Smoking out Health Disparities in Sexual and Gender Minorities: Lessons from the National Adult Tobacco Survey

Lara M. Stepleman¹, Eliot J. Lopez¹, Wendi Rawlins², Vahé Heboyan³

¹Department of Psychiatry and Health Behavior, Medical College of Georgia, Augusta University, Augusta, Georgia, USA
²Department of Psychology, Augusta University, Augusta, Georgia, USA
³Department of Clinical and Digital Health Sciences, College of Allied Health Sciences, Augusta University, Augusta, Georgia, USA

Email: *lsteplem@augusta.edu


Received: June 10, 2017
Accepted: July 31, 2017
Published: August 3, 2017

Abstract

Sexual and gender minorities (SGM) in the United States experience a number of health disparities and unique contributing factors to them. From a research perspective, survey design and implementation challenges, such as lack of inclusion of effective items for assessing SGM status and inadequate sampling methods, remain barriers to studying SGM. The purpose of this commentary is to describe, using the National Adult Tobacco Survey (NATS) survey items and datasets, the primary limitations we encountered when trying to describe SGM tobacco use. Our intent is to demonstrate through use of a national dataset around a specific health disparity, the imperative for researchers to change their data collection strategies and practices around tobacco use and other healthcare priorities. Our team utilized the 2009 NATS dataset as well as the 2012 iteration to highlight significant changes between them regarding demographics, tobacco use, and access to healthcare, in addition to methodological concerns regarding sampling strategies. It is critical that researchers strive to use items for survey research that accurately capture data on marginalized groups. Additionally, careful consideration is warranted regarding strategies to identify members of these populations, changes in item wording, and changes in questions asked over time in an effort to track changes in behavior over time.

Keywords

Health Disparities, LGBT, Tobacco Use, Survey Research, National Adult Tobacco Survey

DOI: 10.4236/jss.2017.58001
1. Introduction

As the visibility of Lesbian, Gay, Bisexual, and Transgender (LGBT) persons, broadly referred to as sexual and gender minorities (SGM), has increased within the United States, so too does our understanding of the magnitude and impact of health disparities experienced across and within these communities. Studies over the past two decades, in particular, have described such health inequities, including higher risk for disability [1], cardiovascular disease [2], and substance abuse [3], and their associations with SGM stigma and discrimination [4] [5] [6]. Further, the national edict to improve SGM health issued by the White House in 2010 with response by the Department of Health and Human Services has arguably become one the most impactful initiatives, not only by defining measurable health objectives, but by situating well-being of SGM and their families as connected to the health and well-being of all Americans [7]. Facilitating the development of impactful individual, community, and societal SGM health interventions also necessitates that national SGM samples can be studied and in comparison to heterosexual and cisgender (birth sex and gender identity concordant) counterparts, when appropriate. This requires a consistent means to purposefully assess sexual orientation and gender identity that ideally becomes standard assessment language used by agencies undertaking this type of work. By-and-large, this has not occurred, especially across national health agencies. To demonstrate assessment and sampling strategy limitations that can impede the investigation of SGM health using national, publically available datasets, we review our research team’s recent experiences in working with one on national tobacco use. Although SGM research barriers are well-described by others [8] [9], tangibly outlining such challenges and possible resolutions around a specific health disparity may provide further clarity and demonstrate the imperative for researchers to change their data collection strategies and practices.

1.1. The Case of Tobacco

Tobacco is not only the leading cause of preventable and premature death in the United States [10] but also results in one of the most robust and detrimental health disparities affecting SGM communities [11]. Data from national surveys, such as the National Adult Tobacco Survey (NATS), as well as from city- and state-wide surveys indicate that exposure to secondhand smoke and cigarette smoking and other tobacco products are considerably higher among SGM in general and especially its sub-populations, such as lesbian and bisexual women [10] [12] and HIV-positive SGM individuals [13].

Even though smoking rates have declined to 17% in 2014 for US [14] adults compared with 43% in 1964, smoking still remains considerably high at 23.9% among SGM [14]. One study found smoking rates as high as 62% among transgender women [15]. Additionally, SGM related stress and trauma, such as anti-gay verbal and physical attacks, same-sex intimate partner violence, and childhood sexual abuse, have each been found to independently increase the
odds of smoking [16] [17]. Finally, the growing e-cigarette trend is also a concern, with limited evidence suggesting that e-cigarette use in the SGM population is about 4.5% compared to 1.9% in the heterosexual/straight community, mirroring disparities in cigarette smoking [10] [18].

