Cochrane Library

Cochrane Database of Systematic Reviews

Search strategies to identify observational studies in MEDLINE and Embase (Review)

Li L, Smith HE, Atun R, Tudor Car L

www.cochranelibrary.com

TABLE OF CONTENTS

[Methodology Review]

Search strategies to identify observational studies in MEDLINE and Embase

Li Li[1], Helen E Smith[2], Rifat Atun[3], Lorainne Tudor Car[2]

[1]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore. [2]Family Medicine and Primary Care, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore. [3]Department of Global Health and Population, Harvard School of Public Health, Boston, MA, USA

Contact: Lorainne Tudor Car, Family Medicine and Primary Care, Lee Kong Chian School of Medicine, Nanyang Technological University, 11 Mandalay Road, Level 18-08 Clinical Sciences Building, Singapore, Singapore. lorainne.tudor.car@gmail.com.

A B S T R A C T

Background

Systematic reviews are essential for decision-making. Systematic reviews on observational studies help answer research questions on aetiology, risk, prognosis, and frequency of rare outcomes or complications. However, identifying observational studies as part of systematic reviews efficiently is challenging due to poor and inconsistent indexing in literature databases. Search strategies that include a methodological filter focusing on study design of observational studies might be useful for improving the precision of the search performance.

Objectives

To assess the sensitivity and precision of a search strategy with a methodological filter to identify observational studies in MEDLINE and Embase.

Search methods

We searched MEDLINE (1946 to April 2018), Embase (1974 to April 2018), CINAHL (1937 to April 2018), the Cochrane Library (1992 to April 2018), Google Scholar and Open Grey in April 2018, and scanned reference lists of articles.

Selection criteria

Studies using a relative recall approach, i.e. comparing sensitivity or precision of a search strategy containing a methodological filter to identify observational studies in MEDLINE and Embase against a reference standard, or studies that compared two or more methodological filters.

Data collection and analysis

Two review authors independently screened articles, extracted relevant information and assessed the quality of the search strategies using the InterTASC Information Specialists' Sub-Group (ISSG) Search Filter Appraisal Checklist.

Main results

We identified two eligible studies reporting 18 methodological filters. All methodological filters in these two studies were developed using terms from the reference standard records.

The first study evaluated six filters for retrieving observational studies of surgical interventions. The study reported on six filters: one Precision Terms Filter (comprising terms with higher precision while maximum sensitivity was maintained) and one Specificity Terms Filter (comprising terms with higher specificity while maximum sensitivity was maintained), both of which were adapted for MEDLINE, for Embase, and for combined MEDLINE/Embase searches. The study reported one reference standard consisting of 217 articles from one systematic review of which 83.9% of the included studies were case series

The second study reported on 12 filters for retrieving comparative non-randomised studies (cNRSs) including cohort, case-control, and cross-sectional studies. This study reported on 12 filters using four different approaches: Fixed method A (comprising of a fixed set of controlled vocabulary (CV) words), Fixed method B (comprising a fixed set of CV words and text words (TW)), Progressive method (CV) (a random choice of study design-related CV terms), and Progressive method (CV or TW) (a random choice of study design-related CV terms, and title and abstracts-based TWs). The study reported four reference standards consisting of 89 cNRSs from four systematic reviews.

The six methodological filters developed from the first study reported sensitivity of 99.5% to 100% and precision of 16.7% to 21.1%. The Specificity Terms Filter for combined MEDLINE/Embase was preferred because it had higher precision and equal sensitivity to the Precision Terms Filter. The 12 filters from the second study reported lower sensitivity (48% to 100%) and much lower precision (0.09% to 4.47%). The Progressive method (CV or TW) had the highest sensitivity.

There were methodological limitations in both included studies. The first study used one surgical intervention-focused systematic review thus limiting the generalizability of findings. The second study used four systematic reviews but with less than 100 studies. The external validation was performed only on Specificity Terms Filter from the first study Both studies were published 10 years ago and labelling and indexing of observational studies has changed since then.

Authors' conclusions

We found 18 methodological filters across two eligible studies. Search strategies from the first study had higher sensitivity and precision, underwent external validation and targeted observational studies. Search strategies from the second study had lower sensitivity and precision, focused on cNRSs, and were not validated externally. Given this limited and heterogeneous evidence, and its methodological limitations, further research and better indexation are needed.

