Sexual orientation and gender identity reporting in highly cited current alcohol research.

Dean J. Connolly, MA (Cantab) MSc,1,2* Santino Coduri-Fulford,3 Connor Tugulu,4 Meron Yalew,5 Elizabeth Moss,6 Justin C. Yang, PhD7,8

1Mortimer Market Centre, Central and North West London NHS Foundation Trust, London, United Kingdom.
2National Addiction Centre, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, London, United Kingdom.
3Queen Elizabeth The Queen Mother Hospital, East Kent Hospitals University NHS Foundation Trust, Margate, United Kingdom.
4Glasgow Royal Infirmary, NHS Greater Glasgow and Clyde, Glasgow, United Kingdom.
5UCL Medical School, University College London, London, United Kingdom.
6Chelsea and Westminster NHS Foundation Trust, London, United Kingdom.
7UCL Division of Psychiatry, University College London, London, United Kingdom.
8Camden and Islington NHS Foundation Trust, London, United Kingdom.

*Correspondence:
Dean J. Connolly MA (Cantab) MSc
National Addiction Centre, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, 4 Winsor Walk, London, United Kingdom, SE5 8AF.

Email: dean.1.connolly@kcl.ac.uk
Phone: +447824118910
Twitter: @DJConnolly94

**Running head:** LGBTQ+ data reporting in alcohol research.

**Word count:** 2,133 (excluding title page, references, tables, figure)

**Abstract word count:** 244

**Tables:** 3

**Figures:** 1

**Key words:** alcohol; citation; data collection; gender identity; sexual orientation; sexual and gender minority
Abstract

Purpose
This study aimed to measure the frequency of high-quality and transparent sexual orientation and gender identity (SOGI) data collection and reporting in highly cited current alcohol use research, using the extant literature to identify community-informed priorities for the measurement of these variables.

Methods
A single search to identify alcohol use literature was conducted on PubMed with results restricted to primary research articles published between 2015 and 2022. The 200 most highly cited studies from each year were identified and their titles and abstracts reviewed against inclusion criteria following deduplication. After full-text review, study characteristics and data indicating quality of SOGI reporting were extracted. The fidelity of the results was verified with a random sample before analyses.

Results
The final sample comprised 580 records. Few studies reported gender identity (n=194; 33.4%) and, of these, 7.2% reported the associated gender identity measure. A two-stage approach to measure gender was adopted in three studies, one study used an open-ended question with a free-text response option, and 13 studies recorded non-binary gender identities (reported by 0.9% of the whole sample). Nineteen (3.3%) studies reported sexual orientation and more than half of these provided the sexual orientation measure. Eight of the 20 studies which reported sexual orientation and/or gender identity measures were classified as sexual and gender minority specialist research.

Conclusions
Culturally competent SOGI reporting is lacking in highly cited current alcohol research. SOGI measures should be disclosed in future research and should provide free-text response options.
Introduction

A substantial body of research suggests that, relative to their majority counterparts, lesbian, gay, bisexual, transgender (trans), queer and other sexual or gender minority (LGBTQ+) people experience a greater prevalence of high-risk or dependent drinking\(^1,2\) and are more likely to experience harms such as alcohol-related blackouts.\(^3\) This disparity is likely to widen as antecedents of problem alcohol use observed in this population (e.g., anti-LGBTQ+ legislation, hate crime and dehumanising media discourse) are increasing in prevalence and severity internationally.\(^2,4^-{10}\)

Despite the high acceptability, among both the general population and LGBTQ+ communities, of sexual orientation and gender identity (SOGI) data collection,\(^11\) SOGI is rarely recorded in electronic health records, health data systems and large epidemiological surveys,\(^12\) resulting in a paucity of quantitative data. Inclusion in these datasets is necessary to monitor trends in drinking, the scale and correlates of harm as well as engagement with and response to health care. A recent United Kingdom National Institute for Health and Care Research (NIHR) review identified no NIHR-funded randomised controlled trials (RCTs) which collected SOGI data, suggesting this disparity is also an issue in experimental research.\(^13\)

Alcohol use research with LGBTQ+ people has largely involved within-group studies which recruit small convenience samples, frequently comprised of people with an additional shared vulnerability (e.g., sex work), primarily aimed at HIV risk reduction or understanding the relationship between LGBTQ+ status and alcohol harm (specialist research).\(^2,14\) The extent to which sexual or gender minority status are recorded or used in general population alcohol
research and how this practice compared to specialist research is unclear and bears further scrutiny.