1.2. The National Adult Tobacco Survey (NATS)

Although there is a solid research foundation on SGM tobacco related health disparities, there is still much to flesh out nationally, especially in subgroups for which collecting representative samples of data can be arduous, such as transgender individuals. Aforementioned, our research team recently confronted these issues and others when investigating SGM-related health disparities and tobacco economics using the NATS from 2009-2010 [19] and 2012-2013 [10]. The NATS was created by the Centers for Disease Control and Prevention (CDC) to assess the prevalence of tobacco use and factors promoting and impeding use among adults. The NATS also establishes a comprehensive framework for evaluating national and state-specific tobacco control programs.

The NATS gauges the extent of tobacco use in adults, evaluates the amount that tobacco use varies as a function of demographics estimates, and the achievement of key short-, intermediate-, and long-term tobacco prevention and control outcome indicators. The NATS administrations were conducted using stratified sampling by state via landline and cell phone numbers. The 2009-2010 sample contained 118,581 adults and the 2012-2013 sample contained 60,192 adults. Demographics including age, gender, marital status, income, education level, state, sexual orientation and race/ethnicity are collected and used for statistical purposes.

The 2009-2010 survey identified smokers, either previously or currently, by asking if respondents have smoked 100 cigarettes in their lifetime to assess smoking status, as well as asking about multiple tobacco product use, cessation and chronic condition information, use of counseling services, and lifetime quit attempts. The 2012-2013 survey assessed tobacco use in many of the categories listed for 2009-2010 as well as for e-cigarettes. Changes from one dataset to the next, as they relate to our experiences with data analyses, are discussed in further detail hereafter.

In describing our team’s SGM-related data challenges with the NATS, it is crucial to underscore that our struggles are by no means unique to this survey; similar and additional challenges are evident in many national, publically-available health datasets, such as the National Youth Tobacco Survey (NYTS), and the Global Adult Tobacco Survey. In fact, the NATS included items about SGM status in both survey administrations and before many other non-SGM specific national health surveys. Still, as these challenges emerged specifically from our work with the NATS data, we can most clearly illustrate why and how changes in survey content and sampling methodology can be advantageous for clarifying and reducing SGM health disparities.
2. Challenges to Understanding SGM Health Behaviors and Disparities in NATS

We review four primary ways in which our group confronted difficulties in trying to capture LGBT health disparities through the NATS: 1) categorization of sexual and gender identity, 2) significant changes in survey items between administrations; 3) sampling methodology, and 4) participant response. Suggested alternative approaches, applicable to other national and international datasets, are offered herein.

2.1. Categorization of Sexual and Gender Identity

In survey research, determining which identity characteristics to collect such that they function as meaningful constructs to a study’s purpose is not a new challenge [20]. Similar issues have arisen around racial and ethnic identity categories [21], and not unlike race and ethnicity, the language of sexual and gender identity has shifted over time, paralleling society’s relationships to these constructs. SGM terminology also has a complicated history tied to perceptions of pathological, criminal, and immoral behavior—a topic beyond the scope of this paper. Even a commonplace term like “homosexual” has been described as too clinical and linked to its history as a mental disorder [22]. Further, there are inherent limits to using SGM terminology across culture given varied understandings of gender and sexuality around the world [23].

SGM invisibility occurs when no SGM items are asked in surveys, such as in the Youth Tobacco Survey and the NYTS. Conversely, the NATS collects SGM data in the 2009-2010 and 2012-2013 administrations, though not in the same way across years. In 2009-2010, participants have the following options in response to “Do you consider yourself to be” with the options of 1) heterosexual or straight; 2) gay or lesbian; 3) bisexual; 4) transgender; 5) respondent does not understand responses; 6) don’t know/not sure; and 7) refused?” Particularly problematic for transgender individuals, this survey strategy forces a choice between sexual orientation and gender identity, which ignores the reality of individuals having both.

In 2012-2013, “transgender” is no longer presented in the first set of response options. Instead, it is presented subsequently if “something else” is selected as the response choice for sexual orientation. If a transgender participant identified first as gay, straight, or bisexual, there is no further opportunity to provide gender identity. Should the participant state “something else,” transgender is grouped with other sexual orientation items, such as “you are not straight, but identify with another label such as queer, trisexual, omnigender, or pansexual.” Again, this approach conflates sexual orientation with gender identity, making it all but impossible to acquire data on tobacco use in transgender individuals.

