Open access Perspective



# Trans and gender diverse identities in adolescent health research: making the most of imperfect data

Emma Soneson , 1 Shona O'Leary, 1 Mina Fazel 10 1,2

# ► Additional supplemental material is published online only. To view, please visit the journal online (https://doi. org/10.1136/bmjment-2024-301150).

<sup>1</sup>Psychiatry, University of Oxford, Oxford, UK <sup>2</sup>Children's Psychological Medicine, Oxford University Hospitals NHS Foundation Trust, Oxford, UK

#### Correspondence to

Dr Emma Soneson, Psychiatry, Oxford University, Oxford, UK; emma.soneson@psych.ox.ac.uk

Received 10 May 2024 Accepted 13 September 2024

#### **ABSTRACT**

Advancing adolescent health research necessitates deliberate design and analysis that accurately captures the rapidly evolving world in which adolescents live and the ways in which they understand and express themselves and their experiences. In this Perspective, we reflect on how researchers might approach existing, imperfect data in a way that is accurate and inclusive of trans and gender diverse (TGD) adolescents. Drawing on our experience of running a large school-based survey of health and well-being (the OxWell Student Survey) and extensive coproduction with three TGD adolescents ('youth advisors'), we present considerations for critically appraising, processing and analysing quantitative gender data to better reflect adolescents' lived experiences. Specific topics discussed include how to assess the strengths and limitations of existing gender data; how to generate meaningful research data from free-text gender descriptions (and the implications of not doing so); and how to analyse data from TGD adolescents, including for those who choose not to disclose their gender. We conclude with a set of 10 recommendations, coproduced with the three youth advisors, for those working with 'imperfect' gender data.

#### INTRODUCTION

Advancing adolescent health research necessitates deliberate design and analysis that accurately captures the rapidly evolving world in which adolescents live and the ways in which they understand and express themselves and their experiences. This is especially relevant to the shifting landscape of gender identity. The most methodologically robust research from the last decade suggests that around 2.5%-8.4% of adolescents identify as trans and/or gender diverse (TGD), with clear upward trends,<sup>2</sup> and that these adolescents have distinct health profiles.<sup>3 4</sup> However, research about TGD adolescents is limited by issues including use of unclear or inappropriate sampling frames, selection bias and lack of information regarding how gender was operationalised.<sup>5</sup> Furthermore, given rapid societal shifts in how gender identity is understood and expressed, it is likely that research from even 5 years ago might not adequately capture the lived experiences of today's adolescents.

In this *Perspective*, we reflect on how researchers might approach existing gender data in a way that is more accurate and inclusive of TGD adolescents. Acknowledging the richness of how the wider social sciences have approached this topic and all those who have already encouraged more nuanced consideration of gender identity, we focus here on

large-scale studies including epidemiological and intervention studies. Furthermore, while others have written extensively about how to design more inclusive studies, 6-10 we consider how researchers can approach existing, imperfect gender data. We draw on our experience of leading the OxWell Student Survey, a large school-based survey of health and well-being, 11 and extensive coproduction with three 17–18-year-old TGD adolescents ('youth advisors'; see online supplemental appendix 1 for details).

# Critically appraising existing data

Inclusively and accurately measuring gender is paramount for all researchers working with adolescents. While this *Perspective* focuses on critically appraising existing data rather than on developing new measurement tools and approaches, we direct interested readers to a range of resources for improving measurement. 6-10 It is further important to note that best practice in this area is constantly evolving, 6 meaning that researchers would do well to consult current guidelines and ensure that TGD adolescents' perspectives are considered through active involvement and/or engagement with resources that have centred youth perspectives. 7 10

When working within the constraints of existing data, the first step in ensuring that analyses accurately reflect the experiences of TGD adolescents is to understand how gender has been measured. Details to consider include who provided the data on gender (informant), what they were asked for (construct), how they were asked (question framing and response options), whether they were asked for any identifiable information (as this might impact the accuracy of their responses), when they were asked (timepoint(s)) and where they were asked (setting/context, including level of privacy). These details form the basis for critically appraising how gender was measured. 12 Researchers may seek to appraise data according to current guidelines alongside principles such as maintaining a balance between inclusivity and data utility, allowing for participant autonomy in describing their gender and minimising potential harms.<sup>68</sup>

We can use OxWell to illustrate how researchers might critically appraise existing gender data. The 2023 OxWell wave asked a single question on gender: 'what is your gender? (options: male—female—other (with an option to self-identify)—prefer not to say)'. With this question, the survey conflated key constructs, using terms associated with sex ('male' and 'female') to describe gender. This kind of misalignment in researchers' intended



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. Published by BMJ.

