

## Protocol

# Retrieving Transgender and Gender Diverse Literature: Protocol for the Development and Validation of 2 Search Hedges

Chi Dinh<sup>1</sup>, BHSc; Zack Marshall<sup>1,2,3</sup>, MSW, PhD; Scott Marsalis<sup>4</sup>, BA, MLiS; Ryn Gagen<sup>5</sup>, BA, MLiS; Preet Kang<sup>6</sup>, BA; Avery Everhart<sup>7</sup>, BA, MA, PhD

<sup>1</sup>Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

<sup>2</sup>School of Social Work, McGill University, Montreal, QC, Canada

<sup>3</sup>O'Brien Institute for Public Health, University of Calgary, Calgary, AB, Canada

<sup>4</sup>Sciences, Agriculture & Engineering, University of Minnesota Libraries, University of Minnesota, Saint Paul, MN, United States

<sup>5</sup>Health Sciences Library, University of Minnesota, Saint Paul, MN, United States

<sup>6</sup>Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, BC, Canada

<sup>7</sup>Department of Geography, Faculty of Arts, University of British Columbia, Vancouver, BC, Canada

## Corresponding Author:

Zack Marshall, MSW, PhD

Department of Community Health Sciences

Cumming School of Medicine

University of Calgary

3280 Hospital Dr NW

Calgary, AB, T2N 4Z6

Canada

Phone: 1 4032206940

Email: [zack.marshall@ucalgary.ca](mailto:zack.marshall@ucalgary.ca)

## Abstract

**Background:** Searching for transgender and gender diverse (TGD) references within large academic databases can be a challenging process, partly due to the dynamic and diverse definitions of words and terminologies used by multiple interest holders. Search hedges are preestablished search strings that aid in the efficacy of identifying and screening relevant articles. Validated search hedges focused on TGD people and topics will aid in identifying relevant literature.

**Objective:** This study aims to develop and validate the sensitivity and precision of 2 interdisciplinary and cross-cultural TGD search hedges designed for retrieving references from MEDLINE and APA PsycInfo, both on the Ovid platform.

**Methods:** Searches were conducted using the finalized search hedges via Ovid on June 7, 2024, yielding 31,055 references from MEDLINE and 22,924 references from APA PsycInfo. A random sample of 2330 records from MEDLINE and 2293 records from APA PsycInfo will be independently screened by at least 2 team members. At the title and abstract screening stage, references will be excluded if they (1) use solely binary terminology to describe gender, (2) focus on psychometric measurement of gender, or (3) focus on intersex or differences of sex development (DSD) topics. References will be included if they (1) report on transgender or gender diverse people, or both, in their sample; or (2) specifically discuss TGD communities or TGD topics. References without an abstract will be categorized as *No\_Abstract*. References in which the TGD population is unclear will be categorized as *LGB\_Maybe\_T* or *Mixed\_Topics*. Only references in the *No\_Abstract*, *LGB\_Maybe\_T*, or *Mixed\_Topics* categories will proceed to the full-text screening phase. In the full-text screening phase, references will be categorized as included if they (1) clearly distinguish between sexual identity and gender identity, (2) mention or discuss TGD topics or experiences in the Methods or Results sections, (3) communicate consideration for participants' gender self-identification and experiences, or (4) consider TGD populations as a distinct subpopulation. The results of the screening process will be used to calculate precision and sensitivity, with a targeted sensitivity of 100% and a targeted precision of 76% for each search hedge.

**Results:** Validation and data analysis are projected to be finished by December 2025, with results expected to be published in 2026.

**Conclusions:** Rigorous and transparent knowledge synthesis processes, starting with a high-quality search hedge, can help inform and equip community members, clinicians, policymakers, and other key decision-makers with scientifically sound evidence.

**International Registered Report Identifier (IRRID):** DERR1-10.2196/76055

## KEYWORDS

evidence synthesis; database searching; search hedge validation; search filter validation; Two-Spirit; transgender; nonbinary; gender minority; information justice

## Introduction

### Background

Knowledge synthesis, also called evidence synthesis, is the process of identifying, evaluating, and critically integrating all available evidence on a topic [1]. This process is a foundational practice for the development of evidence-based decision-making [1]. Effective and comprehensive search strategies are one of the core components of knowledge synthesis [2]. To support robust search strategies, search hedges are increasingly used. A search hedge consists of pre-established search strings designed to comprehensively retrieve references in a topic area from one or multiple bibliographic databases [3]. A search hedge is typically developed by librarians working with subject matter experts and should be validated for its performance [4]. Validated search hedges can help ensure all relevant literature is retrieved, reduce the number of irrelevant articles screened, and increase the efficiency of resource use in the knowledge synthesis process.

In addition to hedges with a specific subject matter focus, more recent publications have emphasized other search attributes, such as geographic, methodological, or population-specific searches [5-7]. While search hedges focused on combined sexual and gender minority groups are available, there are currently no validated and published search hedges that focus on transgender and gender diverse (TGD) communities and topics [8,9]. In this protocol, we use the term *TGD* to describe people who do not identify with their assigned gender at birth. This definition has the potential to also describe intersex people and people with differences of sex development (DSDs), and there are indeed certain overlapping experiences between TGD and intersex people. However, we follow the recommendations of organizations such as interACT: Advocates for Intersex Youth who make plain that being intersex and being TGD are distinct categories of identity and experience. Thus, we will not include articles studying DSDs or intersex communities unless it is clear within the article that the study sample includes people who are both intersex and transgender to respect these identities and experiences as distinct from those of TGD communities.

Additionally, we recognize that *TGD* can be understood as an anglophone export or a category imposed onto autochthonous sex and gender systems that have been devalued, erased, or even criminalized through imperialism and settler colonialism [10,11]. Thus, our search hedges incorporate terminology beyond that centered within the medical literature of the anglophone world and Global North. This decision to incorporate culturally specific terms, while intended to be inclusive, also runs the risk of reproducing the centrality of not only English-language terms but also medicalized models of understanding gender nonconformity and gender diversity.

