



A Global Call to Action for Gender-Inclusive Data Collection and Use

Rajeev Colaço and Stephanie Watson-Grant



Key Findings

- Global surveys and data collection efforts focus mainly on binary cisgender sex-disaggregation of data.
- This binary focus excludes transgender and gender-nonconforming populations, thereby further marginalizing them and silencing their voices.
- We call for global action to ensure that we collect, analyze, and use binary cisgender as well as multinomial transgender and gender-nonconforming data so our data are inclusive and representative of all.

Over the past two decades, surveyors, researchers, and program implementors have made enormous strides in the collection and analysis of sex-disaggregated data in international development programs.^{1,2} The focus on sex-disaggregated data is paramount; a key challenge with aggregated data is that they can mask important differences between cisgender female and cisgender male populations, thereby contributing to false assumptions that cisgender women and cisgender men share the same experiences and benefit from international development and other programs in similar ways.³ The sex-disaggregation of data, which can often be generated with the clicks of a few buttons in today's global digital health data systems such as District Health Information

Software (DHIS2) and Demographic Health Surveys, gives us richness and nuance that allow us to compare and contrast cisgender data.^{1,3}

It is important to acknowledge, however, that when data collection at individual and community levels is limited to begin with, our survey results and data systems (e.g., DHIS2, Demographic Health Surveys) reflect those limits. For example, we know that respondents are more comfortable answering questions from a data collector that they can identify with, such as a data collector of a similar sex.³ Comfort and trust can lead to higher-quality data and consequently better interpretation of data for policy making or programmatic adaptation.⁴

Many surveys and electronic data systems, including ones such as DHIS2 that are used across countries, allow us the flexibility to collect and enter not only sex-disaggregated cisgender data but also gender-disaggregated data from transgender and gender nonconforming people⁵ (see Table 1 for a comparison between sex and gender disaggregation). However, the reality is that we still do not accurately and comprehensively collect gender-disaggregated data in most of our standardized surveys and data collection systems. This holds true in resource-sufficient as well as resource-constrained settings. In the United States, for example, one study has identified that out of 100 federal surveys, only 11 collect some sexual orientation and gender identity data.⁶ Although there is still a focus, understandably and laudably, on women’s and girls’ equity issues,⁷ the global policy landscape has not evolved to include gender data considerations beyond the binary.

Table 1. Key differences between sex- and gender-disaggregated data

Sex-disaggregated data	Gender-disaggregated data
<p>Binary by design:</p> <ul style="list-style-type: none"> • Cisgender^a female • Cisgender male 	<p>Multinomial by design:</p> <ul style="list-style-type: none"> • Cisgender female • Cisgender male • Transgender^b female/trans female/male-to-female • Transgender male/trans male/female-to-male • Nonbinary/genderqueer/gender nonconforming^c • Other (please specify)

Note: We acknowledge that although this is the currently used language and terminology that surveys and data collection methods use to capture gender identities, the terminology is constantly evolving—as it should—to better capture the nuances and diversity of all gender identities.

- ^a Refers to people whose gender identity is congruent with the gender they were assigned at birth.
- ^b Refers to an umbrella term used to describe people whose gender identity is inconsistent with the gender they were assigned at birth.
- ^c Refers to people with a gender identity that falls outside of the gender binary of male and female. They may identify as both male and female or as another gender, or their gender identity may change over time.

Sources: University of San Francisco LGBT Resource Center (<https://lgbt.ucsf.edu/glossary-terms>), RTI Guide to Research with LGBTQ Populations (internal document, 2021), University of California, Los Angeles (UCLA) Williams Institute SOGI Adult Measures Recommendations (<http://williamsinstitute.law.ucla.edu/wp-content/uploads/SOGI-Measures-FAQ-Mar-2020.pdf>).

Complexities of Collecting High-Quality, Confidential Gender-Disaggregated Data and COVID-19’s Gap in Gender Data

Gender-disaggregated data are lacking for several reasons, including political resistance; laws that criminalize and persecute non-cisgender people; stigma and discrimination around collecting gender-disaggregated data; and a lack of adequate awareness and training for both data collectors and respondents around gender definitions, gender-sensitive language, and data protection and confidentiality.^{8,9}

The ongoing COVID-19 pandemic has once again brought to the forefront the importance of de-stigmatizing transgender and gender-nonconforming people, given that marginalized populations are often excluded from adequate testing, care and treatment, and vaccinations.^{7,10} Several studies and guidelines have emphasized the importance of sex-disaggregated COVID-19 data and highlighted sex-specific (cisgender female and cisgender male) differences in immunity, pharmacology, and vaccine outcomes.^{11–13} However, there is a global dearth of studies that examine these differences for transgender and gender-conforming people.