**To cite:** Soneson E, O'Leary S, Fazel M. *BMJ Ment Health* 2024;**27**:1–5.



construct and adolescents' understanding of the question can result in misclassification that impacts findings and their interpretation and application. <sup>8</sup> <sup>13</sup> As queried by our youth advisors, for example, how would binary trans adolescents respond to this question? Would they choose their gender, their sex assigned at birth or 'other'? Would they self-describe? OxWell is not unique in this sense, as many other major studies of young people <sup>14</sup> <sup>15</sup> have also used sex-related terms to describe gender, exemplifying 'how researchers [...] make implicit and probably unconscious assumptions that conflate sex with gender' (Lindqvist, p334). <sup>5</sup>

OxWell's gender question has also been shaped by considerations pertaining to identifiability. One of the study's foremost strengths is non-identifiability, which can help promote greater participation and encourage more honest and accurate responses. Consequentially, OxWell uses a gender question that (1) does not explicitly capture information about gender modality (ie, experience of gender in relation to sex assigned at birth and (2) has a limited set of response options. While many other UK studies, such as Born in Bradford, the Millennium Cohort Study and BeeWell, include some measure of gender modality, our youth advisors still perceived these studies' overall range of response options as restrictive and suggested that gender data are more meaningful when there is a wider range of options (though again, the benefits of granularity must be carefully considered within the bounds of identifiability).

In addition to non-identifiability, there are two other main strengths about OxWell's gender question. First, the survey offers adolescents the opportunity to describe their own gender via a free-text field. Baams and Kaufman describe how '...given the fast-paced developments in how adolescents define and give meaning to their identities and experiences, research struggles to keep up. This makes it even more important to ask adolescents to explain, describe, and define their own experiences and identities as much as possible - while acknowledging that this is often difficult in population-based research' (Baams, p1014).9 Despite these benefits, an option to self-describe is, for a variety of reasons, not always provided (eg, in #BeeWell<sup>4</sup>). Second, OxWell provides a 'prefer not to say' option. While all OxWell survey questions are optional, the inclusion of this explicit option to not disclose gender promotes adolescent autonomy in terms of actively deciding what they want to share, rather than defaulting to being treated as missing data.

We wish to conclude this section by stressing that all data have limitations and the fact that existing data might have captured gender in an imperfect way does not necessarily mean that the data cannot provide useful insights. Instead, researchers can make best use of existing data by being *transparent* and *accountable*, <sup>12</sup> acknowledging the limitations of the data and what those might mean for the findings, their interpretation and their implications for research and practice.

#### Processing free-text gender descriptions

For datasets containing free-text gender descriptions, it is important to ensure that these responses are properly processed into meaningful research data. In the 2023 OxWell wave, 5.1% (1503/29271) of secondary school and further education students selected a gender other than 'male' or 'female': 3.1% (905/29271) selected 'prefer not to say', 0.2% (61/29271) selected 'other' and chose not to self-describe and 1.8% (537/29271) selected 'other' and provided a free-text description. We processed responses in two stages: first, by separating out free-text responses not considered to represent gender identities, and second, by categorising remaining responses.

#### Stage 1: Determining what is (not) gender

The first stage involved identifying responses *not* representing gender. In OxWell, 35.9% (193/537) of all free-text responses were not considered to represent gender identities, the vast majority of which were deemed to be likely disingenuous (our youth advisors labelled these 'silly or malicious'). These included responses that indicated students were likely not taking the survey seriously (eg, 'dinosaur'); ones that mocked the idea of TGD identities using common derogatory tropes (eg, 'attack helicopter', 'refrigerator' and 'plastic shopping bag'; youth input was particularly important here given how quickly these change); and a range of slurs. A small minority (<5%) of responses represented likely misunderstandings (eg, sexualities, religions, nationalities). We collectively refer to these responses as 'likely disingenuous/not gender'.