TGD people face systematic marginalization and oppression unique from those of the cisgender heterosexual population and their sexual minority peers [12,13]. Misconceptions and disinformation regarding the experiences, quality of life, and health of TGD people further fuel the discrimination against this population [14-16]. Rigorous and transparent knowledge synthesis processes, starting with high-quality search hedges, can help inform and equip community members, clinicians, policymakers, and other key decision-makers with scientifically sound evidence.

### Validation Structure

#### Overview

Developing structured search strategies focused on TGD populations is complex for several reasons, and our nonstandard screening process was developed to address these complexities. In addition to the *Include* and *Exclude* categories in a standard literature screening process, 2 additional categories were created for the title and abstract screening phase: *Mixed\_Topics* and *LGB\_Maybe\_T*. The purpose of the *Mixed\_Topics* category is to differentiate articles related to TGD people and topics from those related to cisgender people and topics. The purpose of the *LGB\_Maybe\_T* category is to differentiate articles on gender identity topics relevant to TGD people from those on Two-Spirit, lesbian, gay, bisexual, transgender, queer, intersex, asexual, or additional sexual and gender diverse (2SLGBTQIA+) topics.

#### Mixed\_Topics Category

One challenge in identifying and retrieving literature relevant to TGD populations is the variations in context and usage of polysemic terms such as *gender diversity* or *gender inclusivity* (for a complete list of polysemic terms relevant to these hedges, see [Multimedia Appendix 1](#)). For example, the term *gender inclusivity* can be used to describe the inclusion of TGD people in historically underrepresented fields, but it can also be used to refer to the inclusion of cisgender women in fields and occupations historically dominated by men. Due to this usage in both cisgender and TGD contexts, inclusion of these terms in a search hedge has the potential to increase the number of nonrelevant references being retrieved, imposing additional burden on the subsequent literature screening process. However, exclusion of these terms risks compromising the comprehensiveness of the search hedges. Thus, the category *Mixed\_Topics* was created to evaluate the accuracy and usefulness of including these polysemic terms.

#### LGB\_Maybe\_T Category

An additional challenge with searching for TGD literature is the conflation of gender identity and sexual orientation: authors often report gender identity inconsistently, if at all [17-19]. This lack of specificity means that gender minority populations are often homogenized within the broader 2SLGBTQIA+ acronym without differentiating the experiences of gender minority

groups as distinct from those of sexual minority groups. While sexual and gender minority groups share close social and political struggles, and TGD people can also have a sexual minority identity (eg, a nonbinary bisexual individual), people with gender minority identities often experience unique forms of discrimination stemming from cisnormativity [20]. Therefore, articles that are relevant to 2SLGBTQIA+ populations are not guaranteed to be relevant to TGD populations, and articles that only use the term 2SLGBTQIA+ (or its variations) in their titles and abstracts do not provide sufficient information to determine their relevancy to TGD people and topics. The category *LGB\_Maybe\_T* was created to (1) confirm the relevance of articles using the acronym 2SLGBTQIA+ to the TGD population and (2) analyze patterns of usage of TGD terms in 2SLGBTQIA+ research.

## Precision and Sensitivity

Precision and sensitivity are 2 measurements commonly used in the validation process of search hedges and together offer quantifiable reference points in the performance of a hedge. Precision is the ability of a search hedge to minimize the retrieval of nonrelevant articles [21,22]. Precision is calculated as the proportion of relevant over total retrieved articles [22]. Sensitivity, also called relative recall, is the ability of a search hedge to retrieve all relevant articles [21,22]. Sensitivity is calculated as the proportion of articles from a gold-standard set that are retrieved by a search hedge [22]. The most common method of establishing a gold-standard set for the validation of a search hedge is through manually searching a journal, a set of journals, or a database [22]. However, the main disadvantage of this method is its high time consumption and low article yield [22]. Another approach to finding a gold standard is using a previously assembled collection [22]. In this study, the gold-standard set consists of 144 TGD articles from Knowsy (search year: 2019; for a full list of articles in the gold-standard set, see [Multimedia Appendix 2](#)). Knowsy is an online dataset of scoping reviews, systematic reviews, and structured reviews focused on 2SLGBTQIA+ topics [23,24]. Articles in Knowsy were identified by systematically searching 4 academic databases (MEDLINE, APA PsycInfo, Embase, and CINAHL) and selected in accordance with the standard 2-step screening process performed by 2 independent, trained reviewers [23,24]. Each reference in Knowsy has been labeled with standardized category tags for topics, populations, and type of knowledge synthesis by a trained reviewer, with independent validation by a second reviewer [23,24].

Special consideration also needs to be paid to ensure ethical and responsible research when working with a marginalized community, such as gender minority communities. Accountability and transparency in the search and research process are 2 key components to ensure that researchers do not further alienate already marginalized people. One of the best practices to accomplish these goals is the publication of a protocol at the start of the research process [25]. A protocol offers detailed background information on the search and screening process and increases the reproducibility of the results. It also provides an accountability mechanism as deviations from the protocol need to be identified when reporting study findings. As validations of search hedges are still uncommon, protocols

for the validation of these hedges are even more limited [26-28]. The scant literature regarding search hedge validation methods could further discourage this type of research from being conducted as lack of established methods can increase the amount of time and resources dedicated to this process. Thus, publishing a search hedge validation protocol that includes details about the methods used aids in knowledge and resource sharing, as well as facilitating the peer review process and increasing research rigor. In this paper, we present a protocol for the validation of 2 TGD search hedges using a novel validation structure.

## Objectives

This protocol describes the methods for a search hedge validation. The objectives are to (1) validate 2 interdisciplinary and cross-cultural TGD search hedges designed for retrieving references in MEDLINE and APA PsycInfo, (2) determine the usage pattern of TGD polysemic terms in the literature, and (3) assess the usage pattern of TGD terms in 2SLGBTQIA+ literature.