Existing studies with LGBTQ+ samples have identified several key characteristics which define good SOGI data collection and reporting practice. The first relates to the inclusion of non-binary/gender diverse people in data collection and reporting. While the data collected with this approach is unlikely to generate sufficient data for stratified analysis, non-binary/gender diverse participants should still be identified and their data summarised and reported with a view to informing future meta-analyses.

The literature overwhelmingly supports the use of a two-stage approach (i.e., asking gender identity and birth-registered sex/assigned gender) to understand both gender identity and trans status. Quantitative and qualitative work provide evidence of high sensitivity and specificity with this method, which is largely accepted within the LGBTQ+ community. An open-ended response option when measuring SOGI was generally endorsed in the extant literature, with one study participant advising “put a line and let [us] put what [we] want [our] damn self”, highlighting the potential for identity invalidation with aggregated response categories.

This aim of this study was to measure the frequency of high-quality and transparent SOGI data collection and reporting in highly cited current alcohol use literature. A secondary aim was to identify patterns in current practice with measures described above as indicators of good practice.

Methods
Alcohol was selected as the sole substance of interest because it is the most ubiquitously used intoxicant globally,\textsuperscript{20} and there is a substantial and growing body of specialist LGBTQ+-alcohol literature.\textsuperscript{2,21} Highly cited current literature was investigated for two reasons. The first related to the observation that specialist literature was cited infrequently, relative to the whole sample. By investigating highly cited research, we were able to address the risk that specialist research in the sample might inflate estimates of SOGI representation. Secondly, LGBTQ+ people deserve to be represented in and benefit from the most impactful research, as a matter of health equity. If this is not currently the case, it must be highlighted and addressed. As well, we focused on current literature (published 2015 onwards) which is meant to represent up-to-date methodological approaches and is most relevant for practitioners and policymakers.

This article presents a secondary analysis of SOGI recording and reporting in highly cited current alcohol use research. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) informed the sampling (i.e., search, title-abstract screening) and data extraction.\textsuperscript{22}

\textit{Search strategy}
A single search was conducted on 31st May 2022. PubMed, anticipated to yield the most highly cited literature, was searched with MeSH terms “alcoholism” (MeSH ID: D000437) and “binge drinking” (MeSH ID: D063425) and key words “alcohol use disorder”, “alcohol consumption”, and “alcohol dependence”, all combined with the Boolean classifier “OR”. Results were then restricted to years 2015-2022, inclusive, and the following article type filters applied: Clinical Study, Clinical Trial, Clinical Trial, Phase III, Clinical Trial, Phase IV, Comparative Study, Controlled Clinical Trial, Multicenter Study, Observational Study,
Pragmatic Clinical Trial, Randomized Controlled Trial, Twin Study, Validation Study, Humans.

Bibliographic data from all records identified by the search were downloaded to Zotero, a reference management software. A software ‘add-on’, Zotero Citation Counts Manager 1.3.0, was applied to extract, from Crossref, the number of times each record had been cited. The reliability of the citation count was tested using a random sample (n=30) from a pilot search (100% concordance). The 200 most highly cited records from each year were retained and collated to give a study sampling frame of 1,600 records.

**Inclusion and exclusion criteria**

With no restriction by study design, all original qualitative, quantitative, or mixed-methods studies, with more than ten human participants, and at least one alcohol use variable were included. Research with non-human subjects and all forms of journal communication not classified as original research, including (systematic) reviews, case reports and series, were excluded. Conference proceedings, books and chapters were also excluded.

**Record selection and data extraction**

Bibliographic data for the entire sample were uploaded to Rayyan, a systematic review software, which has a partially automated de-duplication function. All authors were given access to the database and a minimum of two independently assessed each abstract against the inclusion criteria. Conflicts were resolved by the first author who made the final decision regarding inclusion.
Included records were then divided equally among all authors. A study’s eligibility was confirmed by a review of the full text and a piloted data extraction table was populated with data from each study by at least one author. These data included study characteristics (DOI, authors, year of publication) and a list of pre-specified criteria suggesting high- or low-quality SOGI data collection. These criteria began with determination of whether SOGI was measured (with SOGI measures collected verbatim, if available). Also assessed was the use of a two-stage approach to measuring gender identity and whether an open-ended response option was provided for both gender identity and sexual orientation measures. The recording and reporting of non-binary/gender diverse participants’ data was assessed, as was the frequency of two sets of response options to measure gender identity which represent poor practice: 1. male, female, transgender; 2. male, female, prefer not to say. The operationalisation of each variable was reviewed with and agreed by the whole team to ensure consistency.