One resolve has been to assess transgender identity within the gender/sex item. The NATS does not do this and instead largely reinforces the binary no-
tion of gender. In either case, the long-standing awareness of gender identity and biological sex as discrete [24] is not reflected in these strategies nor that transgender is one of many gender identities to which people now ascribe. In the 2009-2010 NATS instruments, it is unclear whether an individual was providing their birth sex or gender identity, making it impossible to determine if someone was a trans man or woman to examine distinct tobacco use and health disparities, such as trans men’s risk for gynecologic cancers [25], which is believed to be increased through tobacco use.

The term LGBT itself also may unintentionally reinforce that these data be collected within the same survey item, a practice we have used and was once thought acceptable and progressive relative to asking nothing about sexual orientation and gender identity. The shorthand of LGBT (and similar acronyms) serves these communities collectively to bring needed societal awareness and visibility; however, the true utility of these labels unravels when trying to understand the unique and sometimes very different health disparities experienced within SGM subpopulations and becomes all but obsolete at the individual level.

Support for better data collection has already begun, such as the American Lung Association (ALA) calling on the Health and Human Services Secretary to incorporate the proposed “Data Standards for Sex” by looking at sex from a social perspective rather than genetic and/or biological. They also stated that the new standard for public health surveys should include items on sexual and gender identity as part of core demographics [26].

The NATS would benefit from independent questions about birth sex, gender identity, and sexual orientation. We agree with the two-step assigned sex and gender identity protocol developed in 1997 by the Transgender Health Advocacy Coalition [27] and since 2011 has been used by the CDC in their electronic surveillance system. Step one inquires about current gender identity with the options of 1) male; 2) female; 3) trans male/trans man; 4) trans female/trans women; 5) gender queer/gender nonconforming; 6) different identities (please state). Step two inquires about sex assigned at birth with the choices of 1) male; 2) female.

In terms of sexual orientation, the Williams Institute [28] has recommended the item developed by the National Center for Health Statistics, which uses the stem “Do you consider yourself to be” with the options of “a) heterosexual or straight; b) gay or lesbian; c) bisexual?” While the Williams Institute recommends not using an “other” category and does not specify providing additional options such as pansexual, these options should be considered [28]. The Williams Institute report was published in 2009 and SGM culture has been transforming rapidly. In our interaction with patients in clinical settings, we find more young adults, in particular, identifying as pansexual. This is especially true for transgender individuals who may find that this better describes their sexual orientation.

Though the evolution of SGM assessment will continue, we believe the above to be a low-burden solution to collection of SGM identity data for un-
derstanding tobacco health disparities. Logically, survey context matters, so additional items about sexual behavior and attraction among others may be appropriate dependent on the purpose of the survey or the target population. Similarly, studies specific to SGM individuals may require in-depth data collection about these communities that is impractical in national general population studies.

2.2. Significant Changes in Survey Items between Administrations

As access to healthcare increases in the United States, it is imperative to track how individuals utilize these resources over time. This is particularly true for SGM individuals, who are historically underserved. However, with the Affordable Care Act (ACA) passed into law in 2010, as well as increased awareness of SGM health disparities, makes it crucial to assess SGM healthcare in order to provide intervention specific to their needs [5].

Focusing specifically on tobacco use, the 2009-2010 NATS dataset asked several questions about access to healthcare and how care providers may assist in attempts to quit smoking, via counseling, prescriptions for nicotine replacement, or appropriate referrals to cessation programs. However, these items, or reasonable approximations of these items, are notably absent from the 2012-2013 survey, making trends unfeasible to track, particularly in light of federal changes to healthcare access. Further, as one of our goals was to assess potential increases in healthcare utilization for SGM following the ACA and disparities or changes in the role of care providers in cessation for SGM, the removal of these items eliminated opportunities for such inquiry.

Items are also significantly altered from one dataset to the next. For instance, in 2009-2010, participants were asked “How old were you when you smoked a whole cigarette for the first time?” while participants in 2012-2013 were asked about “part or all of a cigarette.” Though the latter captures more data, it also confounds identifying the initiation of tobacco use between surveys for comparison. This seemingly benign change in survey wording has large ramifications; research suggests that smoking during adolescence increases sensitivity to the rewarding aspects of nicotine, increasing the likelihood of nicotine addiction [29]. However, altered wording changes the definition of the initial use of cigarettes, making interpretation between these data difficult.