## Methods

### Ethical Considerations

Ethics approval will not be sought for this study because it will only include analysis of published articles. This study will not involve human participants, medical records, patient information, observations of public behaviors, or secondary data analyses.

### Search Hedge Design

To develop and validate the 2 search hedges, we followed a 2-stage process: a preliminary search and a finalized search.

#### Stage 1: Preliminary Search Hedges

Two preliminary search hedges were developed in 2023 by 2 librarians with the goal of obtaining new insights into the effectiveness of the searches and inform sample size calculations for the full validation study. The MEDLINE search produced 49,796 results, and the APA PsycInfo search produced 38,344 references. References were imported into EPPI-Reviewer (EPPI-Centre) and randomly allocated into screening groups [29]. Two team members then screened 0.5% (239/49,796) of the references from MEDLINE and 0.5% (192/38,344) of the references from APA PsycInfo.

In terms of precision and sensitivity, the preliminary search hedges identified 37.7% (90/239) of articles from MEDLINE and 37.5% (72/192) of articles from APA PsycInfo that met the eligibility criteria, achieving a precision of 39% for both databases. The hedges retrieved all articles from the gold-standard set, achieving a 100% sensitivity rate. Additionally, 5.7% (11/192) of the records from APA PsycInfo were not available in full text.

#### Stage 2: Finalization of the Search Hedges

On the basis of the precision and sensitivity of the preliminary search hedges, the team refined the search strategies to reduce the number of nonrelevant results. On June 7, 2024, we conducted 2 searches using the finalized search hedges, which

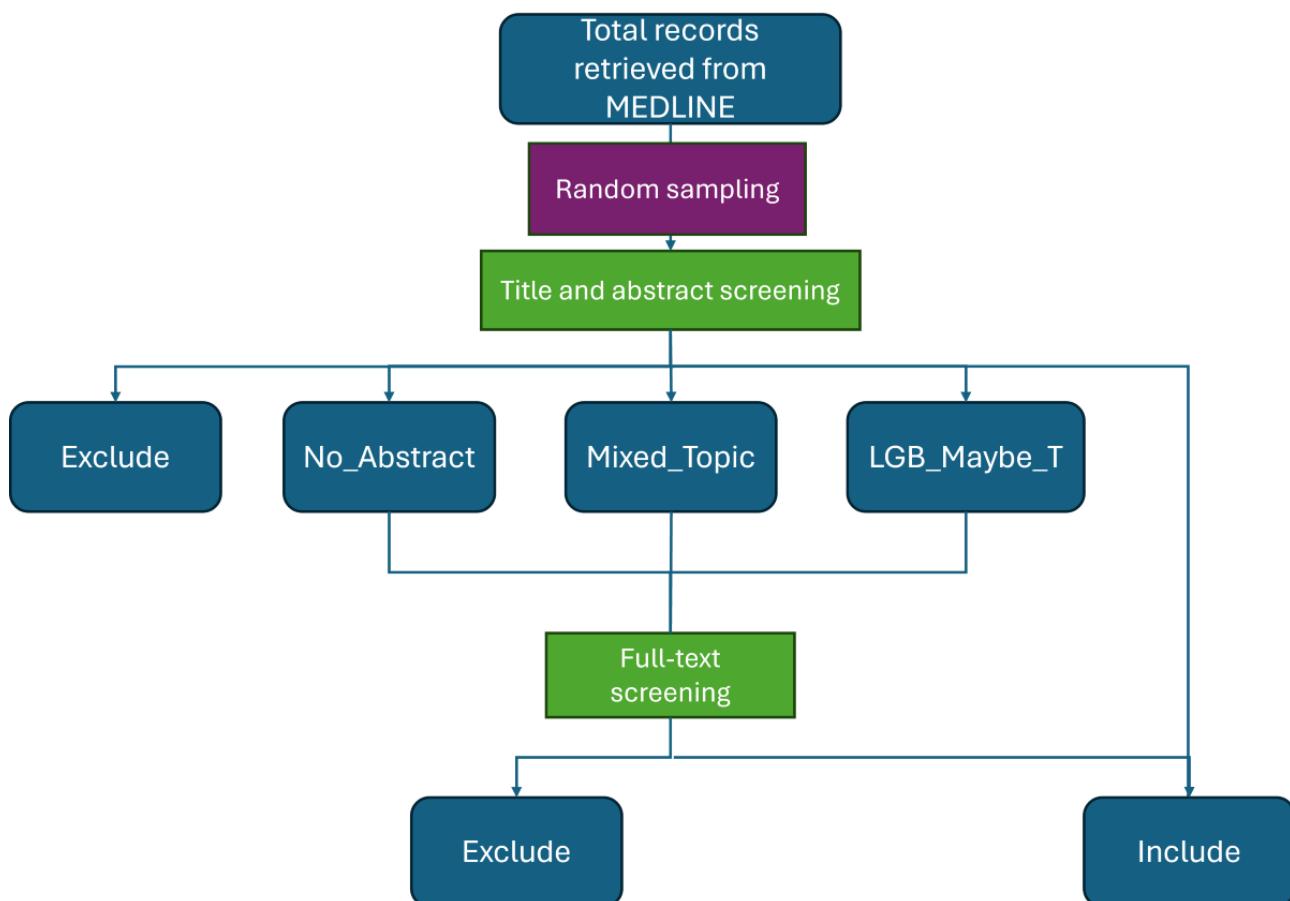
produced 31,055 references from MEDLINE and 22,924 references from APA PsycInfo ([Multimedia Appendix 3](#)).

### Sample Size Calculations

Informed by the preliminary results of our pilot hedges, we assume that a similar proportion of sampled records will meet the inclusion criteria (approximately 39%) and expect approximately 10% of records to be unretrievable (although our pilot search in APA PsycInfo only showed 6% of records to be unretrievable). For the MEDLINE hedge, a sample size of 2330 will achieve a targeted sensitivity of 100% (99% CI 99.4%-100%). For the APA PsycInfo hedge, a sample size of 2293 will achieve a targeted sensitivity of 100% (99% CI 99.3%-100%).