Analysis

A random sample (5%) of the data was reviewed by the first author to confirm fidelity. The number of records fulfilling each pre-specified outcome was described as a count and percentage of eligible records (e.g., the number of articles reporting a gender identity measure was reported as a percentage of those which reported participants’ gender identity). Verbatim SOGI measures were tabulated.

Results

Search results

A total of 29,096 records were identified from 2015 (n=3,737), 2016 (n=3,710), 2017 (n=3,939), 2018 (n=3,918), 2019 (n=3,883), 2020 (n=4,156), 2021 (n=4,210) and 2022
(n=1,543). Following deduplication (N=1,450), title-abstract screening (N=621), full-text review and retrieval (N=581) and exclusion of one retracted article, the final sample comprised N=580 (Figure 1).

**SOGI recording and reporting practices**

One hundred and ninety-four studies (33.4%) reported participants’ gender identity. Of these, 14 (7.2%) reported the gender identity measure (e.g., the question/query/prompt and response options given to study participants; Table 1). Three studies (21.4%; 0.5% of whole sample) adopted the two-stage approach described in the Introduction, and one used an open-ended question with a free-text response option.37

Thirteen studies (6.7% of those reporting gender identity) recorded non-binary identities29–31,32,34–36,38–44 and, of these, five (38.5%; 0.9% of whole sample) reported non-binary participants’ data.34,39–42 Five studies listed ‘male’, ‘female’ or ‘transgender’ as mutually exclusive response options.26–28,30,33 One study gave ‘male’, ‘female’, or ‘prefer not to say’ as mutually exclusive response options.45 Sexual orientation was reported in 19 studies (3.3%). Eleven of these (57.9%) reported the sexual orientation measure (Table 2).26–28,33,36,46–51 Of the 20 unique studies reporting sexual orientation and/or gender identity measures, eight studies (40.0%) were SOGI specialist research.33,36,39,46,48–51 There was no additional SOGI specialist research in the wider sample.

Table 3 summarises the characteristics of studies reporting sexual orientation and/or gender identity measures. Of these studies, eleven recruited participants from the United States. Sixteen were cross-sectional surveys. Most obtained representative (n=8) or convenience
samples (n=7) and young people (n=7) were investigated as frequently as the general adult population (n=7).

**Discussion**

*Summary of key findings*

One third of studies reported participants’ gender identity. However, markedly fewer (<1%) reported the associated measure with a high proportion indicating a poor understanding of gender identity. Only one used an open-ended gender identity measure. Non-binary/gender diverse participants’ data was recorded in a small minority of studies. Less than half of these reported stratified analyses of non-binary/gender diverse participants’ data. While very few studies reported sexual orientation, a greater proportion of sexual orientation than gender identity measures were reported. Almost half of the studies reporting a SOGI measure were classified as specialist research.

*Findings in context*

Corroborating the findings presented here, a 2015 study similarly identified poor representation of LGBTQ+ people in alcohol research. Examining research published in 2007 and 2012, the authors found that sexual orientation was reported in 2.3% (PsycINFO) and 6.4% (PubMed) of sampled “substance abuse” articles from 2012. Non-binary gender was reported in 2.3% (PsycINFO) and 1.9% (PubMed) of the same sample. The authors observed a negligible improvement from 2007. Comparing these results with the findings of the present (2015-2022) study, we observed lower rates of reporting in our study compared with earlier studies and the reasons for this are unclear.
Gender identity was recorded and reported with much greater frequency than sexual orientation. Failure to consider sexual orientation is a glaring omission in highly cited current alcohol use literature. There are long-established disparities in alcohol harm experienced by sexual minority groups. Associated with discrimination and the cultural significance of the “gay bar”, LGB+ people are more likely than heterosexual counterparts to drink alcohol, report heavy episodic or daily use, and meet criteria for alcohol use disorders.