P L A I N   L A N G U A G E   S U M M A R Y

Search strategies to identify observational studies in MEDLINE and Embase

Background

Systematic reviews collect, analyse and summarise research to answer defined research questions. The evidence from systematic reviews is considered as the most reliable evidence and is often used to inform healthcare-related decision making. Systematic reviews on prognosis, cause, risk factors and complications of a disease include a specific type of study design: observational studies. Searching for literature usually involves the use of MEDLINE and Embase databases and can result in a large number of articles to check for the review. Therefore, in order to focus their search, researchers often use a set of terms relating to study designs, known as methodological filters. However, when using these, researchers may miss relevant studies but find many irrelevant articles. Performance of filters are evaluated by sensitivity, which is the ability of the search filter to retrieve all the relevant studies that exist, and precision, which is the ability of the search filter to retrieve only relevant studies. Given the importance of including evidence from observational studies in systematic reviews, we aimed to assess studies evaluating methodological strategies for identifying observational studies in the two main databases of health literature, MEDLINE and Embase.

Study characteristics

We found two eligible studies reporting on 18 methodological filters, including six MEDLINE, six Embase and six combined MEDLINE/Embase filters. The firsts study focused on filters on observational studies of surgical interventions. The second study focused on filters for a specific subtype of observational studies: comparative non-randomised studies.

Key results

Six filters from the first study showed sensitivity of 99.5% to 100% and precision of 16.7% to 21.1%. One type of filter was evaluated by two additional systematic reviews (i.e. externally validated) and found that this retrieved 85.2% to 100% of the articles in the reference standard. Twelve filters from the second study had lower sensitivity (48% to 100%) and much lower precision (0.09% to 4.47%).

Quality of evidence

The included studies had several limitations. The first study used only one systematic review for search strategy development and focused on observational studies of surgical interventions, which might limit the generalizability of the findings to other literature searches. The reference standard in the second study, although encompassing four different systematic reviews, included a limited number of studies, which might affect the accuracy of the performance assessment. Both studies were published 10 years ago and labelling and indexing of observational studies has changed since then.

BACKGROUND

Description of the problem or issue

Systematic reviews of the literature have become vital decision-making aids for clinicians, researchers, policy makers and patients (Gough 2012a; Ligthelm 2007; Manchikanti 2009; Wilczynski 2007). They provide a formal synthesis of a large and ever increasing body of research literature. Systematic reviews typically address specified questions and can, as a result, help to (1) establish links between available information and potentially beneficial (or harmful) interventions, (2) compare and contrast conflicting results, and (3) identify gaps in medical knowledge (Manchikanti 2009; Wilczynski 2007).

In order to achieve their objectives, systematic reviews rely on the use of explicit strategies to search for relevant evidence and on methodological criteria against which to evaluate this evidence (Wilczynski 2007). When searching for relevant studies, researchers (including those conducting a systematic review) can make use of structured search strategies that can facilitate this process (Wilczynski 2007). However, searching for specific studies on a given topic can be challenging, particularly when searching for a specific study design. Wilczynski 2007 attributes this to the spread of relevant papers across numerous scientific journals, the inherent limits in indexing, and the lack of search skills amongst database users.

Systematic reviews vary in many respects, including the types of research questions they ask (Gough 2012b). The review question, in turn, determines the methods and the types of data that are most appropriate to answer the question (Gough 2012a; Gough 2012b). Systematic reviews assessing the effectiveness of interventions are typically best answered by data from randomised controlled trials (RCTs) (Glasziou 2001; Ligthelm 2007). Systematic reviews asking questions of aetiology and risk, prediction and prognosis, or frequency of rare outcomes or complications are usually best answered by data from observational studies (Furlan 2006; Glasziou 2001).

However, there are circumstances in which evidence from observational studies is needed in order to assess the effectiveness of interventions or safety outcomes: when data from RCTs are insufficient or when the findings of RCTs appear to be contradictory (Fraser 2006; Furlan 2006; Manchikanti 2009). Improvements in observational-study methods and statistical analyses have made observational studies an important source of evidence, particularly with regards to the side effects or adverse events associated with health interventions (Ligthelm 2007; Manchikanti 2009; Wieland 2005). As argued by Ligthelm 2007, observational studies can complement data from RCTs in order to provide an evidence base for clinical decision-making or for policy-making. While searching for RCTs has become a relatively simple task since the 1990s (Lefebvre 2013), limitations in indexing practices still make the identification of observational studies particularly challenging.

