

# Searching for sex- and gender-sensitive tuberculosis research in public health: finding a needle in a haystack

Bilkis Vissandjee<sup>1</sup>  
 Assia Mourid<sup>2</sup>  
 Christina A Greenaway<sup>3</sup>  
 Wendy E Short<sup>4</sup>  
 Jodi A Proctor<sup>5</sup>

<sup>1</sup>Faculty of Nursing, Public Health Research Institute, Université de Montréal, Montréal, Montréal, QC, Canada; <sup>2</sup>Allied Health Library, Université de Montréal, Montréal, QC, Canada; <sup>3</sup>Department of Medicine, McGill University, Division of Infectious Diseases, Jewish General Hospital, Montréal, QC, Canada; <sup>4</sup>Faculty of Humanities and Social Sciences, School of Social Sciences, University of Queensland, St Lucia, QLD, Australia; <sup>5</sup>School of Social Work, McGill University, Montréal, QC, Canada

**Abstract:** Despite broadening consideration of sex- and gender-based issues in health research, when seeking information on how sex and gender contribute to disease contexts for specific health or public health topics, a lack of consistent or systematic use of terminology in health literature means that it remains difficult to identify research with a sex or gender focus. These inconsistencies are driven, in part, by the complexity and terminological inflexibility of the indexing systems for gender- and sex-related terms in public health databases. Compounding the issue are authors' diverse vocabularies, and in some cases lack of accuracy in defining and using fundamental sex–gender terms in writing, and when establishing keyword lists and search criteria. Considering the specific case of the tuberculosis (TB) prevention and management literature, an analysis of sex and gender sensitivity in three health databases was performed. While there is an expanding literature exploring the roles of both sex and gender in the trajectory and lived experience of TB, we demonstrate the potential to miss relevant research when attempting to retrieve literature using only the search criteria currently available. We, therefore, argue that for good clinical practice to be achieved; there is a need for both public health researchers and users to be better educated in appropriate usage of the terminology associated with sex and gender. In addition, public health database indexers ought to accept the task of developing and implementing adequate definitions of sex and gender terms so as to facilitate access to sex- and gender-related research. These twin advances will allow clinicians to more readily recognize and access knowledge pertaining to systems of redress that respond to gendered risks that compound existing health inequalities in disease management and control, particularly when dealing with already complex diseases. Given the methodological and linguistic challenges presented by the multidimensional and highly contextual nature of definitions of sex and gender, it will be important that this review task be undertaken using a multidisciplinary approach.

**Keywords:** sex, gender, tuberculosis, literature search, indexing, databases, terminological accuracy, keywords search

## Introduction

It is well documented that different social notions of sex and gender mean that different individuals are impacted (positively or detrimentally) and experience varied vulnerabilities, symptoms, and responses to particular disease conditions and treatments.<sup>1–3</sup> Today, multinational groups such as the World Health Organization (WHO) advocate that good public health science accounts for social determinants of health such as sex and gender whenever possible.<sup>4</sup> Johnson et al<sup>5</sup> recommended that “all health fields and disciplines incorporate sex and gender as a matter of ethics and science”, that is, that researchers should always critically consider the influence of sex and gender on health

Correspondence: Bilkis Vissandjee  
 Faculty of Nursing, Université de Montréal, PO Box 6128, Station Centre-Ville, Montréal, QC H3C 3J7, Canada  
 Tel +1 514 343 6837  
 Email [bilkis.vissandjee@umontreal.ca](mailto:bilkis.vissandjee@umontreal.ca)

outcomes throughout research development, design, and reporting.<sup>6</sup> Bottorff et al<sup>7</sup> highlight the importance of critically considering sex and gender as socioeconomic determinants of health that contribute to broader health inequalities.

In the context of this general acknowledgment of the need to consider sex and gender as essential factors in the design and delivery of programs, policies, and practices that achieve better health outcomes, sex- and gender-based analysis in health research has become a much more common practice.<sup>8</sup> It is, therefore, expected that literature on sex and gender correlations should be represented both in the literature retrievable from health databases and in the controlled vocabulary systems used to index articles. However, while the evidence for the contribution of sex and gender to health outcomes and inequities is growing, accessing such data remains challenging. This paper identifies failures by authors in applying concepts as fundamental as the differentiation between sex and gender, as well as the potential for individual interpretation of terms which are specified for use within database search criteria, resulting in a mismatch between search results and user expectations.

The aim of this paper is to highlight the discrepancies by way of a use-efficacy analysis of the subject headings (terms of a controlled vocabulary or of a thesaurus called also descriptors) relevant to the concepts of sex and gender contained in the thesauri of three major health databases: Medical Literature Analysis and Retrieval System Online (MEDLINE), Excerpta Medica Database (EMBASE), and Cumulative Index to Nursing and Allied Health Literature (CINAHL). The paper first provides an overview of definitions of sex and gender, followed by a discussion of the importance of taking into consideration sex and gender as determinants in tuberculosis (TB) prevention and management. Using the example of searching for sex and gender impacts within the literature pertaining to TB prevention, management, and control, the efficacy of the three existing database search engines was analyzed. The steps adopted for the analysis of the three major health databases such as MEDLINE, EMBASE, and CINAHL, are presented.

## Defining sex and gender

Sex is a relatively well-defined term, widely understood as the biological and physiological characteristics that distinguish men, women, and intersex individuals. It is included in most epidemiological and clinical studies as a basic descriptor (male, female, and intersex) and a potential “confounding” variable. In contrast, gender is a more complex term and, as it encapsulates a continuum of sociocultural constructions,

remains variably and contextually defined.<sup>9–12</sup> WHO defines gender as encompassing: “features of males and females that are socially constructed, [and] distinct from those features that are biologically determined (sex linked).”<sup>13</sup> In common usage, gender and “sexual orientation” are often conflated, with gender incorrectly understood as a collective spectrum of identity expressions, including heterosexual, homosexual, bisexual, transgender, and intersex.<sup>14,15</sup> However, Hammarström et al<sup>15</sup> clarified that gender is not only distinct to biological sex but relates to a continuum of socially constructed norms filled by male, female, and intersex individuals, which is formed independent of an individual’s preferences with respect to the gender(s) of their sexual partner(s). The nuanced attributes and the changes over time in constructions of gender are multifaceted and constructed within specific cultures to encapsulate notions of gender appropriate identity, performance, roles and responsibilities, relations, sexuality, behaviors, and institutionalized ontologies.<sup>5</sup> Therefore, neither sex nor gender is a binary division. It is within this fluid context that it can be seen that sex and gender can become social determinants of vulnerability with significant health impacts: transgender women, for example, are found to have “a higher prevalence of drug use, HIV, drug use [sic], and sex work than the general population.”<sup>16</sup> As understandings of gender have developed, these component facets have been formalized into systematic and diverse fields of research.<sup>5</sup>