Poor recording and reporting of ethnicity has also been observed. A recent systematic review examining RCTs of pharmacotherapies for alcohol use disorder found that 49.0% of included records had not reported their participants’ ethnicity. While the difference in population size and the circumstances of their exclusion preclude direct comparison, it appears both these minoritised groups are underrepresented in alcohol research.

**Strengths and limitations**

A limitation of this study was its partial adherence to PRISMA (i.e., double, independent title-abstract review, pre-specified inclusion/exclusion criteria and study variables, and piloted data extraction with accuracy checking). Limited resources precluded double screening of full-texts and whole sample double data extraction.

Conflation between gender and sex in the primary literature meant it was frequently difficult to determine whether participants’ gender identity had been reported. Interchangeable use of terms “sex” and “gender”, “female” and “woman”, “male” and “man” between and within articles may have resulted in misestimation of the frequency of gender identity reporting. These findings may not be generalisable to the wider alcohol use literature as only the most highly cited studies from one database were sought and included. However, inclusion in the
most impactful research is a matter of health equity. Minority groups are entitled to be represented in studies which are more readily translated into public, group, or individual health interventions.

**Implications for policy, research, and practice**

Inclusion of community- and expert-informed SOGI measures in all observational and experimental research should be enforced by grantors and research ethics committees. Healthcare providers should be supported to adapt their electronic records to collect SOGI data with cultural competence being mindful that approaches will likely evolve with time.

The exclusion of LGBTQ+ people from alcohol research, through inadequate recruitment, data collection or statistical stratification may mean that caution is required when administering interventions in the absence of valid outcome (favourable or adverse) data. A person-centred approach to supporting alcohol service users is required.

Despite LGBTQ+ people reporting dissatisfaction with the use of broad catch-all response options to supposed gender identity measures (e.g., ‘transgender’), these are not uncommon. In future, researchers should consider using a qualitative measure to empower participants to disclose their exact gender identity. Code to categorise these data for analysis has been trialled with success.\(^\text{17}\)

**Conclusion**

Transparent and culturally competent SOGI reporting is lacking in highly cited current alcohol research. Alcohol researchers must comprehensively assess and document SOGI to
fully understand and appropriately respond to the disproportionate alcohol-related harm experienced in LGBTQ+ communities.

**Acknowledgments**

None.

**Author Contributions**

**Dean Connolly:** Conceptualisation (lead), Formal analysis (lead), Investigation (lead), Methodology (lead), Project administration (lead), Supervision (lead), Writing – Original Draft (lead), Writing – Review & Editing (lead)  

**Santino Coduri-Fulford:** Investigation (equal), Writing – Review & Editing (equal)  

**Connor Tugulu:** Investigation (equal), Writing – Review & Editing (equal)  

**Meron Yalew:** Investigation (equal), Writing – Review & Editing (equal)  

**Elizabeth Moss:** Investigation (equal), Writing – Review & Editing (equal)  

**Justin Yang:** Formal analysis (equal), Investigation (equal), Methodology (lead), Writing – Review & Editing (lead)

**Disclaimer**

None.

**Author disclosures**

The authors have no conflicts of interest to declare.

**Funding**


This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

References


<table>
<thead>
<tr>
<th>Authors (year)</th>
<th>Gender identity measure (i.e., question or stem) and response options</th>
</tr>
</thead>
</table>
| Ebert et al. (2019),26 Auerbach et al. (2018, 2019)27,28 | Gender was assessed by asking respondents whether they identified as being male, female, transgender (male-to-female/female-to-male), or “other”

| Agley & Xiao (2021)29                     | Q: Please indicate your gender  
R: male, female, non-binary, transgender |
|---------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Agley et al. (2021)30                   | Q: Please indicate your gender  
R: female, male, transgender, other |
| Callinan et al. (2021)31                | Participants were asked for their “self-identified gender”  
R: male, female, non-binary, not-listed gender |
| Kilian et al. (2021)32                  | Q: please specify your gender  
R: male, female, other |
| Coulter et al. (2018)33                 | Q1 (to ascertain LGBT status): which of the following best describes you? Mark all that apply  
R1: heterosexual (straight), gay or lesbian or bisexual, transgender, not sure (removed from analysis), decline to respond (removed from analysis)  
Researchers “created a three-category measure of gender by coding participants who selected “transgender” as transgender adolescents, and those who did not select this option were coded as cisgender males or cisgender females based on their responses to the following question”  
Q2: what is your sex?  
R2: male, female |
| Davies et al. (2022)34                  | Q1: What is your gender?  
R1: male, female, non-binary, different identity  
Q2: What gender were you assigned at birth?  
R2: male, female |
| Every-Palmer et al. (2020)35            | Q1: which gender do you identify with?  
R1: male, female, gender diverse  
Q2: are you?  
R2: transgender female to male, transgender male to female, intersexed, gender non-conforming, genderqueer, two-spirit, third gender