Not all item changes are negative; in fact, many positive changes are observed in the 2012-2013 NATS survey. A necessary addition, for example, was the inclusion of questions about e-cigarettes in the 2012-2013 dataset, which represents a growing contingent of the population who utilize these as an alternative, or compliment, to cigarettes. Further, though we note the inconsistency in the sexual orientation items, we do commend the NATS research team for inclusion of items that attempt to capture data on these marginalized groups. Still, we caution that decisions to change items be made mindfully and with consideration for the impact that they may have on analyses which track trends over time.
2.3. Sampling Methodology

Probability sampling methods tend to generate very small SGM samples. Data reported in one of the largest single-study surveys of SGM in the US [30] indicates that the national average of persons who identify as an SGM was 3.5%; we should note that participants were asked “Do you identify as lesbian, gay, bisexual, or transgender?” perpetuating the conflation of these groups. However, data from the NATS samples suggest even smaller populations of SGM; the 2009-2010 sample reports a total of 2%, while the 2012-2013 data report a total of 2.9%. Further, only 26% \((n = 642)\) of the SGM sample in the 2009-2010 data were smoking, while 22\% \((n = 386)\) in 2012 were current smokers. Within the subsamples, we find a very small number of transgender participants in both datasets \((2009, n = 96; 2012, n = 17)\). Typically, the four sub-groups constituting LGBT are often collapsed into one to compensate despite the recognized distinct effects of tobacco use on the health within these groups and different concerns regarding psychosocial [31], sexual [32], and medical health [1] [33]. Further, smaller samples limit researchers’ ability to conduct robust and accurate analyses and develop comprehensive models of behavior. Though a thorough examination of sampling procedures is outside the scope of this particular discussion, we offer some suggestions on how researchers may elect to recruit members of marginalized groups.

Historically, sampling of minority groups has been a challenge, particularly when the goal of the research is to provide a fair representation of the population. Therefore, oversampling strategies that aim to compensate for small sample sizes are often utilized. The simplest oversampling approach [34] [35] is to just increase the sample size. However, due to the costs of such an approach, an alternative is to combine data, either over longer periods of time or across different data sets [36]. Considerations that arise due to these approaches include 1) consistently changing data-interests, 2) the daunting task of merging multiple datasets, and 3) a lack of standardized measures across studies [36]. Even though the latter may be addressed by developing a standardized approach to measures at least in federally funded surveys, the former two still remain critical barriers to operationalizing this method.

Additional practical approaches for oversampling SGM can be network, or snowball sampling, and location sampling. The former asks the sampled persons to identify others who are of a certain demographic, while the latter samples persons in specific community locations where these individuals usually congregate. A more detailed discussion on various techniques and their advantages and limitations is provided in Kalton [34] and Meyer and Wilson [37].

Finally, investigators may elect to oversample at block level. Blocks are small geographic areas that are known to be “rich” in the demographics of interest. For example, previous national surveys or polls, such as the recent Gallup poll [30], can be used to identify areas of relatively dense SGM populations, which can be deliberately oversampled. However, this technique may unintentionally
result in oversampling other groups as well. Nevertheless, if combined with screening (respondents are briefly screened for meeting certain “eligibility” criteria for oversampling), this technique may be highly effective.

Given that research suggests that respondents are becoming more open to providing SGM status information in surveys [36], we advocate for the inclusion of standardized sexual orientation and gender identity questions into major survey initiatives to better identify and recruit members of these groups. Furthermore, echoing the calls by leading public health advocates, such as the ALA and Institute of Medicine of the National Academies [36], we suggest the incorporation of mixed-method oversampling techniques into these data collection efforts. Though we recognize that there is an economic cost to implementing oversampling strategies as well as potential validity concerns resulting from non-probability sampling techniques, the potential for gaining a greater understanding of these groups might make such efforts worthwhile. Moreover, data availability and sample size are especially critical for promoting SGM-focused research and extramural funding proposals that address SGM health disparities, since the lack of data on SGM and their sub-groups introduces competitive disadvantages and constraints to advance SGM-focused research.