The calculation of precision for each search hedge requires the specificity rate. Specificity is defined as the proportion of nonrelevant articles correctly not retrieved by a search hedge [21]. In the context of this project, to calculate specificity, all articles indexed in MEDLINE and APA PsycInfo would need to be screened. This is not a feasible option. Thus, specificity will not be calculated. However, we opted for a targeted specificity of 80% (99% CI 77.2%-83%) for MEDLINE and a targeted specificity of 80% (99% CI 77.1%-83%) for APA PsycInfo. These targeted sensitivity and specificity levels correspond to a precision of 76% for both hedges.

**Figure 1.** Flow chart depicting the screening process for references relevant to trans and gender diverse (TGD) subject topics from MEDLINE. A similar process was used for results retrieved from APA PsycInfo. Only references categorized as *No\_Abstract*, *Mixed\_Topics*, or *LGB\_Maybe\_T* will proceed to the full-text screening phase to validate their relevance to TGD subject topics.



### Exclude Category

References will be excluded if they (1) use solely binary terminology to describe gender (eg, boys and girls, men and women, *both genders*, or *opposite gender*) or (2) focus on psychometric measurement of gender without reporting participants' self-identification or narrative description (eg, Bem Sex-Role Inventory, including sex typing using feminine, masculine, and androgynous categories) [30]. Abstracts focusing on intersex or DSD topics will also be excluded.

### Include Category

References will be included if they (1) report on TGD people, or both, in their sample; or (2) specifically discuss TGD communities or TGD topics, such as gender-affirming health care, gender expression, or drag performance. References containing historical terminology referring to Two-Spirit or TGD people, as well as Indigenous and non-Eurocentric gender-expansive terminology, will be included.

### No\_Abstract Category

References without abstracts will be categorized as *No\_Abstract* unless their titles clearly indicate relevancy to TGD populations (eg, using the term *transgender* in the title), in which case they will be categorized as *Include*.

### LGB\_Maybe\_T or Mixed\_Topics Category

For abstracts with unclear subject matter relevant to TGD topic areas, 2 coding options will be available. Abstracts with an unclear TGD or 2SLGBTQIA+ population or TGD or 2SLGBTQIA+-related study topic will be categorized as *Mixed\_Topics* (eg, using the term *gender inclusivity* but with no further clarification of whether this means inclusion of TGD people or only cisgender women). If abstracts use the term *transgender* or *trans* solely associated with or as part of elaborating on sexual and gender identity acronyms (eg, *LGBTQ+* or *2SLGBTQIA+*), these references will be categorized as *LGB\_Maybe\_T*. Abstracts in the *Mixed\_Topics* category will be presumed to relate to the cisgender population with unclear relevance to TGD people and topics until further confirmation via the full text. Abstracts in the *LGB\_Maybe\_T* category will be presumed to relate to the 2SLGBTQIA+ population with unclear relevance to TGD people and topics until further confirmation via the full text.

### Full-Text Eligibility Criteria

During full-text screening, articles will be categorized as *Include* or *Exclude*. Similar to the screening based on title and abstract phase, the objective is to accurately identify whether the article is relevant to TGD subject matter.

Articles that are included will (1) clearly distinguish among sex, sexual identity, and gender identity; (2) communicate consideration for participants' gender self-identification and experiences (eg, self-declaration options in a demographic survey); or (3) consider the TGD population as a distinct subpopulation within the 2SLGBTQIA+ umbrella (eg, report on transgender or gender diverse participants separately in a demographic section or as part of the data analysis strategy). For primary research papers, these considerations should be reflected in the Methods or Results sections.

### Information Sources

No restrictions will be placed on the type of reference, publication date, or language. All efforts will be made to translate articles that are not in English, French, Spanish, or Mandarin using Google Translate or DeepL Translator (DeepL SE), an online translation software [31]. Additional professional translation services will also be sought for articles in languages that are not stated above.

### Study Records

#### Data Management

Results from the MEDLINE and APA PsycInfo searches will be imported to the EPPI-Reviewer online systematic review software [29]. Deduplication will not be conducted as validation and data analysis will be performed separately for each database. Random samples will be allocated using the *Create reference groups* feature in EPPI-Reviewer. Full-text references will be manually or automatically retrieved via Zotero (Corporation for Digital Scholarship), a citation manager, and uploaded to EPPI-Reviewer. EPPI-Reviewer will also be used to track and manage full text retrieval.

#### Selection Process

The review team includes a total of 6 members with lived experience in the TGD community. To ensure intercoder agreement, team member training will consist of a prescreening walkthrough of the eligibility criteria followed by trial coding until the disagreement rate between all coders is equal to or below the recommended 20% [32]. An existing internally circulated team notebook will be created at the start of the screening process to record notable discussions and decisions throughout the process. This notebook will be available to all team members to ensure consistency between coders. Two team members will independently screen each randomly sampled title, abstract, and full text based on the predefined eligibility criteria. Any disagreements will be addressed through discussion facilitated by a third team member, with at least one of the original screening coders present to ensure clarity and context. If differences cannot be resolved through discussion, the project team leads will make the final decision.