| Hegazi et al. (2016)36                  | “gender identity…stated by the PERSON” (during clinical contact/appointment booking etc.)  
R1: man (including trans man), woman (including trans woman), non-binary, other (not listed), not stated  
Q: Is the patient’s gender identity the same as the gender assigned at birth?  
R2: yes, no, not stated, not known |
| Li et al. (2020)37                       | Q: Your gender:  
R: free-text option only

| Thompson et al. (2021)38                | Q: what is your gender identity?  
R: man, woman, non-binary |
| Reisner et al. (2016)39                 | Q not reported (“eligible participants were assigned male sex at birth”)  
R: woman, female, transgender woman, transfemale, male-to-female, other identity on the transfeminine spectrum |

**Notes:** a: responses to this question were not presented in the article. b: though only data for males and females included in report. The exclusion of transgender people was acknowledged as a limitation; LGBT: lesbian, gay, bisexual, transgender; Q: question/query; R: response
<table>
<thead>
<tr>
<th>Authors (year)</th>
<th>Sexual orientation measure (i.e., question or stem) and response options</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ebert et al. (2019), Auerbach et al. (2018, 2019)</td>
<td>Sexual orientation was classified into heterosexual, gay or lesbian, bisexual, asexual, not sure, and other. Additional questions were asked about the extent to which respondents were attracted to men and women and the gender(s) of people they had sex with (if any) in the past 5 years. Respondents were categorized into the following categories: heterosexual with no same-sex attraction, heterosexual with same-sex attraction, non-heterosexual without same-sex sexual intercourse, and non-heterosexual with same-sex sexual intercourse.</td>
</tr>
</tbody>
</table>
| Coulter et al. (2018) | Q: which of the following best describes you?  
R: heterosexual (straight), gay or lesbian or bisexual, transgender, not sure, and decline to respond |
| Hegazi et al. (2016) | The patient's sexual orientation as stated by the patient.  
Q: what is the patient’s sexual orientation?  
R: heterosexual or straight, gay or lesbian, bisexual, other sexual orientation not listed, PERSON asked but does not know or is not sure, not stated (PERSON asked but declined to provide a response) |
| Jones et al. (2020) | Q: which of the following best describes you?  
R: heterosexual (straight), gay or lesbian, bisexual, not sure |
| Evans-Polce et al. (2020) | Individuals were asked to report which category best described them  
R: heterosexual, gay or lesbian, bisexual, not sure |
| Roxburgh et al. (2016) | Q: Do you think of yourself as...  
R: heterosexual or straight, homosexual (gay or lesbian), bisexual, not sure/undecided, something else/other |
| Schuler & Collins (2020) | Q: Which one of the following do you consider yourself to be?  
R: heterosexual (that is, straight), lesbian or gay, bisexual, don’t know |
| Slater et al. (2017) | To assess sexual identity, respondents were shown a preprinted response card and asked to select the category that best described them.  
R: heterosexual (straight), gay or lesbian, bisexual, not sure |
| Gonzales et al. (2016) | Respondents 18 years or older were asked which of the following categories best represents how they thought of themselves:  
R: lesbian or gay, straight (that is, not gay), bisexual, something else, I don’t know the answer, refuse |
Table 3: Characteristics of highly cited alcohol research reporting sexual orientation and/or gender identity (SOGI) measures (published 2015-2022)