2.4. Participant Response

In an effort to capture data on sexual and gender identity, investigators may take for granted participants’ understanding of terms like “LGBT,” “heterosexual,” “transgender,” and other phrases used to identify SGM. For example, in the 2009-2010 NATS databases, an item assessing sexual orientation is presented to participants. One potential response, “Respondent does not understand responses,” was selected by 0.52% of the unweighted sample (n = 610). By its inclusion, the questionnaire developers acknowledge there may be a contingent of participants for whom this terminology is unclear; however, it does not appear that further effort to define these terms in order to obtain more specific data is provided. This particular issue is carried over into the 2012-2013 NATS database, in which a larger proportion of participants (2.3%; n = 1361) reported not understanding the potential responses.

Item wording further complicates participant response. For instance, in the 2012-2013 NATS database, participants may select “something else” as a response to the question on sexual orientation; 0.42% (n = 254) of participants selected this. They are then presented with a follow-up item allowing them to clarify what they meant by “something else;” it is here that participants are presented with “transgender” as a response choice. Given that participants do not have the opportunity to see forthcoming items and options, they may not realize that a response choice better reflecting their self-identity is nested within “something else.” Consequently, they may default to a set of options that does not “fit.”

An additional consideration is allowing the participant to choose their own
description. Open-ended responses, while providing an opportunity to allow participants to select a response that may not be included elsewhere, can often lead to redundant or unusable data. For example, though provided with ample opportunity to decline to respond (“refused” is a viable selection, 2.69%; \( n = 1619 \)), several individuals responded with variations of refusal (“does not want to explain,” “do not want to answer,” etc.). Also mixed in with these responses are items that were offered previously but were not selected, such as “heterosexual,” gender identifiers such as “man” or “female,” and responses that were irrelevant, such as “alien” and “flying unicorn.”

The Williams Institute [28] suggests using terms such as “gay and lesbian” and “bisexual,” without the use of definitions except when respondents do not understand the question. Additionally, they advocated against the use of “other” categories, as these responses are typically discarded from most analyses. Further, recoding these responses is typically time-intensive. To address this, we advocate for greater inclusivity of additional response options for SGM items (e.g., pansexual) in an effort to provide ample opportunity initially to accurately self-identify. Finally, regarding “not sure” responses, the context about participants’ uncertainty is typically unclear, particularly within sexual and gender identity research. Participants may be unsure due to their own indecision regarding identity [38] or they may not understand the question. Researchers may wish to allow participants the opportunity to clarify what aspect of the question or response options they are unsure about, as this may provide better insight into their experience of such items and improve the integrity of the collected data. Considering many of these interviews are conducted via telephone, follow-up items regarding participants’ intent might be included in future surveys incorporating “unsure” as a response option. More specificity regarding “not sure” also may be warranted; incorporating “not sure about my identity” or “not sure what these items mean” options would provide clarification on the individual’s perspective.

3. Conclusions

During the development of this piece, our team recognized our advocacy of clearer, more distinct categories for SGM populations, and anticipated a potential critique regarding the limitation of participant identification, specifically in light of the fluidity of both language and identity. For example, we highlight labels that participants use to self-identify that may not be the most helpful in the discussion of sexual and gender identity, such as “alien” and “flying unicorn,” and only serve to remove people from specific subsamples or create new subsamples with incredibly small numbers. We recognize that labels such as gay, lesbian, and transgender are socially constructed and ultimately mean little outside of their use within our flawed classification systems of individuals. Still, these groups, as we culturally understand them at this historical time point, face specific challenges, have unique experiences specific to their identities, and deal
with specific health disparities. As such, while we recognize that these labels often pose limits, they also provide opportunities to learn more about these specific groups and reduce health disparities.

As researchers continue to investigate the health disparities and health behaviors of SGM, our approach to conceptualizing constructs, asking meaningful questions, identifying target individuals, and collecting and analyzing data must shift accordingly. As has been mentioned, adopting language that can be used across survey samples would help to ensure that we are measuring the same SGM constructs [39] (as this also likely contributes to some of the ongoing debates about SGM representation in the United States). On a larger scale, building an evolving set of best practices that are implemented nationally for conducting research that includes SGM populations in a culturally competent manner, including research question and instrument development, implementation and collection of data as well as analyses, interpretation, and reporting will yield information that more accurately capture health disparities, like tobacco use, to ensure that meaningful tailored interventions can follow.

References