The risk of conflating the terms sex and gender, and not systematically accounting for the complexity underlying these (related) concepts and their genuine differences, is increasingly well documented.<sup>1,2,5,9</sup> Although the sex-disaggregated collection and presentation of health data are quite common, research endeavors identifying differences in experiences of disease do not always address gender as a causal factor in determining differential life trajectories, and one which can promote or inhibit access to and use of health or legal care. Studies show that selected social determinants of health, health-seeking behaviors, and health systems responses affect people of different genders in markedly different ways, sufficient to warrant developing methodological approaches beyond simply presenting data separately for women and for men.<sup>17</sup> The need to increase gender awareness, beyond the provision of simply sex-disaggregated data, is therefore essential in recognizing the causes and impacts of social gender inequalities and in developing tailored health promotion strategies.<sup>18,19</sup> Assisting researchers and clinicians to simultaneously engage contextual issues of social gender awareness, and utilize sex-disaggregated data, provides the imperative for clear and consistent concepts and definitions.<sup>5,18</sup>

## Why “tuberculosis” and sex and gender?

TB is known as one of the world’s deadliest communicable diseases and, in spite of the global decline in rates of infection, constitutes a major proportion of the global disease burden, affecting an estimated 9 million people worldwide in 2013.<sup>4</sup> TB is a health issue with significant sex- and gender-related disparities, which are not always widely recognized.<sup>20–24</sup> It is well known, for example, that TB is a preventable disease with a disproportionate impact on vulnerable populations. What is becoming increasingly apparent is that these vulnerable groups are often located in high disease burden, low- to middle-income countries where gender inequities constitute a major social determinant of access to adequate health care and systems of redress.<sup>25</sup> TB control, therefore, requires integrated, context-specific, sex and gender sensitive interventions, which require comprehensive access to sex- and gender-sensitive health research.<sup>26</sup>

WHO has identified the importance of gender considerations for TB control strategies and programs in multiple publications and resources, highlighting the way that sex and gender interact with other social determinants in limiting access to treatment.<sup>27</sup> As specified by the WHO DOTS (directly observed treatment, short-course) TB control strategy:

Locally appropriate measures should be undertaken to identify and address physical, financial, social and cultural – as well as health systems – barriers to accessing TB treatment services. Particular attention should be given to the poorest and most vulnerable population groups.<sup>27</sup>

Both sex- and gender-based differentials are documented in relation to TB transmission and control. While TB remains one of the top three global causes of death among women aged 15–44 years, the global case notification rate of TB is higher for males than for females.<sup>25</sup> On a global level, statistics indicate that men account for about two-thirds of all active TB notifications (a male:female infection ratio of 2:1); however, this sex-based disaggregation hides variability according to gender, age, and geographic region.<sup>28</sup> To date, there is no consensus as to why men generally have a higher infection rate for TB than women, although it is hypothesized that this may be due to a combination of social and biological factors. For example, social (gendered) factors may include men having exposure to a greater number of contacts during work and social activities outside the home, biological (sex) factors may include a greater susceptibility to TB infection due to male sex hormones.<sup>29–32</sup> This also highlights the need for gendered health studies to be recognized as seeking to

understand all gendered experiences and not to be misunderstood in a reductionist manner to be a study of issues encountered by women only.

Multiple components of gender affect TB control strategies, typically within a contextual web of gender roles and gender relations. For example, in some cultures, while gender roles may increase men’s exposure to TB infection, gender relations dictate that limited health resources are prioritized to care for male family members and women do not receive treatment or diagnosis. This may be further exacerbated where women lack financial autonomy, health literacy, or decision-making power, which prevents their access and use of health information and services for themselves.<sup>33,34</sup> Cultural expectations of appropriate gender roles add a further complexity as women are often expected to care for husbands or male relatives with TB and prioritize their access to nutrition, while the reverse typically does not hold.<sup>26</sup> Stigma can also lead to delays in seeking treatment as the comorbidity between HIV-AIDS and TB has generated a vernacular association between TB and immoral behavior in some low-middle income, high disease burden countries.<sup>22,35</sup> Stigma is often gendered with blame and shame apportioned to women and can result in isolation and psychosocial consequences for women which, in turn, affect access to care and health outcomes.<sup>20,24</sup> Although the gendered social consequences of TB are, therefore, context specific, they compound, exacerbating each other, resulting in individuals experiencing different forms of delay (access and referrals) to complete lack of access to medical care.<sup>22,36</sup>

A first step toward sex and gender awareness in developing research and programs of redress is, therefore, the implementation of clearer definitions of gender, not as an underlying synonym for sex, but rather as a complex relationship between social context, “sex,” “gender,” and illness psychosocial construction. This will assist in moving beyond the mere reporting of sex-related disparities in susceptibility to infection and progression to active disease and toward recognizing and responding to the reality that gender is associated with these disparities and has real physical, social, and economic impacts on TB outcomes.<sup>37</sup>

## Sex and gender sensitivity: MEDLINE (PubMed), EMBASE, and CINAHL’s controlled vocabularies

Databases such as MEDLINE, EMBASE, and CINAHL are powerful resources with sophisticated search tools and subject headings that are regularly updated. In these

databases, advanced searches operate by “mapping” entry terms (search terms or keywords) to subject headings. A controlled vocabulary is a list of terms (words or expressions) that are agreed upon, accepted, and updated, and which are then used by indexers to categorize literature content in a consistent manner. Accurate use of database terms, in turn, improves the process of searching and retrieving information. When controlled vocabulary terms are linked and structured in a specific manner, this list is called a thesaurus. However, database thesauri terms are not similar and some have more specific descriptors than others. When more specific descriptors are available, indexing becomes more precise and access to relevant literature is facilitated. Each of the three databases analyzed exhibits a varying level of sex and gender terminology search-ability, depending on the number and the degree of specificity of their subject headings as reflected in their thesauri. For instance, subject headings including the word gender are less common in MEDLINE (PubMed) compared with CINAHL. Table 1 displays gender-related terms available in these three databases.

It should be noted that for the purposes of this analysis, databases such as ISI Web of Science and some Cochrane Library Databases (eg, Cochrane Database for Systematic Reviews) were not included as they do not offer the ability to search through a controlled vocabulary.

## MEDLINE and PubMed (MeSH)

MEDLINE and PubMed have the same subject headings, called Medical Subject Headings (MeSH). While the terms are structured from general to specific, each is provided with its own “Scope Note”. Scope Notes contain information about

how to use the subject heading and may include: 1) a definition of the term, 2) which concept is included or excluded from it, 3) a suggestion of other subject headings that are more specific or that are related to the chosen subject heading, and 4) which subheading (qualifier or specific aspect of a subject heading) may be combined with this subject heading and which ones may not.