To demonstrate the need to identify and remove disingenuous responses, figure 1 presents the postprocessing distributions of adolescents' scores on the 11-item Revised Children's Anxiety and Depression Scale (RCADS-11)<sup>16</sup> according to gender groupings (TGD (including the n = 61 who chose 'other' but did not provide a free-text description), 'prefer not to say' and responses that were likely disingenuous/not gender). The distributions clearly vary, particularly on the extreme ends of the scale, where there is a higher density of adolescents reporting likely disingenuous genders, potentially suggesting low engagement with the survey more broadly. The preprocessing median and IQR for adolescents who selected a gender other than 'female', 'male' or 'prefer not to say' was 18 (11, 25). After processing, these were 11 (2.8, 24.2) for the 'likely disingenuous/not gender' group and 20 (13, 25) for the TGD group, meaning that failing to identify and separate out likely disingenuous responses would have slightly underestimated TGD adolescents' median self-reported depression and anxiety scores and slightly overestimated the IQR. This potential for misrepresentation has parallels to a previous finding that potentially mischievous responses on the American Youth Risk Behavior Survey resulted in biased estimates of disparities across sexual orientations. 17

# Stage 2: Categorising remaining responses

The second stage focused on categorising remaining free-text responses into analysable groupings. While rapid advances in artificial intelligence and natural language processing might subject to rigorous testing and ethical oversight—facilitate this process in the future, and although others have already suggested technology-assisted approaches,8 we chose to categorise manually to appreciate the nuance and complexity of our data. The relatively limited number of responses helped to make this manual process feasible. We worked iteratively with our youth advisors to develop two sets of categories, one broad and one more granular, each with potential (dis)advantages depending on the research question. These categories are presented in online supplemental table 1; however, categories will likely vary across studies, affected by factors including how questions were structured, populations of interest and sample sizes, and terminology will differ across region, culture and time.<sup>8 10</sup>

Two raters from the OxWell team (both cisgender women) used these categories to independently classify responses with 94% agreement. We found that the process was often not straightforward, and although our youth advisors thought that an opportunity to self-describe was important, they also reported feeling uncomfortable 'putting people into boxes' and believed there was not always enough information to do so accurately. While grouping individuals according to shared

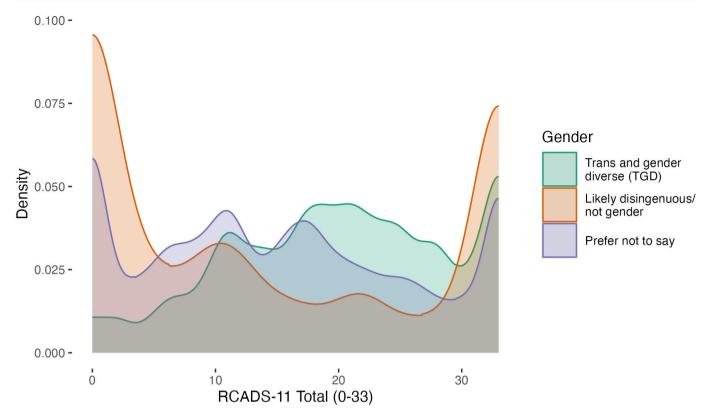


Figure 1 Distributions of the 11-item Revised Children's Anxiety and Depression Scale (RCADS-11) total scores according to gender groupings. Scores (0−33, with higher scores representing greater symptoms of depresison and anxiety) are mean-imputed by subscale, allowing for ≤1 item missing for each depression and anxiety subscales.  $N_{TGD}$ =330 (12% missing data required to calculate total RCADS-11 score);  $N_{Prefer not to say}$ =716 (21% missing data);  $N_{Likely disingenuous/not gender}$ =124 (36% missing data).

characteristics is often necessary in quantitative research, Rioux *et al* aptly describe the need to 'recognise the utilitarian purpose of these methods, where categorisations are made to advance the field but do not fully reflect reality, complex biosociocultural entanglements or any singular existence' (Rioux, p767). <sup>13</sup> Furthermore, while our present discussion primarily relates to quantitative research, we wish to highlight the central role of qualitative research in exploring TGD adolescents' lived experiences, as such approaches can capture nuances not readily apparent in quantitative data.