Steps Toward Collecting and Using Gender-Disaggregated Data: Beacons of Hope

Despite pressing challenges, there are encouraging examples of progress toward collecting, analyzing, and using gender-disaggregated data (see Table 2). For instance, in the United States, the Behavioral Risk Factor Surveillance System (BRFSS)

Table 2. Examples of surveys and data systems collecting gender-disaggregated data

Data system or survey	Key gender-disaggregation characteristics
Behavioral Risk Factor Surveillance System (BRFSS) (https://www.cdc.gov/brfss/index.html) —US geographic focus	<ul style="list-style-type: none"> • Includes options for cisgender female, cisgender male, transgender female, transgender male, and gender nonconforming data
People Living with HIV (PLHIV) Stigma Index (https://www.stigmaindex.org/about-the-stigma-index/) —Global focus	<ul style="list-style-type: none"> • Recruits data collectors from transgender and other populations • Emphasizes training for data collectors around gender sensitization and definitions
Census systems in Bangladesh, India, Nepal, and Pakistan (in phases from 2007 onward; https://www.ispionline.it/en/pubblicazione/third-gender-rights-south-asia-whats-new-25354) —South Asia geographic focus	<ul style="list-style-type: none"> • Includes option for respondents to state that they belong to “Other” gender category, distinct from female or male

collects data from nearly half a million adults each year, gathering data on health-related risk behaviors, chronic health conditions, and use of preventive services. It is often referred to as the largest continuously conducted health survey system in the world.¹⁴ The system has historically collected only cisgender female and cisgender male-specific data, but from 2014 onward it has been collecting gender-disaggregated data in most US states.¹⁵ Studies from BRFSS and other data sources have shown that gender-nonconforming respondents have significantly higher odds of reporting poor self-rated health than any other gender identity group.^{7,16} This is a relevant example of how, when our surveys and data systems are set up to collect and present gender-disaggregated data, those data can then be analyzed and shown to relevant stakeholders and policy makers.

Globally, strides in programming sex and gender-disaggregated data in data systems have been made in the field of HIV prevention and treatment.² One example is the People Living with HIV (PLHIV) Stigma Index, which is a standardized tool to gather evidence on how stigma and discrimination affects the lives of people living with HIV.¹⁷ The tool was first launched in 2008, and since then, it has been implemented in more than 100 countries, with more than 100,000 PLHIV participating in the process. A unique feature of this index is that it was developed to be used by and for people living with HIV. The index collects gender-disaggregated data and also intentionally trains and recruits interviewers who are living with HIV and interviewers from other key populations, such as men who have sex with men, transgender people, and gender nonconforming people.¹⁸ This helps increase familiarity and trust between respondents and data collectors, and thus is a best practice that surveyors and researchers can replicate in other data collection efforts. Studies analyzing PLHIV Stigma Index data show that transgender people experience high levels of stigma.¹⁹ If these data were not collected, it would be impossible to fill knowledge gaps and make informed decisions to benefit people living with HIV.

Another example of positive change comes from South Asian countries such as Bangladesh, India, Nepal and Pakistan. Thanks to concerted advocacy efforts, the past two decades have seen governments of these countries for the first time include the option of a third gender in their official national census and voter rolls.²⁰ Advocacy efforts continue in this area, especially because there is valid criticism that the census options are still limited in that they classify all people who do not identify as female or male into an “other” category, thereby failing to acknowledge and appropriately capture the nuances and fluidity of gender disaggregation (as highlighted in Table 1) of transgender and gender-nonconforming people.²¹

Recommendations to Capture High-Quality Gender-Disaggregated Data

Gender-Inclusive Questions in Data Collection

There have been several recommendations for how surveyors, researchers, and program implementors can improve data collection efforts to capture gender-disaggregated data as accurately as possible. One recommendation that has been getting increasing acceptance among surveyors is to include two sets of questions to capture the nuances of sex and gender-disaggregated data more comprehensively. One question asks respondents their assigned sex at birth (usually a binary option of female or male), and the other question asks respondents what they consider their current gender to be, with a range of response options that reflect the diversity and fluidity of gender.²² Tables 3 and 4 offer some concrete recommendations for gender disaggregation questions. Table 3 includes sex and gender within a single question (which could be confusing for