Considering first the three subject headings related to sex: Sex/, Sex characteristics/, and Sex factors/ as provided through the thesauri of these databases: first, Sex/, a general term referring to differences between males and females is defined in the scope note of the Mesh Database as: “the totality of characteristics of reproductive structure, functions, PHENOTYPE, and GENOTYPE, differentiating the MALE from the FEMALE organism.”<sup>38</sup> Second, Sex characteristics/, known as anatomical/physiological differences between men and women are defined as:

Those characteristics that distinguish one SEX from the other. The primary sex characteristics are the OVARIES and TESTES and their related hormones. Secondary sex characteristics are those which are masculine or feminine, but not directly related to reproduction.<sup>38</sup>

And the third subject heading, Sex factors/, usually referring to potential confounders, biases, and effect modifiers in the field of epidemiology is defined in PubMed as:

Maleness or femaleness as a constituent element or influence contributing to the production of a result [...]. It is used with human or animal concepts but should be differentiated from SEX CHARACTERISTICS, anatomical or physiological manifestations of sex, and from SEX DISTRIBUTION [...].<sup>38</sup>

It can already be seen that the definition for Sex factors/ is open to confusion when compared with the WHO definition of gender.<sup>13</sup> Other scope notes also generate a level of confusion resulting from inaccuracy in the definitions provided. For example, with regard to the scope note for Sex differentiation/, sex and gender are both used (erroneously) in reference to biological differences:

The process in developing sex- or gender-specific tissue, organ, or function after SEX DETERMINATION PROCESSES have set the sex of the GONADS. Major areas of sex differentiation occur in the reproductive tract (GENITALIA) and the brain.<sup>38</sup>

The distinction between “men” and “male” and between “women” and “female” was not captured in the thesaurus

**Table 1** Available MEDLINE, EMBASE, and CINAHL subject headings for gender-related terms

Subject headings	MEDLINE	EMBASE	CINAHL
Gender	NA	Yes	NA
Gender relations	NA	NA	NA
Institutionalized gender	NA	NA	NA
Gender roles	NA	NA	NA
Gender identity	Yes	Yes	Yes
Gender and sex	NA	Yes	NA
Gender bias	NA	Yes	Yes
Sex role	NA	Yes	Yes
Sexism	Yes	Yes	Yes
Masculinity	Yes	Yes	Yes
Femininity	Yes	Yes	Yes
Gender-specific care	NA	NA	Yes

**Abbreviations:** CINAHL, Cumulative Index to Nursing and Allied Health Literature; EMBASE, Excerpta Medica Database; MEDLINE, Medical Literature Analysis and Retrieval System Online; NA, not available.



prior to 1980 when descriptors for Men/ defined as “Human males as cultural, psychological, sociological, political, and economic entities”<sup>38</sup> and for Women/ defined as “Human females as cultural, psychological, sociological, political, and economic entities”<sup>38</sup> were added to allow for more appropriate indexing.

Within the MeSH listings, gender-related terms are not defined with the same specificity accorded to sex-related terms. Referring to selected definitions, Gender identity/ is defined as:

A person's concept of self as being male and masculine or female and feminine, or ambivalent, based in part on physical characteristics, parental responses, and psychological and social pressures. It is the internal experience of gender role.<sup>38</sup>

On searching MeSH terms, researchers find that there are limited descriptors to reflect the more complex and culturally specific facets of gender. For instance, in MEDLINE, the descriptors which are currently available are Gender identity/, Masculinity/, Femininity/, Sexism/. It is worthy of note that Femininity/, defined as “Female-associated sex-specific social roles and behaviors unrelated to biologic function,”<sup>38</sup> and Masculinity/, defined as “Male-associated sex-specific social roles and behaviors unrelated to biologic function”<sup>38</sup> were introduced as MeSH descriptors in 2010. While this appears to be an effort to address issues of gender, it does not consider gender identities other than a male:female dichotomy. The term Sexism/ was introduced in 2013 and defined as “Prejudice or discrimination based on gender or behavior or attitudes that foster stereotyped social roles based on gender”.<sup>38</sup> While the addition of Sexism/ demonstrates a degree of progress, it would be more beneficial if the term “Gender bias” was in fact a subject heading in its own right in MEDLINE as it currently exists in CINAHL and EMBASE.

Prior to the progressive introduction of even these limited additional terms, indexers faced issues in adequately capturing the essence of some topics referring to “gender issues” in published research papers. The most common consequence is a tendency to index papers addressing sex or gender matters with the estimated “closest” descriptor, thereby exacerbating the potential for confusion and a lack of transparency when researching in selected databases. As an example, in the case of Gender Identity/ and Sexism/, Table 2 shows how a selected entry term will be mapped to a specific descriptor.

Although the provision of descriptors has improved overall in MEDLINE, when searching for literature it would be

**Table 2** Terms mapped to the descriptors: gender identity/ and sexism/ in MeSH

Gender identity/entry terms		Sexism/entry terms
Gender identities	Roles, men's	Sex bias
Identity, gender	Man's role	Bias, sex
Sex role	Role, man's	Gender bias
Role, sex	Woman's role	Bias, gender
Roles, sex	Role, woman's	Sex discrimination
Sex roles	Roles, woman's	Discrimination, sex
Gender role	Woman's roles	
Gender roles	Women's role	
Roles, gender	Role, women's	
Gender	Roles, women's	
Men's role	Women's roles	
Role, men's		

**Note:** Data from reference 38.

**Abbreviation:** MeSH, Medical Subject Headings.

beneficial to be able to develop a search query that includes both descriptors and keywords (free text search terms) to enhance the specificity as well as the sensitivity of the searches. This would be beneficial not only for complex terms as discussed here but also useful for all literature searches. Table 3 displays an example of a search strategy using subject headings and keywords in reference to TB, sex, and “gender.”

When applied to TB and a search query with descriptors in reference to sex is used in the period 1980–2015: (exp Tuberculosis/AND (Sex/ OR Sex Characteristics/ OR Sex factors/)), 1,059 papers are retrieved. When a specific equation with descriptors in reference to gender is used: (exp Tuberculosis/ AND (Gender identity/ OR Femininity/ OR Masculinity/ OR Sexism/)), only 10 (n=10) papers are retrieved.

When a free text search is conducted for the terms (Tuberculosis AND Sex) in the title, abstract, or keywords field, the results are 2,149 papers in the period 1980–2015. However, when a more specific query is used in the same fields, such as (Tuberculosis AND (Sex adj4 (Characteristics OR differen\* OR factor\*))), less irrelevant papers are retrieved (n=203). A free text search conducted for the terms (Tuberculosis AND Gender\*) in the title, abstract, or keywords fields resulted in 1,048 articles in the period 1980–2015. However, when a more sensitive query is used such as (Tuberculosis AND (Gender\* OR feminin\* OR masculin\* OR “sex role\*” OR Sexism)), additional articles were retrieved (n=1,058). It is difficult to state that the second equation is “better” or “more sensitive” than the first. The key issue is the lack of consistency and uniformity in the use of the terms by the authors, database