### Data Analysis

#### Intercoder Agreement Calculation

The intercoder agreement in each screening phase will be calculated using the following formula [32]:

$$\text{Percent agreement} = \frac{\text{Number of concordant codes assigned}}{\text{Total number of codes assigned}} \times 100\%$$

#### Data Items and Planned Statistical Tests

Precision and sensitivity will be calculated to evaluate the validity of each search hedge. Precision quantifies the ability of a search hedge to identify relevant TGD articles. This measurement will be calculated as the number of relevant articles retrieved divided by the total number of screened articles [21]:

$$Precision = \frac{\text{Number of relevant articles}}{\text{Total number of screened articles}} \times 100\%$$

Sensitivity assesses the ability of a search hedge to comprehensively retrieve all relevant articles [21]. Sensitivity will be calculated as the proportion of articles successfully retrieved by the search hedge compared to the gold-standard set as predefined from Knowsy:

$$Sensitivity = \frac{\text{Number of gold standard articles retrieved by the hedge}}{\text{Number of articles in the gold standard set}} \times 100\%$$

### Search Term Insights

Polysemic terms identified from the finalized search with unclear relevancy to TGD topics will be assessed. A verbatim search of each term will be conducted in EPPI-Reviewer, and accuracy will be calculated as a proportion of the articles in each category (*Mixed\_Topics*, *LGB\_Maybe\_T*, *Exclude*, and *Include* from both the title and abstract and full-text screening) over the total number of articles containing said term. A higher proportion from the *Mixed\_Topics* or *LGB\_Maybe\_T* category suggests that a term is not being used in a TGD-specific context. A higher proportion from the *Exclude* category suggests that a term does not have high relevance to TGD topics:

$$Search term insight = \frac{\text{Number of articles in each category}}{\text{Total number of articles containing term}}$$

## Results

As of November 2025, we have finished screening all 2330 articles from MEDLINE and all 2293 articles from APA PsycInfo. Data analysis is projected to be finished by December 2025. The results are expected to be published in 2026.

## Discussion

### Expected Findings

While a number of validated and unvalidated search hedges focused on 2SLGBTQIA+ populations exist, to the best of our knowledge, this protocol is the first to detail the process of evaluating 2 search hedges that exclusively focus on the TGD population for MEDLINE and APA PsycInfo [8,9,33,34]. In this paper, we developed a novel screening structure for terminologies with dynamic meanings, with full-text screening only for references where further confirmation was required. We anticipate that both the MEDLINE and APA PsycInfo search hedges will achieve a similar sensitivity of 100% to that of their pilot hedges. We also anticipate that the hedges will achieve higher precision than their pilot counterparts due to the fine-tuning conducted in between. The precision, sensitivity, and search term insight results will provide guidance for librarians and researchers to customize the hedges to best fit their research questions.

While various published and validated search filters exist, most of these hedges focus on retrieving articles based on geographic locations, type of study, and treatment [5,26,27,35]. Validated search hedges focusing on populations, specifically marginalized communities, are particularly scant. Schilperoort et al [34]

documented the development and validation process of search filters for PubMed focusing on queer women. Wafford et al [36] documented the development and validation process of a search filter for PubMed focusing on immigrant populations. Both of these population filters relied on the relative recall method to create their respective gold-standard sets. Relative recall is the process of retrieving structured reviews, such as systematic reviews, scoping reviews, and evidence gap maps, related to the specified population and searching through the citations of each review to further identify relevant articles. Schilperoort et al [34] used an internal validation approach in which the gold-standard set was used to generate the search terms as well as the standard for testing the sensitivity of the filters. On the other hand, Wafford et al [36] used an external validation approach in which the team followed the same process of relative recall but partitioned the retrieved articles into 2 separate sets: a development set from which to derive search terms and a validation set (occupying the same role of a gold standard) to test the filter. External validation offers more rigorous results as it mimics a testing environment closer to how a filter will be deployed in a real-world setting [35]. In this protocol, we are able to use an external validation approach due to the existence of the TGD subset in Knowsy. The Knowsy dataset is entirely independent in development and screening to the 2 current search hedges, allowing for a more stringent validation process.

Our validation method further diverged from the methods used by Schilperoort et al [34] and Wafford et al [36]. Both groups validated the precision of their filters against known sets established through the relative recall process, whereas the existence of the Knowsy dataset as our gold standard allows us to test and calculate the precisions and sensitivities of our hedges against unknown sets, offering a closer resemblance to real-world settings [34,36]. While Wafford et al [36] also used external validation, the team only accessed articles at the title and abstract level. For a large portion of our retrieved articles, the titles and abstracts do not provide sufficient information regarding our population of interest, a similar problem that Schilperoort et al [34] also encountered with a search filter focused on queer women. Thus, without the need to generate a gold-standard or validation set, our screening process can explore the semantics and context of terms through our *Mixed\_Topics* and *LGB\_Maybe\_T* categories while minimizing the number of articles needing to be screened in full text.

### Knowledge Mobilization Plan

Once validated, the search hedges will be submitted to the InterTASC Information Specialists' Sub-Group Search Filter Resource, the McMaster Health Knowledge Refinery, and the University of Alberta Library's health sciences search filter database [37-39]. We plan to periodically update the 2 hedges to keep abreast with indexing changes in their respective databases. Additionally, future plans will include development and adaptation of the hedges for other bibliographic databases, such as Embase, PubMed, CINAHL, and Scopus. Other knowledge mobilization plans include presentations at relevant research and knowledge synthesis conferences such as Moving Trans History Forward and the Global Evidence Summit [40,41].

## Limitations and Considerations

One primary limitation of the hedges is the limited semantic search. This is related to the standard search mechanisms in MEDLINE, APA PsycInfo, and other databases. The default search mechanism for these databases is the keyword search, which relies on exact match or known variants of the search strings [42]. For example, if a search string specifies the term *transgender*, the keyword search will retrieve only items with the exact match. Truncation operators can increase recall by looking for variants based on the word root (eg, plurals) but not semantic meaning. However, a keyword search does not take into account the context in which the term occurred, thus retrieving results with false semantic matches. In this example, articles with phrases such as *not transgender* would still be retrieved by the hedges. We anticipate that this search mechanism will decrease the precision rates of the hedges.