Table 3. Two-step gender status measurement questions

Question 1: What is your sex or gender? (Check ALL that apply)

- (A) Cisgender male
- (B) Cisgender female
- (C) Transgender
- (D) Transgender male/trans male/female-to-male
- (E) Transgender female/trans female/male-to-female
- (F) Nonbinary
- (G) Genderqueer
- (H) Genderfluid
- (I) Gender transitioning
- (J) Gender nonconforming
- (K) Other, please specify: _____
- (L) Decline to state

Question 2: What sex were you assigned at birth? (Check one)

- (A) Male
- (B) Female
- (C) Unknown or question not asked
- (D) Decline to state

Note: We acknowledge that although this is the currently used language and terminology that surveys and data collection methods use to capture gender identities, the terminology is constantly evolving—as it should—to better capture the nuances and diversity of all gender identities. One limitation of these particular questions is that there is no explicit “Intersex” response option, which could result in intersex respondents not being able to adequately identify their gender.

Source: Lombardi & Banik (<https://doi.org/10.1007/s13178-016-0220-6>); Bauer et al. (<https://doi.org/10.1371/journal.pone.0178043>).

data collectors and respondents), and Table 4 separates sex and gender in distinct questions.

Table 4. Question sequence with separate questions for gender and sex measurement

Gender Identity		
1: What sex were you assigned at birth, on your original birth certificate?		
(A)	<input type="checkbox"/>	Male
(B)	<input type="checkbox"/>	Female
2: How do you currently describe yourself?		
(A)	<input type="checkbox"/>	Male
(B)	<input type="checkbox"/>	Female
(C)	<input type="checkbox"/>	Transgender
(D)	<input type="checkbox"/>	Nonbinary
(E)	<input type="checkbox"/>	Gender nonconforming
(F)	<input type="checkbox"/>	None of these
Sexual Identity		
1: Which of the following best represents how you think of yourself:		
(A)	<input type="checkbox"/>	Gay or lesbian
(B)	<input type="checkbox"/>	Straight, that is, not gay or lesbian
(C)	<input type="checkbox"/>	Bisexual
(D)	<input type="checkbox"/>	Something else
(E)	<input type="checkbox"/>	I am not sure yet
(F)	<input type="checkbox"/>	Refused

Note: We acknowledge that although this is the currently used language and terminology that surveys and data collection methods use to capture gender identities, the terminology is constantly evolving—as it should—to better capture the nuances and diversity of all gender identities. One limitation of these particular questions is that there is no explicit “Intersex” response option, which could result in intersex respondents not being able to adequately identify their gender.

Source: University of California, Los Angeles (UCLA) Williams Institute Sexual Orientation and Gender Identity (SOGI) Adult Measures Recommendations (<http://williamsinstitute.law.ucla.edu/wp-content/uploads/SOGI-Measures-FAQ-Mar-2020.pdf>)

Gender-Sensitive Training for Data Collectors and Respondents to Ensure Gender Data Accuracy, Confidentiality, and Anonymity

Merely including gender-disaggregation options in data collection efforts is unlikely to automatically result in the capture of better gender-disaggregated data. Data collectors and respondents need to be primed to the importance of adequate training and gender sensitization when gender questions are asked.^{7,23} Particularly for marginalized respondents, political and societal stigma and discrimination can be barriers to disclosure of gender status. Data collection efforts are increasingly emphasizing culturally appropriate gender sensitization and assurance of anonymity and data protection. For example, the PLHIV Index emphasizes that PLHIV respondents (including transgender respondents

and respondents from other populations) have buy-in and ownership over the data collection and analysis process, as well as in advocacy efforts to reduce and eliminate stigma. In addition, the PLHIV Index recruits diverse teams of data collectors that include transgender and gender non-conforming people.¹³ This is a best practice that other data collection efforts can emulate, as studies have shown that respondents are more comfortable answering questions from data collectors that they can identify with on a personal level. We need to continue to balance the desire for high-quality gender data with the importance of ensuring data privacy, confidentiality, and anonymity. We also need to ensure that adequate human, financial, and time resources are allocated to gender-sensitive data collection efforts, including sensitization, training, and quality testing of gender-disaggregation questions and instruments. These efforts can generate gender evidence that in turn can educate policy makers about the complexity of gender and support marginalized gender communities in ways that can help reduce stigma.