**Table 3** MeSH, Emtree, and CINAHL headings and free text (keywords) for basic tuberculosis and gender/sex searches

	MEDLINE	EMBASE	CINAHL
<b>Subject headings</b>	<b>MeSH (controlled vocabulary of MEDLINE)</b>	<b>EMTREE (controlled vocabulary of EMBASE)</b>	<b>CINHAL Headings (controlled vocabulary of CINAHL)</b>
Tuberculosis	exp <sup>a</sup> Tuberculosis/	exp Tuberculosis/	(MH <sup>38</sup> "Tuberculosis")
Gender	Gender identity/ Masculinity/ Femininity/	Gender/ Masculinity/ Femininity/ Gender Identity/ Gender bias/ Sex role/ Sexism/	(MH "Gender Identity") (MH "Femininity") (MH "Masculinity") (MH "Gender Bias") (MH "Gender Specific Care") (MH "Sex Role") (MH "Sexism") (MH "Sex Factors")
Sex	Sex factors/ Sex characteristics/ Sex/	Sex difference/ Sex/	(MH sex)
<b>Free text</b>			
Tuberculosis	Tuberculosis. ab,kw,ti <sup>c</sup>	Tuberculosis. ab,kw,ti.	TI <sup>d</sup> (tuberculosis) OR AB <sup>d</sup> (tuberculosis) OR MW <sup>d</sup> (tuberculosis)
Gender	Gender <sup>se</sup> . ab,kw,ti.  (Gender OR masculin* OR feminin* OR "sex role*" OR sexism). ab,kw,ti.	Gender*. ab,kw,ti.  (Gender OR masculin* OR feminin* OR "sex role*" OR sexism). ab,kw,ti.	TI gender* OR AB gender* OR MW gender* TI (masculin* or feminin* or "sex role*" or sexism) OR AB (masculin* or feminin* or "sex role*" or sexism) OR MW (masculin* or feminin* or "sex role*" or sexism)
Sex	Sex. ab,kw,ti	Sex. ab,kw,ti	TI sex OR AB sex OR MW sex
Sex specificity	(Sex adj <sup>4f</sup> (characteristics OR differen* OR factor*)). ti,ab,kw.	(Sex adj <sup>4</sup> (characteristics OR differen* OR factor*)). ti,ab,kw.	TI ((sex N <sup>4g</sup> (characteristics or differen* or factor*))) OR AB ((sex N <sup>4</sup> (characteristics or differen* or factor*))) OR MW ((sex N <sup>4</sup> (characteristics or differen* or factor*)))

**Notes:** <sup>a</sup>Explode a subject heading in MEDLINE and EMBASE (retrieves results using the selected descriptor and all of its more specific descriptors). <sup>b</sup>Explode a subject heading in CINAHL. <sup>c</sup>Runs a search through these fields: abstract (ab), keywords (kw), and title (ti) in MEDLINE and EMBASE. <sup>d</sup>Runs a search through these fields: title (TI), abstract (AB), keywords (MW) in CINAHL. <sup>e</sup>Truncation: searches for variations of words that are formed with different suffixes. <sup>f</sup>Retrieve records that contain two searched words (in any order) and with a maximum of *n* words between them in MEDLINE and EMBASE. <sup>g</sup>Retrieve records that contain two searched words (in any order) and with a maximum of *n* words between them in CINAHL.

**Abbreviations:** CINAHL, Cumulative Index to Nursing and Allied Health Literature; EMBASE, Excerpta Medica Database; Emtree, controlled vocabulary of EMBASE; MEDLINE, Medical Literature Analysis and Retrieval System Online; MeSH, Medical Subject Headings; MH, label for the subject heading field in CINAHL; MW, runs a search through the keyword and subject heading fields in CINAHL.

designers, and users of databases. Such sustained inconsistency affects the relevance of papers retrieved using keyword searches.

It is also worth reiterating in this discussion that any search using the terms sex or gender may be compromised by the widespread misuse of foundational terms and errors in indexing. As indicated earlier, whether in reference to descriptors or to keywords, there is widespread misuse on the one hand, and lack of systematic clarification on the other hand, of the terms sex and gender from the authors, the database designers, and database users trying to retrieve evidence. At the same time, given that the growing tendency in empirical evidence accounts for the differences between gendered experience, there is a growing expectation of greater inclusion of these complex terms.

## EMBASE (EMTREE)

Similar to MEDLINE's MeSH, EMBASE has its own hierarchically structured controlled vocabulary: Emtree. By allowing searchers to access correlated, searchable headings, Emtree allows a wider vocabulary for sex and gender terms. However, as Emtree does not provide scope notes for its subject headings, their meanings are open to interpretation, a situation which can lead to a potential counterproductive effect of losing the distinctiveness of the terms when articles are indexed. When searching EMBASE for actual descriptors for the topic of "sex," the available subject headings are sex/ and sex difference/, which is fewer than the number of subject headings provided by MEDLINE. With regard to gender-related topics, Gender/, Masculinity/, Femininity/, Gender Identity/, Sex Role/, Sexism/, and Gender Bias/ are available.

This database also offers a precombined term: Gender and Sex/; however, it is possible that this combined term could intensify confusion and inconsistent application in a domain where terms are not defined. An Emtree search for “men” or “women” automatically mapped to Male/ and Female/ replacing social terms with their biological counterparts. This is another example of a counterproductive search conflating gender and “sex,” since both terms are commonly used to index nearly all studies conducted with human subjects, resulting in over 6 million articles indexed under each term.

When a sex and gender search for TB is entered into EMBASE, the results are as follows: A sex search for papers published during the 1980–2015 period using the descriptor query: (exp Tuberculosis/ AND (Sex/ OR Sex difference/)) generates 1914 results. When a gender search for papers published during the same period is conducted using the descriptors: (exp Tuberculosis/ AND (Gender/ OR Gender Identity/ OR Gender Bias/ OR Sexism/ OR Sex role/ OR Masculinity/ OR Femininity/)), the result in EMBASE is 879 papers. A free text search conducted for the terms (Tuberculosis AND Sex) in the title, abstract, or keywords fields expands to 2,833 papers. A more refined or specific free text search using the terms (Tuberculosis AND (Sex adj4 (Characteristics OR differen\* OR factor\*))) narrows the results to 283 papers. When a free text search conducted for the terms (Tuberculosis AND Gender\*) in the title, abstract, or keywords fields resulted in 1,577 papers between 1980 and 2015. However, for a more sensitive query such as (Tuberculosis AND (Gender\* OR feminin\* OR masculin\* OR “sex role\*” OR Sexism)) conducted in the title, abstract, or keywords, a few more papers were retrieved, for example, 1,592 was the case with MEDLINE. It is, therefore, shown that when dealing with Emtree, similar challenges are encountered as those identified with the use of MEDLINE.

## CINAHL (CINAHL Headings)

The third database reviewed, CINAHL, shares similarities with the two other databases in that it provides readers with scope notes for a controlled vocabulary (MEDLINE) while displaying a wider coverage of gender descriptors (EMBASE). However, the terms sex and gender are often inappropriately merged within the scope notes, demonstrating a failure to clearly understand sex- and gender-related terms. This conflation creates problems for researchers and indexers, as resources are not indexed with relevant distinct descriptors. For example, the CINAHL Headings scope note for the descriptor (Subject label, abbreviated MH) “Sex

Factors” reads: “Use when gender is discussed as a factor in relation to some specific subject or problem, usually a statistical concept.”<sup>39</sup> The intent of distinguishing sex from gender is lost in this case when gender is absorbed into a sex-related statistical concept. This is contrary to the scope note of the descriptor (MH “Sex Role”), which validates the socially constructed understanding of gender by stating: “The behavior and appearance that one presents in terms of what the culture considers to be masculine or feminine.”<sup>36</sup> The gender-related term (MH “Gender Identity”) is clearly differentiated from sex-related terms being defined as:

Those aspects of appearance and behavior which are attributed to masculinity or femininity and determined by cultural and familial factors to which the individual is exposed. Differentiate from biologically determined sex characteristics.<sup>39</sup>

In addition, similar to Emtree, coverage of gender-related terms in CINAHL is greater than the ones available in MEDLINE as opposed to sex-related terms. However, it remains problematic in that gender continues to be described only in terms of masculine and feminine in the database definitions.