Another consideration when using the search hedges is the trade-off between sensitivity and precision. The primary user

audience for the search hedges are information professionals supporting evidence syntheses, and thus, the hedges were designed with sensitivity as the top priority. The trade-off of this prioritization is precision, and certain types of reviews with time and resource constraints, such as rapid reviews, should keep this in mind. By examining specific search terms that are anticipated to retrieve more nonrelevant articles, we hope to provide librarians and other readers with additional information to make guided adaptations of the hedges that best fit individual research questions and purposes.

## Conclusions

This protocol provides a blueprint for future studies to respectfully and comprehensively search literature related to self-described identities used by TGD communities. This protocol, along with other published articles of its kind, will contribute to the openness of science and the practice of sharing resources within the research community.

---

## Acknowledgments

The authors would like to extend their heartfelt gratitude to Arthur Dave Miller (Sexual Health and Gender (SHaG) lab, Dalhousie University), Meng Wang (University of Calgary), and Atlas Davis (University of Calgary) for providing their valuable insights in developing the search hedges. This project received no direct grant from any funding agency in the public, commercial, or not-for-profit sectors.

---

## Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

---

## Authors' Contributions

Conceptualization: AE, CD, RG, SM, ZM

Data curation: RG, SM

Formal analysis: AE, CD

Investigation: AE, CD, RG, SM, ZM

Methodology: AE, CD, PK, RG, SM, ZM

Project administration: CD

Supervision: AE, CD, ZM

Validation: AE, CD

Visualization: CD

Writing—original draft: CD, ZM

Writing—review and editing: AE, CD, PK, RG, SM, ZM

---

## Conflicts of Interest

The gold-standard dataset comes from Knowsy, a free online portal of 2SLGBTQIA+ focused systematic reviews, scoping studies, and evidence maps. The studies in the dataset are from the 2019 iteration of Knowsy and were identified from initial search strategies used to retrieve sexual and gender minority research. ZM cofounded Knowsy and now coleads this initiative with CD.

---

## Multimedia Appendix 1

Lists of polysemic terms identified during pilot review.

[\[DOCX File, 1915 KB-Multimedia Appendix 1\]](#)

---

## Multimedia Appendix 2

List of trans and gender diverse gold standard references from Knowsy (searched in 2019).

[\[DOCX File, 48 KB-Multimedia Appendix 2\]](#)

---

## Multimedia Appendix 3

Trans search hedges in Ovid MEDLINE and APA PsycInfo.

[[DOCX File](#) , [2041 KB-Multimedia Appendix 3](#)]