Limitations of This Call to Action

The limitations to this call to action largely reflect the challenges associated with the lack of adequate and high-quality gender-disaggregated data across the world. For example, a key limitation behind the availability of gender-disaggregated data in some countries and regions is that it is not legal in those settings to acknowledge the presence of noncisgender people. Even where the law might not prohibit collection of gender-disaggregated data, including in resource-sufficient settings such as the United States, there is insufficient political will to collect gender-disaggregated data or lack of sensitization and training to ensure that gender-disaggregated data are collected in a thoughtful, respectful, and confidential manner. Lastly, because gender-disaggregated data are still sparsely available, there is not a rich repository of evidence in the global literature to show how gender-disaggregated data collection and analysis can lead to wide-reaching policy change. Having said that, experience from South Asian countries—as described elsewhere in this brief—shows that continued focus and advocacy efforts by and for marginalized gender populations has resulted in changes in discriminatory practices and the inclusion of noncisgender populations in national censuses.

A Call for Sustaining Gender Advocacy Efforts and Policy Change

Advocacy, sensitization, and training have helped us make huge strides and can continue to play an important role in ensuring that our surveys and data collection processes are adequately designed and set up to collect both sex and gender-disaggregated data. We still have a long way to go, however.

Surveyors, researchers, program implementors, and gender rights advocates can work together to design and implement surveys and data collection instruments that explicitly capture gender-disaggregated data. In addition, they can help ensure that both data collectors and respondents are adequately sensitized, trained, and resourced in gender definitions and respectful, nonjudgmental, and confidential data collection. Through these collective efforts, our data can be representative of all citizens and populations: cisgender, transgender, and gender nonconforming. Gender-disaggregated data in particular can make often invisible, hidden, and marginalized transgender and gender-nonconforming people more visible to national-level policy makers and governments. The evolution from largely binary sex-disaggregation to more-inclusive, multinomial gender-disaggregation that is representative of gender fluidity is an important step toward strengthening data quality and making those data more inclusive and equitable.