When a sex and/or gender inclusive search for TB is put into practice in CINAHL, the result differs significantly from MEDLINE and EMBASE. A sex inclusive search for papers published during the 1980–2015 period using the descriptors (MH “Tuberculosis+”) AND (MH sex) gave only three (n=3) papers after a systematic validation of the strategy by at least two of the authors of this paper. This manual validation was applied as the descriptors that contained the word sex lead to scope notes, which referred to gender or included the word gender in the definition. Given that ambiguity, the majority of papers were not selected, except for the ones that were considered to “truly” discuss sex as a variable and not as a proxy for “gender.”

With regard to a gender-based search for publications between the period 1980 and 2015 conducted using the descriptors ((MH “Tuberculosis+”) AND ((MH “Sex Role”) OR (MH “Gender Bias”) OR (MH “Gender Identity”) OR (MH “Sexism”) OR (MH “Femininity”) OR (MH “Masculinity”) OR (MH “Sex Factors”) OR (MH “Gender Specific Care”))), 211 papers were retrieved. The descriptor “Gender Specific Care” is included in the equation in order to attempt to find all papers considering gender-sensitive interventions and programs. However, the scope note associated with this descriptor is “Care that is sensitive to sex-specific

needs. General only; consider also MEN'S HEALTH and WOMEN'S HEALTH",<sup>36</sup> further demonstrating the lack of clear differentiation between sex and gender.

When a free text search is conducted for the term (Tuberculosis AND Sex) in the title, abstract, or keywords fields, 448 papers are identified. A more refined free text search using the terms (Tuberculosis AND (Sex N4 (Characteristics OR differen\* OR factor\*))) resulted in 256 papers.

When a free text search was conducted for the terms (Tuberculosis AND Gender\*) in the title, abstract, or keywords fields, it resulted in 190 papers in the period 1980–2015. However, when a more sensitive query such as (Tuberculosis AND (Gender\* OR feminin\* OR masculin\* OR "sex role\*" OR Sexism)) is conducted in the title, abstract, or keywords, a few additional papers (195) were retrieved.

As with the points made above, a search strategy using CINHALL is confronted by the issues raised with MEDLINE and EMBASE, with regards to the potential mishandling by authors, designers of databases, and database users of the concept of gender as a simple synonym for biological sex.

## What have we learnt? Where is the needle in the haystack?

### MeSH, Emtree, and CINAHL Headings: searching for specific terminology

The diversity of the impacts of sex (biology) and gender (socially and contextually constructed) in health is a field of growing awareness and multifaceted complexity. Awareness of these impacts and the need for appropriate methods of disease redress is simply good science. With current database search tools, however, this information is like a needle in a very large haystack of research. It, therefore, behoves all database users to pay close attention to the definitions of sex and gender and the important, and distinctly different impacts of each on health and disease experience.

Analysis of sex and gender terms in medical databases was immediately limited by the controlled vocabularies built-in by database developers. For the selected three databases, the observed lack of uniformity in the terms associated with controlled vocabularies in reference to sex and gender created an opportunity for confusion. The inconsistency of coverage of terms in the Emtree and MeSH vocabularies was equally observed.<sup>40</sup> With searches using the term Knowledge Translation in CINAHL, diverse vocabularies and inconsistent terminology usage have been associated with difficulties in retrieving relevant papers using subject heading searches.<sup>41</sup> The analysis additionally highlighted the inconsistent use and the lack of clarity regarding, sex, gender,

and related terminologies by authors, reviewers, editors, and end-users. This, in turn, affects the task of indexers in assigning appropriate subject headings and perpetuates a cycle of confusion in seeking to access the sex- and gender-sensitive literature.

Bekhuis et al<sup>42</sup> recommend using free text searches to capture larger numbers of related literature results rather than the controlled vocabulary approach evaluated in this paper. Compounding these general criticisms, the lack of conceptual clarity of operational definitions and the use of individual and inaccurate interpretations are among the key limitations to undertaking comprehensive sex- and gender-based searches of literature for a given health topic. Furthermore, the novice database user can be easily confused as the databases do not form a consensus around definitions, for example, sex factors/ in MEDLINE refers more explicitly to sex, whereas in CINHALL this descriptor refers to gender. Too often, gender remains a synonym for biological sex. Similarly, descriptors such as "sex role", which represents one of the components of gender, are used (incorrectly) as a term referring to sex.<sup>5</sup> Despite the increasing understandings of gender in non-health-related disciplines in the three databases reviewed, it remains poorly defined and fails to capture the subtleties associated with the multiple socially constructed facets of gender as this concept continues to evolve in the scientific arena.

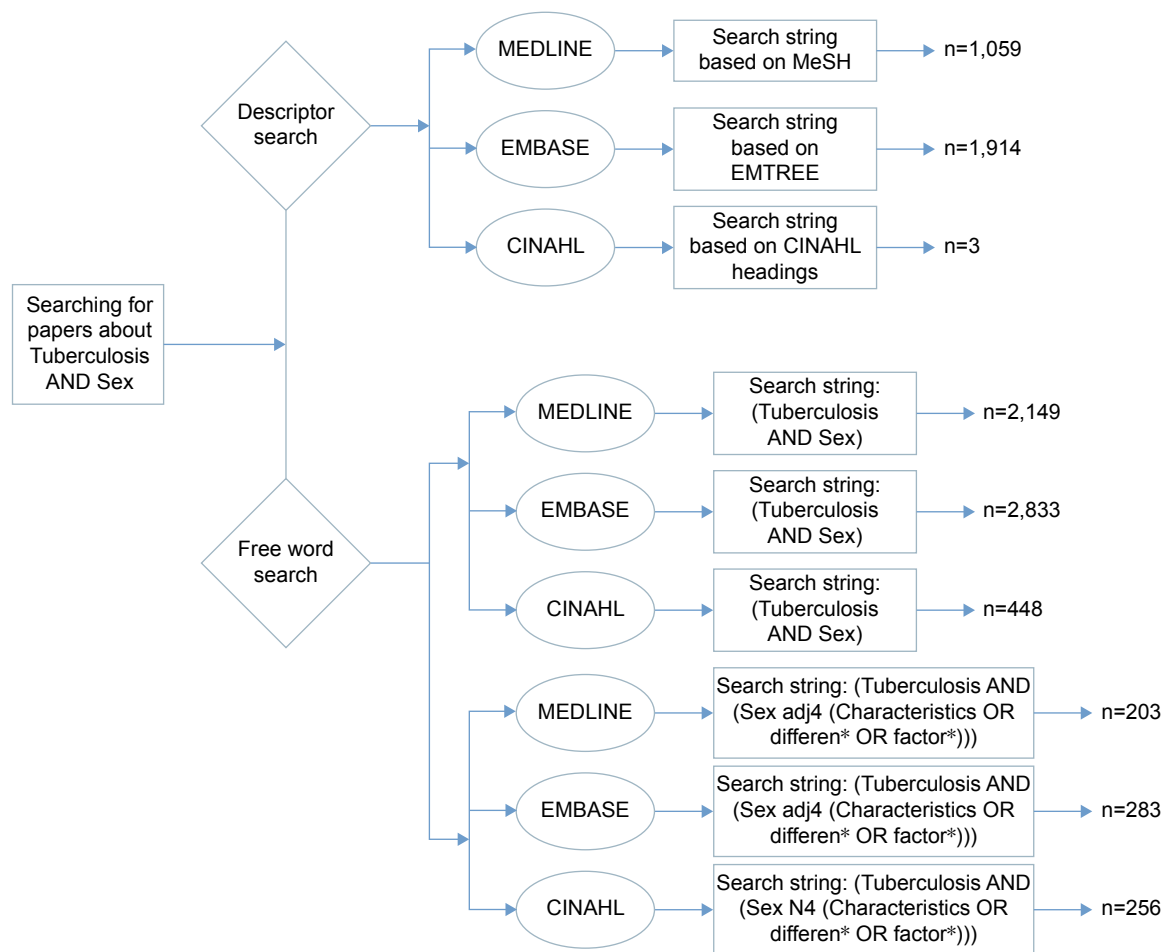
In the case of TB, there is an increasing literature on how sex and/or gender operate differentially in the trajectory of the disease development. This analysis has illustrated the ambiguity, contradiction, and potential for confusion in retrieving specific and sensitive literature in reference to gender and sex as it pertains to TB. To assist in providing an overview of the disjunctions in search result findings, Figures 1 and 2 show the number of papers identified in each of the three databases for the search parameters described above. It is important to recognize the potential risk to quality clinical practice of failing to identify all the relevant determinants of disease that combine to exacerbate existing health inequalities in the management and control of an already complex condition.