<table>
<thead>
<tr>
<th>Authors (year)</th>
<th>Country of study</th>
<th>Study design</th>
<th>Sample type</th>
<th>Study population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ebert et al. (2019),26 Auerbach et al. (2018, 2019)27,28</td>
<td>Multi-national (8 countries)</td>
<td>Cross-sectional survey</td>
<td>Convenience</td>
<td>College students</td>
</tr>
<tr>
<td>Agley &amp; Xiao (2021)29</td>
<td>United States</td>
<td>Cross-sectional survey</td>
<td>Crowdsourced</td>
<td>MTurk users</td>
</tr>
<tr>
<td>Agley et al. (2021)30</td>
<td>United States</td>
<td>Randomised controlled experiment</td>
<td>Crowdsourced</td>
<td>MTurk users</td>
</tr>
<tr>
<td>Callinan et al. (2021)31</td>
<td>Australia</td>
<td>Cross-sectional survey</td>
<td>Convenience</td>
<td>Adults who drink alcohol at least monthly</td>
</tr>
<tr>
<td>Kilian et al. (2021)32</td>
<td>Multi-national (21 countries)</td>
<td>Cross-sectional survey</td>
<td>Convenience/purposive</td>
<td>General population ≥18 years old</td>
</tr>
<tr>
<td>Coultet et al. (2018)33</td>
<td>United States</td>
<td>Cross-sectional survey</td>
<td>Representative</td>
<td>Middle and high school students</td>
</tr>
<tr>
<td>Davies et al. (2022)34</td>
<td>Multi-national (13 countries)</td>
<td>Cross-sectional survey</td>
<td>Convenience</td>
<td>Adults reporting alcohol use in 30 days preceding survey</td>
</tr>
<tr>
<td>Every-Palmer et al. (2020)35</td>
<td>New Zealand</td>
<td>Cross-sectional survey</td>
<td>Representative</td>
<td>General population 18-90 years old</td>
</tr>
<tr>
<td>Hegazi et al. (2016)36</td>
<td>United Kingdom</td>
<td>Retrospective case notes analysis</td>
<td>Convenience</td>
<td>MSM attending one of two SHCs</td>
</tr>
<tr>
<td>Jones et al. (2020)37</td>
<td>United States</td>
<td>Cross-sectional survey</td>
<td>Representative</td>
<td>High school students</td>
</tr>
<tr>
<td>Li et al. (2020)37</td>
<td>China</td>
<td>Cross-sectional survey</td>
<td>Convenience</td>
<td>WeChat/Weibo (social media platform) users 15-35 years old</td>
</tr>
<tr>
<td>Thompson et al. (2021)38</td>
<td>Canada</td>
<td>Cross-sectional survey</td>
<td>Representative</td>
<td>General population ≥19 years old</td>
</tr>
<tr>
<td>Reisner et al. (2016)39</td>
<td>United States</td>
<td>Cross-sectional survey</td>
<td>Convenience</td>
<td>Transgender women 16-29 years old</td>
</tr>
<tr>
<td>Evans-Polce et al. (2020)40</td>
<td>United States</td>
<td>Cross-sectional survey</td>
<td>Representative</td>
<td>LGB+ people 18-50 years old</td>
</tr>
<tr>
<td>Roxburgh et al. (2016)41</td>
<td>Australian</td>
<td>Cross-sectional survey</td>
<td>Multistage stratified</td>
<td>General population ≥15 years old</td>
</tr>
<tr>
<td>Schuler &amp; Collins (2020)42</td>
<td>United States</td>
<td>Cross-sectional survey</td>
<td>Representative</td>
<td>General population ≥12 years old</td>
</tr>
<tr>
<td>Slater et al. (2017)43</td>
<td>United States</td>
<td>Cross-sectional survey</td>
<td>Representative</td>
<td>General population ≥18 years old</td>
</tr>
<tr>
<td>Gonzales et al. (2016)44</td>
<td>United States</td>
<td>Cross-sectional survey</td>
<td>Representative</td>
<td>General population ≥18 years old</td>
</tr>
</tbody>
</table>

Notes: MTurk: Amazon Mechanical Turk; LGB+: lesbian, gay, bisexual or other sexual minority; MSM: men who have sex with men; SHC: sexual health clinic
Fig 1. Flowchart of record selection

Search results (N=29,096)
2015 (n=3,737)
2016 (n=3,710)
2017 (n=3,939)
2018 (n=3,918)
2019 (n=3,883)
2020 (n=4,156)
2021 (n=4,210)
2022 (n=1,543)

Citation counter

N=1,600 highly cited records

Deduplication

N=1,450 records

Abstract screening & article retrieval

Final sample N=580