References

1. Victora C, Boerma T, Requejo J, Mesenburg MA, Joseph G, Costa JC et al. Analyses of inequalities in RMNCH: rising to the challenge of the SDGs. *BMJ Glob Health* 2019;4 Suppl 4:e001295. <https://doi.org/10.1136/bmjgh-2018-001295>
2. Sherwood J, Sharp A, Cooper B, Roose-Snyder B, Blumenthal S. HIV/AIDS National Strategic Plans of Sub-Saharan African countries: an analysis for gender equality and sex-disaggregated HIV targets. *Health Policy Plan* 2017;32(10):1361–7. <https://doi.org/10.1093/heapol/czx101>
3. Morgan R, George A, Ssali S, Hawkins K, Molyneux S, Theobald S. How to do (or not to do)... gender analysis in health systems research. *Health Policy Plan* 2016;31(8):1069–78. <https://doi.org/10.1093/heapol/czw037>
4. Hunt J. Introduction to gender analysis concepts and steps. *Dev Bull (Canberra)* 2004;64:100–6.
5. Begum T, Khan SM, Adamou B, Ferdous J, Parvez MM, Islam MS et al. Perceptions and experiences with district health information system software to collect and utilize health data in Bangladesh: a qualitative exploratory study. *BMC Health Serv Res* 2020;20(1):465. <https://doi.org/10.1186/s12913-020-05322-2>
6. Velte KC. Straightwashing the Census. *Boston Coll Law Rev* 2020;61(1):69–127. <https://lawdigitalcommons.bc.edu/bclr/vol61/iss1/3>
7. Romanelli M, Lindsey MA. Patterns of healthcare discrimination among transgender help-seekers. *Am J Prev Med* 2020;58(4):e123–31. <https://doi.org/10.1016/j.amepre.2019.11.002>
8. Spade D. *Normal life: administrative violence, critical trans politics, and the limits of law*. 2nd ed. Durham (NC): Duke University Press; 2015.
9. Day S, Mason R, Lagosky S, Rochon PA. Integrating and evaluating sex and gender in health research. *Health Res Policy Syst* 2016;14(1):75. <https://doi.org/10.1186/s12961-016-0147-7>
10. Kramer J. In Covid vaccine data, LGBTQ people fear invisibility. *The New York Times*. 2021 May 7. <https://www.nytimes.com/2021/05/07/health/coronavirus-lgbtq.html>
11. Vijayasingham L, Bischof E, Wolfe J; Gender and COVID-19 Research Agenda-setting Initiative. Sex-disaggregated data in COVID-19 vaccine trials. *Lancet* 2021;397(10278):966–7. [https://doi.org/10.1016/S0140-6736\(21\)00384-6](https://doi.org/10.1016/S0140-6736(21)00384-6)
12. Womersley K, Ripullone K, Peters SA, Woodward M. Covid-19: male disadvantage highlights the importance of sex disaggregated data. *BMJ* 2020;370:m2870. <https://doi.org/10.1136/bmj.m2870>
13. Pan American Health Organization. Why data disaggregation is key during a pandemic. 2020 [cited 04 Nov 2021]. Available from: <https://iris.paho.org/bitstream/handle/10665.2/52002/Data-Disaggregation-Factsheet-eng.pdf?sequence=17>
14. Centers for Disease Control and Prevention. Behavioral Risk Factor Surveillance System. [cited 2021 Nov 04]. Available from: <https://www.cdc.gov/brfss/index.html>
15. Centers for Medicare & Medicaid Services. Behavioral Risk Factor Surveillance System. [cited 2021 Nov 04]. Available from: <https://www.cms.gov/About-CMS/Agency-Information/OMH/resource-center/hcps-and-researchers/data-tools/sgm-clearinghouse/brfss>
16. Lagos D. Looking at population health beyond “male” and “female”: implications of transgender identity and gender nonconformity for population health. *Demography* 2018;55(6):2097–117. <https://doi.org/10.1007/s13524-018-0714-3>
17. Global Network of People Living With HIV (GNP+). What is the People Living with HIV Stigma Index? [cited 2021 Nov 04]. Available from: <https://www.stigmaindex.org/about-the-stigma-index/what-is-the-people-living-with-hiv-stigma-index/>
18. Global Network of People Living With HIV (GNP+). The principles of the People Living with HIV Stigma Index. [cited 2021 Nov 04]. Available from: <https://www.stigmaindex.org/about-the-stigma-index/the-principles-of-the-people-living-with-hiv-stigma-index/>
19. Friedland B, Gottert A, Hows J, Baral S, Sprague L, Nyblade L, et al. The People Living with HIV Stigma Index 2.0: generating critical evidence for change worldwide. *AIDS (London)* 2020;34Suppl 1.1S5–S18.
20. Ahmed I. Third gender rights in South Asia: what's new? Italian Institute for International Political Studies. 2020 May 9. Available from: <https://www.ispionline.it/en/publicazione/third-gender-rights-south-asia-whats-new-25354>
21. Sharma G, Banerji A. Nepal's next census to count LGBT+ people for the first time. *Reuters*. 2020 Feb 6. Available from: <https://www.reuters.com/article/us-nepal-lgbt-census/nepals-next-census-to-count-lgbt-people-for-the-first-time-idUSKBN201030>
22. Lombardi E, Banik S. The utility of the two-step gender measure within trans and cis populations. *Sex Res Soc Policy* 2016;13(3):288–96. <https://doi.org/10.1007/s13178-016-0220-6>
23. University of San Francisco LGBT Resource Center. General definitions. [cited 2021 Nov 04]. Available from: <https://lgbt.ucsf.edu/glossary-terms>

About the Authors

Rajeev Colaço, DrPH, is the director of MERLA (Monitoring, Evaluation, Research, Learning, and Adapting) in the International Development Group at RTI International.

Stephanie Watson-Grant, DrPH, is the deputy director of CHISU (Country Health Information Systems and Data Use) at John Snow International.

RTI Press Associate Editor

Phillip Graham

Acknowledgments

We would like to acknowledge the following people and entities for gender-inclusive input and contributions: Derek Kunaka, Alison Mitchell LeFew, Mzi Nduna, Laura Nyblade, Ina Wallace, Anna Wetterberg, Annie Gering, Claire Korzen, and RTI University's Equity, Diversity, Inclusion and Belonging (EDIB) Resource Library.

RTI Press Research Briefs and Policy Briefs are scholarly essays on policy, methods, or other topics relevant to RTI areas of research or technical focus.

RTI International, 3040 East Cornwallis Road, PO Box 12194

Research Triangle Park, NC 27709-2194 USA

+1.919.541.6000

rtipress@rti.org

www.rti.org

©2021 RTI International. RTI International is a trade name of Research Triangle Institute. RTI and the RTI logo are U.S. registered trademarks of Research Triangle Institute.



This work is distributed under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 license (CC BY-NC-ND), a copy of which is available at <https://creativecommons.org/licenses/by-nc-nd/4.0>

www.rti.org/rtipress