates. *Epidemiology*. 1990;1:239–246. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2081259. Accessed May 17, 2018.
30. Gatos KC. A literature review of cervical cancer screening in transgender men. *Nurs Womens Health*. 2018;22:52–62.
31. Quinn VP, Nash R, Hunkeler E, et al. Cohort profile: Study of Transition, Outcomes and Gender (STRONG) to assess health status of transgender people. *BMJ Open*. 2017;7:e018121.
32. Beckwith N, Reisner SL, Zaslow S, Mayer KH, Keuroghlian AS. Factors associated with gender-affirming surgery and age of hormone therapy initiation among transgender adults. *Transgend Health*. 2017;2:156–164.
33. Kailas M, Lu HMS, Rothman EF, Safer JD. Prevalence and types of gender-affirming surgery among a sample of transgender endocrinology patients prior to state expansion of insurance coverage. *Endocr Pract*. 2017;23:780–786.
34. Joint R, Chen ZE, Cameron S. Breast and reproductive cancers in the transgender population: a systematic review. *BJOG.* 2018;125:1505–1512.
35. Deebel NA, Morin JP, Autorino R, Vince R, Grob B, Hampton LJ. Prostate cancer in transgender women: incidence, etiopathogenesis, and management challenges. *Urology*. 2017;110:166–171.
36. Obedin-Maliver J, Makadon HJ. Transgender men and pregnancy. *Obstet Med.* 2016;9:4–8.
37. Light AD, Obedin-Maliver J, Sevelius JM, Kerns JL. Transgender men who experienced pregnancy after female-to-male gender transitioning. *Obstet Gynecol*. 2014;124:1120–1127.
38. Ellis SA, Wojnar DM, Pettinato M. Conception, pregnancy, and birth experiences of male and gender variant gestational parents: it's how we could have a family. *J Midwifery Womens Health*. 2015;60:62–69.
39. American Cancer Society: Cancer screening guidelines. Available at: https://www.cancer.org/healthy/find-cancer-early/cancer-screening-guidelines/american-cancer-society-guidelines-for-the-early-detection-of-cancer.html. Accessed May 17, 2018.