#### Analysing data from TGD adolescents

Our final section overviews some key considerations for analysing data collected from TGD adolescents. While the list of potential considerations is vast, we provide below five specific examples in relation to our OxWell experience.

# **Example 1: Analytical constraints**

When working with gender data, researchers may find themselves trying to balance inclusivity with the analytical constraints of current statistical approaches. For example, presenting participants with a wide range of gender response options and the opportunity to self-describe can support adolescents' autonomy while also leading to analyses that are underpowered to detect effects for adolescents with less common genders. There is no 'one-size-fits-all' approach here: one potential strategy for avoiding under-powered analyses is to report descriptive statistics for all gender subgroups and inferential statistics only for subgroups of a predetermined size, <sup>8</sup> and another is to use categorisations with varying levels of granularity according to the purpose of the analysis. In OxWell, we attempted the latter

approach through our 'broad' and 'granular' categorisations (online supplemental table 1) but acknowledge the limitations of using researcher-derived aggregated categories versus primary data.

#### Example 2: Undisclosed gender

Within OxWell, we found it challenging to determine how to analyse the sizeable group who choose not to disclose their gender (ie, select 'prefer not to say'; n=905, ~3% of our sample). We have found that this group has distinct response patterns, generally showing vulnerabilities greater than boys and girls but less than TGD adolescents (figure 1). Although there are many reasons why adolescents might choose not to disclose their gender, our youth advisors suggested that this group likely contains a high proportion of TGD adolescents. We explored this group in relation to OxWell's question on adolescents' concerns about their gender identity and found that 46.8% of these adolescents report being worried about their gender identity, compared with 73.3% of TGD respondents, 7.2% of girls and 5.2% of boys. This comparison highlights the importance of ensuring these adolescents are not treated as missing data, but rather that they are explicitly included in analyses and their experiences fully explored.

#### Example 3: Validated questionnaires

Another issue pertains to the gendered nature of commonly used measures. Ideally, questionnaires should be inclusive of TGD adolescents, but when working with existing data, this is not always possible. For example, the RCADS-11 has separate clinical cut-off scores for girls and boys, but not TGD adolescents. <sup>16</sup> In OxWell, 90.6% of TGD adolescents



Box 1 10 coproduced recommendations for critically appraising, processing, analysing and reporting gender diversity in adolescent health research

### Overarching principles

- Whenever possible, work in partnership with trans and gender diverse (TGD) adolescents to help ensure planned research is as inclusive as possible and that it accurately reflects their lived experience.
- 2. Ensure transparency by explicitly describing how gender has been measured, processed and analysed (eg, in relation to current guidelines and underlying principles of equality, diversity and inclusion). Consider reporting the gender identities of the research team members involved.
- Be mindful and deliberate about minimising the risk of harm for TGD adolescents (eg, misclassification, erasure), particularly where gender data have not been collected in an appropriate way.

# Critical appraisal

- 4. When working with existing data, determine and describe how gender was measured, including who was asked, what they were asked, how they were asked (including level of identifiability), when they were asked and where they were asked.
- Critically examine the strengths and limitations of the methodology (both in general and specifically for the research question(s) at hand) and consider how these might impact analyses and interpretation of findings.

#### Processing

- Determine whether free-text responses are likely to be genuine, as inclusion of disingenuous answers can have substantial implications for findings and interpretation.
- 7. Select or develop a system of how to categorise free-text responses into analysable categories. Depending on the research question(s), this may involve formulating multiple sets of categories. Ensure that at least two researchers are involved in processing free-text gender descriptions. Have a clear strategy for resolving disagreements in categorisations and report agreement in any outputs.

# **Analysis**

- 8. Use more granular gender categories where appropriate and statistically feasible, avoiding researcher-derived aggregated gender categories where possible. If both sex and gender data are available, use gender in analyses unless there is a strong rationale not to do so.
- Where available, use questionnaires that are inclusive of TGD adolescents (eg, ones that do not include only 'girl' and 'boy' options in the scoring guidelines or in participantfacing text).
- Consider how to analyse data from adolescents who choose not to disclose their gender. In some cases, it might be appropriate to include this group (or a part of it) within the TGD category for analyses.

would meet the RCADS-11 cut-off used for boys and 84.5% would meet the cut-off used for girls. There may be another, more appropriate cut-off to use for TGD adolescents, though this would require additional psychometric evaluation. Until then, it is important to be transparent about analytical decisions to ensure that data from TGD adolescents can be analysed. For example, in OxWell, we have decided to use the

more conservative RCADS-11 cut-off (ie, the one for girls) for TGD adolescents.