## Moving toward finding the needle

Given the results of this analysis and its alignment with the findings of existing literature into academic health-related database use, until database systems are improved it is recommended that:

1. Users access at least all three databases in parallel when undertaking research incorporating sex and gender.





**Figure 1** Observed ambiguities in search results for tuberculosis and sex-related papers.

**Abbreviations:** CINAHL, Cumulative Index to Nursing and Allied Health Literature; EMBASE, Excerpta Medica Database; Emtree, controlled vocabulary of EMBASE; MEDLINE, Medical Literature Analysis and Retrieval System Online; MeSH, Medical Subject Headings.

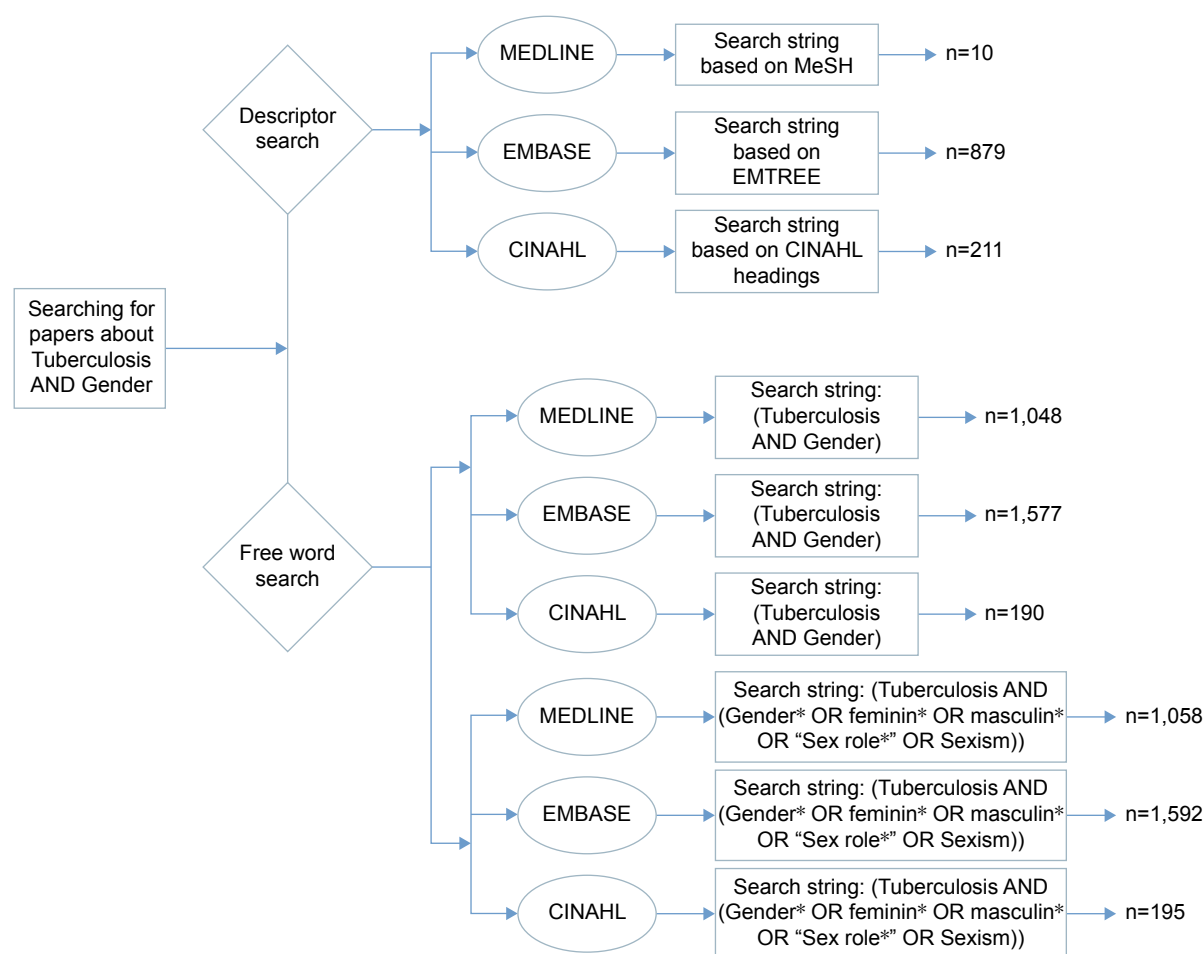
In spite of the limitations within each system, there are advantages in each database's unique resource base.<sup>43</sup>

2. All related descriptors of a concept such as sex or gender are included in the query according to the selection offered by the thesaurus of that database.
3. Controlled vocabulary searches be supplemented by free text searches in order to maximize the sensitivity of the search and that all relevant data are located.

It is recognized that in the short term, this will generate large identified data sets for users and additional work scrutinizing to identify appropriate resources. As health research moves forward, tools are becoming increasingly available to help researchers and users of research to better identify research that considers sex and gender including, for example, briefing notes,<sup>1,44</sup> resource guides,<sup>45,46</sup> web resources such as the gendered innovations website created by Stanford University<sup>47</sup> and The Canadian Institutes of Health Research online tool for peer reviewers.<sup>48</sup> Questions from the CIHR

tool for peer reviewers are designed to support researchers in defining and operationalizing sex and gender in relation to their studies and are displayed in Table 4.<sup>48</sup> Efforts such as these will greatly enhance public health research while uncovering the uneven burden of both infectious diseases and chronic conditions as they affect individuals of all sexes and genders in a differential manner.