Description of the methods being investigated

The use of a search strategy in health-related bibliographic databases is the method required by Cochrane and other evidence-based healthcare organisations to identify relevant study reports for a systematic review. MEDLINE and Embase are the principal databases of biomedical scientific literature. Together, they contain abstracts for many millions of published articles, with the extent of coverage depending on the specific topic. Records in these databases can be searched electronically for words in the title or abstract, and for assigned index terms. The latter are controlled vocabulary terms that indexers assign to each record after reviewing them (Lefebvre 2011). Searching these two databases is usually the minimum requirement for anyone wishing to conduct a systematic review; although the degree of overlap varies; between 10% to 87% depending on the topic under consideration (Manchikanti 2009).

Search strategies can be complemented by including search filters. These refer to a predefined combination of terms that have been designed to retrieve a selection of records on the basis of a particular concept (ISSG 2017). Filters used to retrieve records on the basis of their study design are often referred to as methodological filters. The combination of search filters with content terms in turn determines the performance properties of a search strategy, namely, its sensitivity, precision (or positive predictive value (PPV)) and specificity (Doust 2005; Fraser 2006).

How these methods might work

Evaluating a search strategy relies on the availability of a reference standard against which to compare its performance. There are different approaches to development of reference standards. Handsearching is a manual method of searching for eligible studies through a selection of sources such as journals, databases, conference abstract books etc. (Lefebvre 2013). Relative recall approach compares the records retrieved by a search strategy that includes a methodological filter with those retrieved by a search strategy without a filter (Sampson 2006). Handsearching has been seen as a traditional gold standard for assessment of search strategy performance. Yet, given the exponential growth of the literature and research journals over the last couple of decades it is becoming less efficient and practical. Handsearching is time-consuming and may have limited generalizability (Hayward 1997; Doust 2005). Relative recall approach is less resource intensive and an increasingly popular method for evaluating search strategies (Durão 2015; Golder 2018; Yousefi-Nooraie 2013). While a relative recall approach can have certain limitations such as a highly specific reference set confounded by the topic-only searches, this can be mitigated through the use of multiple reviews in the evaluation (Sampson 2006).

In the context of systematic reviews, sensitivity and precision are the most relevant performance properties of a search filter (Sampson 2006). Sensitivity, also referred to as recall, is defined as the number of relevant records in a database identified by the search strategy as a proportion of the total number of relevant records in the database (Sampson 2006). The precision of a search strategy refers to the number of relevant reports identified by the search strategy as a proportion of the total number of records yielded by the search (Doust 2005; Furlan 2006 Sampson 2006).

Review authors should aim for search strategies that have both high sensitivity and high precision (Sampson 2006). In addition, authors should identify and include all possibly relevant reports (high sensitivity) in order to reduce the likelihood of bias in their systematic reviews, and to reduce random error in meta-analyses (Edwards 2002; Robinson 2002). At the same time, they should attempt to retrieve as few irrelevant records as possible (high precision) in order to minimise the burden on the resources

available (Gough 2012a; Gough 2012b; Sampson 2006). However, in reality there are trade-offs between these two properties.

An ideal methodological filter could help review authors to achieve this balance by maintaining the sensitivity of a content-only search strategy while increasing its precision (Doust 2005; Fraser 2006). Applying methodological filters to a search strategy could in theory limit the number of records retrieved in a search, while avoiding the exclusion of relevant papers. At the same time, a methodological filter could limit the number of records that need to be evaluated for inclusion in the review. However, by reducing the number of hits, methodological filters could increase the likelihood of missing relevant records that would otherwise be obtained for, and included in a systematic review.

Why it is important to do this review

Using the study design to search health-related literature can identify the study type of primary interest in an efficient and time-saving manner (Littleton 2004). Specifying the types of study design is relatively easy for RCTs owing to initiatives such as the Cochrane Central Register of Controlled Trials (CENTRAL) database of trials, the introduction of the Consolidated Standards of Reporting Trials (CONSORT) statement (which is linked to better reporting of RCTs in the titles and abstracts), appropriate indexing terms in MEDLINE and Embase, and the publication of highly sensitive filters (Fraser 2006; Lefebvre 2013).