#### Example 4: Comparison with reference data

A more inclusive approach to gender also complicates comparison with reference data, as while research is increasingly considering TGD identities, this is often not the case for administrative data. For example, in the UK, routinely collected education data are limited to 'girl' and 'boy', which makes it challenging to understand sample representativeness or calculate weighted estimates reflective of the target population. For an interesting discussion of the technicalities of survey weighting, we refer readers to Kennedy and colleagues' simulation study and thoughtful consideration of ethics, accuracy, practicality and flexibility in survey-based assessment of gender identity.

# Example 5: Comparison over time

A final consideration pertains to comparability over time. This consideration is not unique to gender—those who have conducted longitudinal research will be familiar with the trade-off between including new and improved measures and ensuring comparability across waves. In OxWell, a repeated cross-sectional study, changes in how we have measured gender, while small, have already complicated comparison across waves. Such inconsistencies are often unavoidable when trying to foster more inclusive gender measurement, and we would not recommend that researchers continue to use outdated gender measures solely to maintain consistency. However, it is important that researchers are aware of this limitation and transparent about how they address it.

#### RECOMMENDATIONS

Based on our experiences with OxWell, we worked with our youth advisors to coproduce recommendations for how to be more inclusive of TGD adolescents in health research using existing, imperfect data (box 1), building on complementary articles exploring other aspects of TGD young people's perspectives on research.<sup>7 10</sup> Cross-cutting themes included:

- 1. The value of coproduction and lived experience in this rapidly evolving area.
- 2. The importance of capturing the rich diversity of gender identities throughout the research cycle.
- 3. The need for transparent reporting of how gender has been measured, processed and analysed.

#### **CONCLUSIONS**

Inclusive and accurate understanding of gender must be a priority for all adolescent health researchers. In the context of this rapidly evolving landscape, we encourage researchers not to shy away from existing, imperfect data, but rather to carefully consider their strengths and limitations and how decisions pertaining to measurement, processing and analysis might impact findings, interpretation and application. Transparency and accountability can help enable a richer understanding of today's adolescents.

**X** Emma Soneson @emma\_soneson and Mina Fazel @minafazeloxford

Acknowledgements We are immensely grateful for the valuable contributions and continued partnership of our youth advisors: Alice Merivale, Cerys Owen and Nick Warrier (all of whom provided written consent to be acknowledged). We also acknowledge Dr Holly Bear for her recruitment and facilitation of the wider BrainWaves YPAG and Tanya Manchanda for her efforts in categorising the free-text gender descriptions. We would like to thank all of the students who participated in the OxWell Student Survey, as well as the many school, local authority and research partners who put considerable effort into making this study possible, particularly Liverpool City Council, Milton Keynes City Council, NHS Frimley ICB,



NHS Buckinghamshire, Oxfordshire and Berkshire West ICB and Oxfordshire County Council in running the survey. Finally, we are grateful to our three peer reviewers, who provided exceptionally thoughtful feedback and in so doing greatly strengthened this Perspective.

**Contributors** ES, SOL and MF conceived the Perspective. ES and SOL jointly planned and facilitated the sessions with the youth advisors (Alice Merivale, Cerys Owen and Nick Warrier; see Acknowledgements). ES performed all statistical analyses. MF secured funding for the OxWell Student Survey and SOL secured funding for the research presented in this Perspective. ES drafted the manuscript, which was revised and approved by all authors. ES is the guarantor of the work.

**Funding** The 2023 OxWell Student Survey was supported by the NIHR Applied Research Collaboration Oxford and Thames Valley at Oxford Health NHS Foundation Trust (MF) and the Westminster Foundation; this research in this Perspective was supported by the NIHR Oxford Health Biomedical Research Centre (award number: CQR02480, HQ01.01). The views expressed here are those of the authors and not necessarily those of the NIHR or the Department of Health & Social Care.