While sex is a term used in a more straightforward fashion, gender is utilized with nuanced definitions. However, improvement by thesaurus editors of the clarity of terminology definitions in line with up-to-date sex and gender definitions will allow greater precision both for the indexers who perform subject analysis of papers that populate databases, as well as the users themselves. Similarly, authors need to be increasingly attentive to the appropriateness of the keywords used. Peer reviewers are also part of this cycle and should validate the suitability of the author's keywords and terms used (in full text papers) in order that papers be



**Figure 2** Observed ambiguities in search results for tuberculosis and gender-related papers.

**Abbreviations:** CINAHL, Cumulative Index to Nursing and Allied Health Literature; EMBASE, Excerpta Medica Database; Emtree, controlled vocabulary of EMBASE; MEDLINE, Medical Literature Analysis and Retrieval System Online.

**Table 4** Questions to guide improved integration of sex- and gender-related concepts in health research

Study element	Relevant questions
Approaching the research issue	<p>Are the concepts of gender and/or sex used in the proposed research project?</p> <ul style="list-style-type: none"> <li>If yes, has the applicant explicitly defined the concepts of gender and/or sex? Is it clear what aspects of gender and/or sex are being examined in the study?</li> <li>If no, do you consider this to be a significant oversight? Given your knowledge of the relevant literature, are there plausible gender and/or sex factors that should have been considered? If you consider sex and/or gender to be highly relevant to the proposed research, your feedback to the applicant should reflect this</li> </ul>
Drafting research questions and hypotheses	Does the research question(s) or hypothesis/es make reference to gender and/or sex, or relevant groups or phenomena? (eg, differences between males and females, differences among women, seeking to understand a gendered phenomenon, such as masculinity)
Reviewing the literature	Does the literature review cite prior studies that support the existence of significant differences between women and men, boys and girls, or males and females?
Applying the methods	<p>Does the literature review point to the extent to which past research has taken gender or sex into account?</p> <p>Is the sample appropriate to capture gender- and/or sex-based factors? Is it possible to collect data that are disaggregated by sex and/or gender? Are the inclusion and exclusion criteria well justified with respect to sex and/or gender?</p> <p>(Note: this pertains to human and animal subjects and non-organismic biological systems)</p> <p>Is the data collection method proposed in the study appropriate for investigations of sex and/or gender?</p> <p>Is the analytic approach appropriate and sufficiently rigorous to capture gender- and/or sex-based factors?</p>
Ethics	Does the study design account for the relevant ethical issues that might have particular significance with respect to gender and/or sex? (eg, inclusion of pregnant women in clinical trials)

**Notes:** Canadian Institutes of Health Research; Excerpt from Integrating Gender and Sex in Health Research: A Tool for CIHR Peer Reviewers. Available from: <http://www.cihr-irsc.gc.ca/e/43216.html>; 2014. All rights reserved. Reproduced with the permission of the Canadian Institutes of Health Research, 2016.<sup>48</sup>

better indexed. Such disciplined and sustained actions will ultimately contribute to the ongoing terminology development as well as the specificity and sensitivity within indexing systems. These changes can only advance good clinical practice and increase quality of care by admitting the need to clearly address contextual determinants of sex and gender both in understanding disease trajectory and in developing effective means of redress.

## Acknowledgments

The authors give sincere thanks for financial support to the Shastri Indo-Canadian Institute, Collaborative grant and to the Canadian Institute of Health Research, Knowledge Synthesis Grant.

Sincere thanks also go to Aleha Aziz, Anisa Benjamin, Alisha Nicole Apale, Cristina Perez, and Alexandrine Lussier for their contributions as research assistants.

## Author contributions

All authors contributed significantly to the conception and design of the study, data analysis, writing and drafting of the manuscript, and revision of the publication. All authors have approved the final version to be published and have agreed to be accountable for ensuring that questions related to the accuracy or integrity of the work will be appropriately investigated and resolved.

## Disclosure

The authors report no conflicts of interest in this work.

## References

1. Doull M, Welch V, Puil L, et al. Development and evaluation of 'briefing notes' as a novel knowledge translation tool to aid the implementation of sex/gender analysis in systematic reviews: a pilot study. *PLoS One*. 2014;9(11):e110786.
2. Hawkes S, Buse K. Gender and global health: evidence, policy, and inconvenient truths. *Lancet*. 2013;381(9879):1783–1787.
3. Stronks K, Wieringa NF, Hardon A. Confronting diversity in the production of clinical evidence goes beyond merely including under-represented groups in clinical trials. *Trials*. 2013;14:177.
4. World Health Organization (WHO). Global tuberculosis report. Available from: [http://www.who.int/tb/publications/global\\_report/en/\(2014a\)](http://www.who.int/tb/publications/global_report/en/(2014a)). Accessed November 5, 2014.
5. Johnson JL, Greaves L, Repta R. Better science with sex and gender: facilitating the use of a sex and gender-based analysis in health research. *Int J Equity Health*. 2009;8:14.
6. Bahaadinbeigy K, Yogesan K, Wootton R. MEDLINE versus EMBASE and CINAHL for telemedicine searches. *Telemed J E Health*. 2010;16(8):916–919.
7. Bottorff JL, Oliffe JL, Robinson CA, Carey J. Gender relations and health research: a review of current practices. *Int J Equity Health*. 2011;10:60.
8. Chief Public Health Agency of Canada (The). Report, Chapter 4: incorporating sex and gender into health interventions. In: Public Health Agency of Canada; 2012. Available from: <http://www.phac-aspc.gc.ca/cphorsphc-respacsp/2012/chap-4-eng.php>. Accessed November 10, 2014.
9. Hammarström A, Annandale E. A conceptual muddle: an empirical analysis of the use of 'sex' and 'gender' in 'gender-specific medicine' journals. *PLoS One*. 2012;7(4):e34193.
10. Hammarström A, Johansson K, Annandale E, et al. Central gender theoretical concepts in health research: the state of the art. *J Epidemiol Community Health*. 2014;68(2):185–190.
11. Muehlenhard C, Peterson Z. Distinguishing between sex and gender: History, current conceptualizations, and implications. *Sex Roles*. 2011;64(11):791–803.
12. Stewart M, Kushner KE, Gray J, Hart DA. Promoting gender equity through health research: impacts and insights from a Canadian initiative. *Glob Health Promot*. 2013;20(1):25–38.
13. World Health Organization (WHO). Tuberculosis and Gender. Available from: [http://www.who.int/tb/challenges/gender/page\\_1/en/\(2014b\)](http://www.who.int/tb/challenges/gender/page_1/en/(2014b)). Accessed November 16, 2014.
14. Fredriksen-Goldsen KI, Simoni JM, Kim H, et al. The health equity promotion model: reconceptualization of lesbian, gay, bisexual, and transgender (LGBT) health disparities. *Am J Orthopsychiatry*. 2014;84(6):653–663.
15. Hammarström A, Johansson K, Annandale E, et al. Central gender theoretical concepts in health research: the state of the art. *J Epidemiol Community Health*. 2014;68(2):185–190.
16. Hoffman BR. The interaction of drug use, sex work, and HIV Among transgender women. *Subst Use Misuse*. 2014;49(8):1049–1053.
17. Runnels V, Tudiver S, Doull M, Boscoe M. The challenges of including sex/gender analysis in systematic reviews: a qualitative survey. *Syst Rev*. 2014;3:33.
18. Reitmanova S, Spitzer D. The need for gender-based analysis in health research of tuberculosis in Canada: The case of (in) visible minority foreign-born women. *Women's Health and Urban Life*. 2013;12(2):9–22.
19. Nowatzki N, Grant KR. Sex is not enough: the need for gender-based analysis in health research. *Health Care Women Int*. 2011;32(4):263–277.
20. Atre S, Kudale A, Morankar S, Gosoni D, Weiss MG. Gender and community views of stigma and tuberculosis in rural Maharashtra, India. *Glob Public Health*. 2011;6(1):56–71.
21. Jain NK, Thakkar MS, Jain N, Rohan KA, Sharma M. Chronic obstructive pulmonary disease: does gender really matter? *Lung India*. 2011;28(4):258–262.
22. Kaur M, Sodhi SK, Kaur P, Singh J, Kumar R. Gender differences in health care seeking behaviour of tuberculosis patients in Chandigarh. *Indian J Tuberculosis*. 2013;60:217–222.
23. Kaur N. A prospective study of gender disparity in tuberculosis transmission and control. *Int J Res Advent Technol*. 2014;2:1.
24. Shiotani R, Hennink M. Socio-cultural influences on adherence to tuberculosis treatment in rural India. *Glob Public Health*. 2014;9(10):1239–1251.
25. Krishnan L, Akande T, Shankar AV, et al. Gender-related barriers and delays in accessing tuberculosis diagnostic and treatment services: a systematic review of qualitative studies. *Tuberc Res Treat*. 2014;2014:215059.
26. Khan KB. Understanding the gender aspects of tuberculosis: a narrative analysis of the lived experiences of women with TB in slums of Delhi, India. *Health Care Women Int*. 2012;33(1):3–18.
27. World Health Organization (WHO). *The Five Elements of Dots*; 2014. Available from: <http://www.who.int/tb/dots/whatisdots/en/>. Accessed November 15, 2014.
28. Codlin AJ, Khawaja S, Chen Z, et al. Gender differences in tuberculosis notification in Pakistan. *Am J Trop Med Hyg*. 2011;85(3):514–517.
29. Greenaway C, Sandoe A, Vissandjée B, et al. Tuberculosis evidence review for newly arriving immigrants and refugees. *CMAJ*. 2011;183(12):E939–E951.
30. Mukherjee A, Saha I, Sarkar A, Chowdhury R. Gender differences in notification rates, clinical forms and treatment outcome of tuberculosis patients under the RNTCP. *Lung India*. 2012;29(2):120–122.
31. Neyrolles O, Quintana-Murci L. Sexual inequality in tuberculosis. *PLoS Med*. 2009;6(12):e1000199.