The situation is different when dealing with observational studies. Indexing using Medical Subject Headings (MeSH) intervention terms is limited; and when used, these terms are usually applied inconsistently (Fraser 2006;Wieland 2005). Despite the introduction of statements such as the Strengthening the Reporting of Observational studies in Epidemiology (STROBE) guidelines, reporting of methodological detail is still poor in observational studies, contributing to the problems in indexing and searching (Fraser 2006; Manchikanti 2009; Vandenbroucke 2014). Furthermore, in MEDLINE, the search term "observational studies" is available as a 'Publication type' as well as a topic and how this is applied to the record depends on the interpretation of the indexer. The lack of appropriate search terms for observational studies has greatly contributed to the exclusion of methodological components from search strategies (Fraser 2006). As a consequence of this, searches often yield a large number of irrelevant records, leading to the inefficient use of resources and the time needed to complete a review increases (Doust 2005). For this reason, we explored the literature for developments in search approaches that could lead to the efficient identification of observational studies.

In addition, there appear to be no agreed universal standard criteria for the creation of a search strategy (Lemeshow 2005); although guidelines are available to anyone thinking of undertaking a systematic review, particularly in relation to RCTs. In its *Handbook for Systematic Reviews of Interventions* (Higgins 2011), Cochrane presents their Highly Sensitive Search Strategy (HSSS) for identifying RCTs in MEDLINE. The work of another group, the InterTASC Information Specialists' Sub-Group (ISSG), focuses on the identification, assessment and testing of search filters that are intended to select studies depending on their design or focus (ISSG 2017). They offer various resources related to study designs such as RCTs, observational studies, diagnostic studies, and economic evaluations, among others.

Attempts have been made to appraise the evidence for search filters. A Cochrane Review (Beynon 2013) has evaluated the performance of search filters designed to retrieve diagnostic test accuracy (DTA) studies in MEDLINE and Embase. However, we could not identify a similar protocol or review on search strategies for observational studies. A systematic review in this area could help to identify the specific features of a search strategy that could improve the identification of observational studies. As a result, this work could contribute to the creation of evidence-based standards for the formulation of search strategies, building on the previously published protocol for this review (Marcano Belisario 2013).

OBJECTIVES

To assess the sensitivity and precision of a search strategy with a methodological filter to identify observational studies in MEDLINE and Embase.

METHODS

Criteria for considering studies for this review

Types of studies

We included studies that compared search strategies that contained methodological filters for identifying observational studies in MEDLINE and Embase against a search strategy without a methodological filter, i.e. studies using a relative recall approach. We included studies evaluating performance properties of search strategies for observational studies that used methodological filters for searching observational studies in MEDLINE, Embase and both MEDLINE and Embase.

We excluded studies that compared the effectiveness of the same strategy across different search interfaces, studies that focused on the retrieval of observational studies from bibliographic databases other than MEDLINE or Embase, or studies that did not provide enough data to calculate the sensitivity or precision of a search strategy being evaluated. We did not exclude studies on the basis of their language or time of publication. We excluded studies using a handsearching approach to develop reference standard. We also excluded studies that did not focus explicitly on search strategies for retrieval of observational studies such as those focusing on retrieval of prognostic studies, quality improvement studies or epidemiological studies. We also excluded studies in which the reference standard included non-observational studies.

Types of data

We included data from published, unpublished and grey literature comparing a search strategy containing a methodological filter against a reference standard, or comparing two or more search strategies, for retrieving observational studies in MEDLINE and Embase.

Types of methods

We focused on methodological filters for searching for observational studies in MEDLINE or Embase. We defined observational studies as studies with the relevant, classic epidemiological designs (i.e. case-control, cohort and cross-sectional studies) as well as other study designs such as case series, controlled before-and-after (CBA) and interrupted time series (ITS) (Viswanathan 2013)

We categorised the systematic reviews that were used to assess the filers as development systematic reviews and validation systematic reviews. Development systematic reviews are systematic reviews used for the development of the methodological filters (Beynon 2013). Validation systematic reviews are systematic reviews employed to assess the performance of the developed methodological filters externally (Sampson 2006).