## References

1. Grimshaw J. A knowledge synthesis chapter. Canadian Institutes of Health Research. URL: [https://www.cihr.gc.ca/e/documents/knowledge\\_synthesis\\_chapter\\_e.pdf](https://www.cihr.gc.ca/e/documents/knowledge_synthesis_chapter_e.pdf) [accessed 2025-03-04]
2. Rethlefsen ML, Farrell AM, Osterhaus Trzasko LC, Brigham TJ. Librarian co-authors correlated with higher quality reported search strategies in general internal medicine systematic reviews. *J Clin Epidemiol*. Jun 2015;68(6):617-626. [doi: [10.1016/j.jclinepi.2014.11.025](https://doi.org/10.1016/j.jclinepi.2014.11.025)] [Medline: [25766056](#)]
3. Campbell S. What is the difference between a filter and a hedge? *J Eur Assoc Health Inf Libr*. Apr 01, 2016;12(1):4-5. [[FREE Full text](#)] [doi: [10.7939/R35T3GF4C](https://doi.org/10.7939/R35T3GF4C)]
4. Damarell RA, May N, Hammond S, Sladek RM, Tieman JJ. Topic search filters: a systematic scoping review. *Health Info Libr J*. Dec 21, 2018;36(1):4-40. [doi: [10.1111/hir.12244](https://doi.org/10.1111/hir.12244)] [Medline: [30578606](#)]
5. Ayiku L, Levay P, Hudson T, Craven J, Finnegan A, Adams R, et al. The Embase UK filter: validation of a geographic search filter to retrieve research about the UK from OVID Embase. *Health Info Libr J*. Jun 2019;36(2):121-133. [[FREE Full text](#)] [doi: [10.1111/hir.12252](https://doi.org/10.1111/hir.12252)] [Medline: [30912233](#)]
6. Damarell RA, Lewis S, Trenerry C, Tieman JJ. Integrated Care Search: development and validation of a PubMed search filter for retrieving the integrated care research evidence. *BMC Med Res Methodol*. Jan 21, 2020;20(1):12. [[FREE Full text](#)] [doi: [10.1186/s12874-020-0901-y](https://doi.org/10.1186/s12874-020-0901-y)] [Medline: [31964347](#)]
7. Ioerger M, Flanders RM, Goss KD, Turk MA. Developing a systematic search strategy related to people with disability: a brief report testing the utility of proposed disability search terms in a search about opioid use. *Disabil Health J*. Apr 2019;12(2):318-322. [doi: [10.1016/j.dhjo.2018.11.009](https://doi.org/10.1016/j.dhjo.2018.11.009)] [Medline: [30470478](#)]
8. Kennedy M. Filter to retrieve studies related to LGBTQIA2S+ from the Ovid APA PsycInfo database. University of Alberta. Apr 4, 2025. URL: [https://docs.google.com/document/d/1ZNt2GkBy1SEpgDnxH0HCqBts-XStVr90VTsAk6zpX8M/edit?usp=sharing&usp=embed\\_facebook](https://docs.google.com/document/d/1ZNt2GkBy1SEpgDnxH0HCqBts-XStVr90VTsAk6zpX8M/edit?usp=sharing&usp=embed_facebook) [accessed 2025-09-04]
9. Kennedy M. Filter to retrieve studies related to LGBTQIA2S+ from the Ovid MEDLINE database. University of Alberta. Apr 4, 2025. URL: [https://docs.google.com/document/d/1ZNt2GkBy1SEpgDnxH0HCqBts-XStVr90VTsAk6zpX8M/edit?tab=t\\_0#heading=h.qi55eeyvgzy9](https://docs.google.com/document/d/1ZNt2GkBy1SEpgDnxH0HCqBts-XStVr90VTsAk6zpX8M/edit?tab=t_0#heading=h.qi55eeyvgzy9) [accessed 2025-11-28]
10. Boellstorff T, Cabral M, Cárdenas M, Cotten T, Stanley EA, Young K, et al. Decolonizing transgender: a roundtable discussion. *Transgender Stud Q*. Aug 01, 2014;1(3):419-439. [doi: [10.1215/23289252-2685669](https://doi.org/10.1215/23289252-2685669)]
11. Everhart AR. Bones without flesh and (trans)gender without bodies: querying desires for trans historicity. *Hypatia*. Feb 23, 2023;37(4):601-618. [[FREE Full text](#)] [doi: [10.1017/hyp.2022.50](https://doi.org/10.1017/hyp.2022.50)]
12. Deal C, Doshi RD, Gonzales G. Gender minority youth experiencing homelessness and corresponding health disparities. *J Adolesc Health*. May 2023;72(5):763-769. [doi: [10.1016/j.jadohealth.2022.11.229](https://doi.org/10.1016/j.jadohealth.2022.11.229)] [Medline: [36646565](#)]
13. Delozier AM, Kamody RC, Rodgers S, Chen D. Health disparities in transgender and gender expansive adolescents: a topical review from a minority stress framework. *J Pediatr Psychol*. Sep 01, 2020;45(8):842-847. [doi: [10.1093/jpepsy/jsaa040](https://doi.org/10.1093/jpepsy/jsaa040)] [Medline: [32626901](#)]
14. Billard TJ. The politics of transgender health misinformation. *Political Commun*. Jan 08, 2024;41(2):344-352. [doi: [10.1080/10584609.2024.2303148](https://doi.org/10.1080/10584609.2024.2303148)]
15. Coldwater KD. Decoding the Misinformation-Legislation Pipeline: an analysis of Florida Medicaid and the current state of transgender healthcare. *J Med Libr Assoc*. Oct 02, 2023;111(4):750-761. [[FREE Full text](#)] [doi: [10.5195/jmla.2023.1724](https://doi.org/10.5195/jmla.2023.1724)] [Medline: [37928129](#)]
16. McNamara M, McLamore Q, Meade N, Olgun M, Robinson H, Alstott A. A thematic analysis of disinformation in gender-affirming healthcare bans in the United States. *Soc Sci Med*. Jun 2024;351:116943. [doi: [10.1016/j.socscimed.2024.116943](https://doi.org/10.1016/j.socscimed.2024.116943)] [Medline: [38759383](#)]
17. Connolly MD, Zervos MJ, Barone CJ, Johnson CC, Joseph CL. The mental health of transgender youth: advances in understanding. *J Adolesc Health*. Nov 2016;59(5):489-495. [doi: [10.1016/j.jadohealth.2016.06.012](https://doi.org/10.1016/j.jadohealth.2016.06.012)] [Medline: [27544457](#)]
18. Edmiston EK, Donald CA, Sattler AR, Peebles JK, Ehrenfeld JM, Eckstrand KL. Opportunities and gaps in primary care preventative health services for transgender patients: a systemic review. *Transgend Health*. Nov 01, 2016;1(1):216-230. [[FREE Full text](#)] [doi: [10.1089/trgh.2016.0019](https://doi.org/10.1089/trgh.2016.0019)] [Medline: [28861536](#)]
19. Suen LW, Lunn MR, Katuzny K, Finn S, Duncan L, Sevelius J, et al. What sexual and gender minority people want researchers to know about sexual orientation and gender identity questions: a qualitative study. *Arch Sex Behav*. Oct 2020;49(7):2301-2318. [[FREE Full text](#)] [doi: [10.1007/s10508-020-01810-y](https://doi.org/10.1007/s10508-020-01810-y)] [Medline: [32875381](#)]
20. Berger I, Ansara YG. Cisnormativity. In: Goldberg AE, Beemyn G, editors. The SAGE Encyclopedia of Trans Studies. Thousand Oaks, CA. SAGE Publications; 2021.
21. Baratloo A, Hosseini M, Negida A, El Ashal G. Part 1: simple definition and calculation of accuracy, sensitivity and specificity. *Emerg (Tehran)*. 2015;3(2):48-49. [[FREE Full text](#)] [Medline: [26495380](#)]

22. Jenkins M. Evaluation of methodological search filters--a review. *Health Info Libr J*. Sep 2004;21(3):148-163. [FREE Full text] [doi: [10.1111/j.1471-1842.2004.00511.x](https://doi.org/10.1111/j.1471-1842.2004.00511.x)] [Medline: [15318913](https://pubmed.ncbi.nlm.nih.gov/15318913/)]

23. Farley N, Marshall Z. Creating a database for LGBTQ2SI+ knowledge synthesis (Knowsy): a mapping review of evidence and gaps to inform future research. 2020. Presented at: What Works Global Summit; October 30, 2020; Online. URL: <https://web.archive.org/web/20210515083935/https://www.wwgs2020.org/30-october.html>

24. Home. Knowsy. URL: <https://knowsy.ca/> [accessed 2025-03-17]

25. Otake PJ, Childs JD. Why publish study protocols? *Phys Ther*. Sep 01, 2014;94(9):1208-1209. [FREE Full text] [doi: [10.2522/ptj.2014.94.9.1208](https://doi.org/10.2522/ptj.2014.94.9.1208)]

26. Muente C, Pachanov A, Hirt J, Ayiku L, Pieper D. Validated geographic search filters for bibliographic databases: a scoping review protocol. *JBI Evid Synth*. Mar 01, 2024;22(3):441-446. [doi: [10.11124/JBIES-23-00445](https://doi.org/10.11124/JBIES-23-00445)] [Medline: [38344846](https://pubmed.ncbi.nlm.nih.gov/38344846/)]