32. Ramsay A, Bonnet M, Gagnidze L, Githui W, Varaine F, Guerin PJ. Sputum, sex and scanty smears: new case definition may reduce sex disparities in smear-positive tuberculosis. *Int J Tuberc Lung Dis*. 2009; 13(5):613–619.
33. Kulkarni P, Kudale A, Arasu K, Lab M, Darby W, Rangan S. Tuberculosis knowledge and awareness in tribal-dominant districts of Jharkhand, India: implications for ACSM. *Public Health Action*. 2014; 4(3):189–194.
34. Onifade DA, Bayer AM, Montoya R, et al. Gender-related factors influencing tuberculosis control in shantytowns: a qualitative study. *BMC Public Health*. 2010;10:381.
35. Vissandjee B, Pai M. The socio-cultural challenge in public health interventions: the case of tuberculosis in India. *Int J Public Health*. 2007;52(4):199–201.
36. Yang WT, Gounder CR, Akande T, et al. Barriers and delays in tuberculosis diagnosis and treatment services: does gender matter? *Tuberc Res Treat*. 2014;2014:461935.
37. Vissandjée B, Lingam L. Enjeux liés aux approches sensibles à l'interaction entre le genre et la migration : construire des partenariats, revoir les pratiques pour un accès équitable à des soins de santé [Challenges associated with interventions sensitive to the interaction between gender and migration: Building partnerships and reviewing practices to ensure equitable access to health care]. In: Suárez-Herrera JC, Blain MJ, editors. *La Recherche En Santé Mondiale: Perspectives Socio-Anthropologiques* [Global Health Research: Socio-Anthropological Perspectives]. Montréal, Canada: Collection des cahiers scientifiques de l'ACFAS; 2012:284. French.
38. MeSH. US National Library of Medicine, National Center for Biotechnology Information. Bethesda MD. Available from: <http://www.ncbi.nlm.nih.gov/mesh>. Accessed November 5, 2015.
39. CINAHL Plus with Full Text: CINAHL Headings. EBSCO Industries, Inc., EBSCOhost. Birmingham, AL. Available from: <http://web.b.ebscohost.com/ehost/mesh?sid=f8d72c85-48ad-41ef-95ae-fe040fbcc0a6%40sessionmgr120&vid=4&hid=128>. Accessed November 5, 2015.
40. Richter RR, Austin TM. Using MeSH (medical subject headings) to enhance PubMed search strategies for evidence-based practice in physical therapy. *Phys Ther*. 2012;92(1):124–132.
41. Lokker C, McKibbin KA, Wilczynski NL, et al. Finding knowledge translation articles in CINAHL. *Stud Health Technol Inform*. 2010; 160(Pt 2):1179–1183.
42. Bekhuis T, Demner-Fushman D, Crowley RS. Comparative effectiveness research designs: an analysis of terms and coverage in medical subject headings (MeSH) and emtree. *J Med Libr Assoc*. 2013; 101(2):92–100.
43. Névéal A, Doğan R, Lu Z. Author keywords in biomedical journal articles. *AMIA Annu Symp Proc*. 2010;2010:537–541.
44. Johnson J, Beaudet A. Sex and gender reporting in health research: why Canada should be a leader. *Can J Public Health*. 2012;104(1): e80–e81.
45. Johnson J, Sharman Z, Vissandjee B, Stewart DE. Does a change in health research funding policy related to the integration of sex and gender have an impact? *PLoS One*. 2014;9(6):e99900.
46. Sharman Z, Johnson J. Towards the inclusion of gender and sex in health research and funding: an institutional perspective. *Soc Sci Med*. 2012;74(11):1812–1816.
47. Blom K. Review of a gendered innovation: can a website sell the benefits of sex and gender to scientists? *Women's Health Urban Life*. 2013;12:120–126. Available from: <http://genderedinnovations.stanford.edu/>. Accessed November 16, 2014.
48. Canadian Institutes of Health Research (CIHR). Integrating gender and sex in health research: a tool for CIHR peer reviewers. In: Peer review: policies and procedures; 2014. Available from: <http://www.cihr-irsc.gc.ca/e/43216.html>. Accessed December 10, 2014.

## International Journal of Women's Health

### Publish your work in this journal

The International Journal of Women's Health is an international, peer-reviewed open-access journal publishing original research, reports, editorials, reviews and commentaries on all aspects of women's healthcare including gynecology, obstetrics, and breast cancer. The manuscript management system is completely online and includes

Submit your manuscript here: <http://www.dovepress.com/international-journal-of-womens-health-journal>

Dovepress

a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.