Competing interests None declared.

Patient consent for publication Not applicable.

**Ethics approval** The OxWell Student Survey, from which data in this Perspective were drawn, was approved by the research ethics committee of the University of Oxford (R62366/RE0014).

Provenance and peer review Not commissioned; externally peer reviewed.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

#### ORCID iDs

Emma Soneson http://orcid.org/0000-0003-1666-3012 Mina Fazel http://orcid.org/0000-0001-9342-2365

#### **REFERENCES**

- 1 Fazel M, Soneson E. Current evidence and opportunities in child and adolescent public mental health: a research review. J Child Psychol Psychiatry 2023;64:1699–719.
- 2 Zhang Q, Goodman M, Adams N, et al. Epidemiological considerations in transgender health: A systematic review with focus on higher quality data. Int J Transgend Health 2020;21:125–37.
- 3 Tebbe EA, Budge SL. Factors that drive mental health disparities and promote well-being in transgender and nonbinary people. *Nat Rev Psychol* 2022;1:694–707.
- 4 Marquez J, Humphrey N, Black L, et al. Gender and sexual identity-based inequalities in adolescent wellbeing: findings from the #BeeWell Study. BMC Public Health 2023;23:2211.
- 5 Lindqvist A, Sendén MG, Renström EA. What is gender, anyway: a review of the options for operationalising gender. Psychology & Sexuality 2021;12:332–44.
- 6 Lowik A, Cameron JJ, Dame J, et al. Gender & sex in methods & measurement: tool #4: asking about & measuring participants' genders & sexes. 2022.
- 7 Wilson J, Sargeant C, Jago N, et al. Co-creating a new theory of gender beyond the binary: A Delphi study in the United Kingdom. J LGBT Youth 2024;21:549–73.
- 8 Cameron JJ, Stinson DA. Gender (mis)measurement: Guidelines for respecting gender diversity in psychological research. Soc Pers Psych 2019;13.
- 9 Baams L, Kaufman TML. Sexual Orientation and Gender Identity/Expression in Adolescent Research: Two Decades in Review. J Sex Res 2023;60:1004–19.
- 10 Dixon M, Hawke LD, Relihan J, et al. Let's Talk Gender: ten Things Transgender and Nonbinary Youth Want All Researchers to Know. J Am Acad Child Adolesc Psychiatry 2022;61:960–4.
- 11 Mansfield KL, Puntis S, Soneson E, et al. Study protocol: the OxWell school survey investigating social, emotional and behavioural factors associated with mental health and well-being. BMJ Open 2021;11:e052717.
- 12 Lowik A, Cameron JJ, Dame J, et al. Gender & sex in methods & measurement research equity toolkit: tool #6: working with pre-existing, secondary & older data. 2022.
- 13 Rioux C, Paré A, London-Nadeau K, et al. Sex and gender terminology: a glossary for gender-inclusive epidemiology. J Epidemiol Community Health 2022;76:764–8.
- Millennium Cohort Study. MCS age 23 interview questionnaire. Centre for Longitudinal Studies. 2023. Available: https://cls.ucl.ac.uk/wp-content/uploads/2023/ 11/MCS8-combined-questionnaire-October-2023.pdf [accessed Apr 2024]
- 15 Born in Bradford. BiB age of wonder module one (231) 2023-4. Born in Bradford; 2023. Available: https://borninbradford.nhs.uk/wp-content/uploads/AoW\_YPSurvey\_ 231-1.pdf [Accessed 25 Apr 2024].
- 16 Radez J, Waite P, Chorpita B, et al. Using the 11-item Version of the RCADS to Identify Anxiety and Depressive Disorders in Adolescents. Res Child Adolesc Psychopathol 2021;49:1241–57.
- 17 Cimpian JR, Timmer JD, Birkett MA, et al. Bias From Potentially Mischievous Responders on Large-Scale Estimates of Lesbian, Gay, Bisexual, or Questioning (LGBQ)—Heterosexual Youth Health Disparities. Am J Public Health 2018:108:5258—65
- 18 Kennedy L, Khanna K, Simpson D, et al. Using sex and gender in survey adjustment. 2020.