27. Wilczynski NL, Morgan D, Haynes RB, Hedges Team. An overview of the design and methods for retrieving high-quality studies for clinical care. *BMC Med Inform Decis Mak*. Jun 21, 2005;5:20. [FREE Full text] [doi: [10.1186/1472-6947-5-20](https://doi.org/10.1186/1472-6947-5-20)] [Medline: [15969765](https://pubmed.ncbi.nlm.nih.gov/15969765/)]

28. Lunny CA, Salzwedel DM, Liu T, Ramasubbu C, Gerrish S, Puil L, et al. Protocol for the validation of four search strategies for retrieval of clinical practice guidelines in MEDLINE, Embase and PubMed. *PeerJ Preprints*. Preprint posted online on August 28, 2018. [doi: [10.7287/peerj.preprints.27149v1](https://doi.org/10.7287/peerj.preprints.27149v1)]

29. Thomas J, Graziosi S, Brunton J, Ghouze Z, O'Driscoll P, Bond M, et al. EPPI-Reviewer: advanced software for systematic reviews, maps and evidence synthesis. EPPI Centre. 2023. URL: <https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=2967> [accessed 2025-11-28]

30. Bem SL. Bem Sex Role Inventory. American Psychological Association. Sep 12, 2011. URL: <https://psycnet.apa.org/doiLanding?doi=10.1037%2Ft00748-000> [accessed 2025-11-28]

31. DeepL translate: the worlds most accurate translator. DeepL. URL: <https://www.deepl.com/translator> [accessed 2025-03-04]

32. Gisev N, Bell JS, Chen TF. Interrater agreement and interrater reliability: key concepts, approaches, and applications. *Res Social Adm Pharm*. 2013;9(3):330-338. [doi: [10.1016/j.sapharm.2012.04.004](https://doi.org/10.1016/j.sapharm.2012.04.004)] [Medline: [22695215](https://pubmed.ncbi.nlm.nih.gov/22695215/)]

33. Parker R, Foster M, Wanner A, Lackey M. Design and validation of a search filter for LGBTQ+ populations. *J Can Health Libr Assoc*. Jul 24, 2018;39(2):94. [FREE Full text]

34. Schilperoort HM, Hickner A, Morgan-Daniel J, Parker RM. Development and validation of LGBTQIA+ search filters: report on process and pilot filter for queer women. *J Med Libr Assoc*. Apr 18, 2025;113(2):123-132. [doi: [10.5195/jmla.2025.2002](https://doi.org/10.5195/jmla.2025.2002)] [Medline: [40342307](https://pubmed.ncbi.nlm.nih.gov/40342307/)]

35. Lefebvre C, Glanville J, Beale S, Boachie C, Duffy S, Fraser C, et al. Assessing the performance of methodological search filters to improve the efficiency of evidence information retrieval: five literature reviews and a qualitative study. *Health Technol Assess*. Nov 2017;21(69):1-148. [FREE Full text] [doi: [10.3310/hta21690](https://doi.org/10.3310/hta21690)] [Medline: [29188764](https://pubmed.ncbi.nlm.nih.gov/29188764/)]

36. Wafford QE, Miller CH, Wescott A, Kubilius RK. Meeting a need: development and validation of PubMed search filters for immigrant populations. *J Med Libr Assoc*. Jan 16, 2024;112(1):22-32. [FREE Full text] [doi: [10.5195/jmla.2024.1716](https://doi.org/10.5195/jmla.2024.1716)] [Medline: [38911528](https://pubmed.ncbi.nlm.nih.gov/38911528/)]

37. Health sciences search filters. University of Alberta. 2025. URL: <https://guides.library.ualberta.ca/health-sciences-search-filters> [accessed 2025-11-28]

38. Hedges project. McMaster University. URL: <https://hiruweb.mcmaster.ca/hkr/hedges/> [accessed 2025-09-18]

39. The ISSG search filters resource. The InterTASC Information Specialists' Sub-Group. URL: <https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home> [accessed 2025-09-18]

40. Moving trans history forward. University of Victoria. 2025. URL: <https://www.uvic.ca/mthf2025/index.php> [accessed 2025-09-18]

41. Leaders unite at Global Evidence Summit. Joanna Briggs Institute. URL: <https://jbi.global/news/article/global-evidence-summit> [accessed 2025-09-18]

42. Malve A, Chawan PM. A comparative study of keyword and semantic based search engine. *Int J Innov Res Sci Eng Technol*. Nov 2015;4(11):11156-11161. [FREE Full text] [doi: [10.15680/IJIRSET.2015.0411039](https://doi.org/10.15680/IJIRSET.2015.0411039)]

## Abbreviations

**DSD:** difference of sex development

**2SLGBTQIA+:** Two-Spirit, lesbian, gay, bisexual, transgender, queer, intersex, asexual, and additional sexual and gender diverse terminologies

**TGD:** transgender and gender diverse

Edited by J Sarvestan; submitted 20.Apr.2025; peer-reviewed by N Smallheiser; comments to author 30.Jul.2025; accepted 10.Nov.2025; published 23.Jan.2026

*Please cite as:*

Dinh C, Marshall Z, Marsalis S, Gagen R, Kang P, Everhart A

*Retrieving Transgender and Gender Diverse Literature: Protocol for the Development and Validation of 2 Search Hedges*

*JMIR Res Protoc* 2026;15:e76055

*URL:* <https://www.researchprotocols.org/2026/1/e76055>

*doi:* [10.2196/76055](https://doi.org/10.2196/76055)

*PMID:* [41575780](https://pubmed.ncbi.nlm.nih.gov/41575780/)

©Chi Dinh, Zack Marshall, Scott Marsalis, Ryn Gagen, Preet Kang, Avery Everhart. Originally published in JMIR Research Protocols (<https://www.researchprotocols.org>), 23.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.researchprotocols.org>, as well as this copyright and license information must be included.