Types of outcome measures

The primary outcomes were the performance properties of the methodological filters, sensitivity and precision.

- Sensitivity: the number of relevant reports in a database that were identified by the search strategy as a proportion of the total number of relevant reports identified by the reference standard.
- Precision: the number of relevant studies identified by the search strategy divided by the total number of records retrieved by the search strategy.

We assessed the performance properties of the methodological filters against a reference standard. A reference standard in our review encompasses a set of studies from a development systematic review. We extracted, tabulated and compared the sensitivity and/or precision of a search strategy including a methodological filter. Sensitivity and/or precision were calculated if the relevant 2 x 2 data were available (Table 1).

There was no secondary outcome in this review.

Search methods for identification of studies

Electronic searches

We conducted database search on 27 April 2018 with the end date of publication on 31 March 2018 for studies that compared search strategies that contained methodological filters for retrieving observational studies in MEDLINE and Embase against those without methodological filters. We used search strategy outlined in Appendix 1 for MEDLINE (Ovid, from 1946). We modified the search strategy to suit each database. In front of each step. we added "#" for Embase (Ovid, from 1974; Appendix 2) and "S", the search ID, for CINAHL (from 1937; Appendix 3). We conducted "Advanced Search", followed by "Search Manager" in Nanyang Technological University (NTU) Cochrane Library database for Cochrane Reviews (from 1992; Appendix 4).

Searching other resources

We searched Google Scholar with three search strategies outlined in Appendix 5, and grey literatures in OpenGrey (Appendix 6). We also scanned citations of the included studies for any other eligible studies (Horsley 2011) and contacted authors for additional published and unpublished studies (Young 2011).

Data collection and analysis

Selection of studies

The retrieved references were imported into EndNote X7 and duplicates were removed. Two review authors, LL and LTC, screened the titles and abstracts, and assessed the full-text reports independently for compliance with the inclusion and exclusion criteria for this review. All disagreements were resolved through discussion.

Data extraction and management

Review authors LL and LTC discussed together to develop the structured data extraction form and extracted data from included studies independently. The data extraction forms were compared, and the discrepancies were discussed and followed up according to the original reports.

The following information was extracted from each included study:

- general information about the study (e.g. study authors, journal of publication, original language of publication, year of publication);
- study methods including aim of the study and study design;
- number and type of methodological filters being compared;
- number of records yielded by each of the methodological filters in MEDLINE, Embase or both;
- sensitivity and precision of each of the methodological filters; and
- reference standard method against which each methodological filter was compared.

Assessment of the quality of evidence in included studies

The UK InterTasc Information Specialists' Sub-Group (ISSG) Search Filter Appraisal Checklist (Glanville 2008) is a comprehensive appraisal tool which reports filter design methods and search performance measures, such as sensitivity and precision. It assesses the scope of the filters (limitations, generalizability and obsolescence), and the methods used to develop the filter, including the generation of the reference standard. The checklist includes most of the quality assessment categories we outlined in the protocol (Marcano Belisario 2013): (1) information and objective of the search filter; (2) identification of a reference standard; (3) search filter development; (4) performance assessment (or internal validation as it is referred to in the checklist), (5) external validation and (6) potential limitations and comparisons.

Data synthesis

We adopted a narrative synthesis approach to present the results. We created tables to summarise the data extracted from the included studies with separate tables for those methodological filters implemented in MEDLINE, in Embase, and in the two databases combined. We tabulated the performance measures of filters grouped by development systematic reviews, and validation systematic reviews, included in the eligible studies. In each table, we included information about the reference standard, the performance of the filters (i.e. sensitivity and precision), external validation etc. We compared the differences between the search strategies for each of the primary outcomes of interest: sensitivity or precision.

R E S U L T S

Description of studies

We conducted electronic database searches in April 2018 and retrieved a total of 2285 records from MEDLINE, Embase, CINHAL, the Cocharne Library and Open Grey. We also screened the first 500 records from each of the three Google Scholar searches. Upon de-duplication of records and screening of title and abstracts, we screened full texts for 28 records. In the end, two studies (Fraser

2006; Furlan 2006) were found eligible with five published records,
of which three were conference abstracts (Figure 1).

Figure 1.   PRISMA Study flow diagram
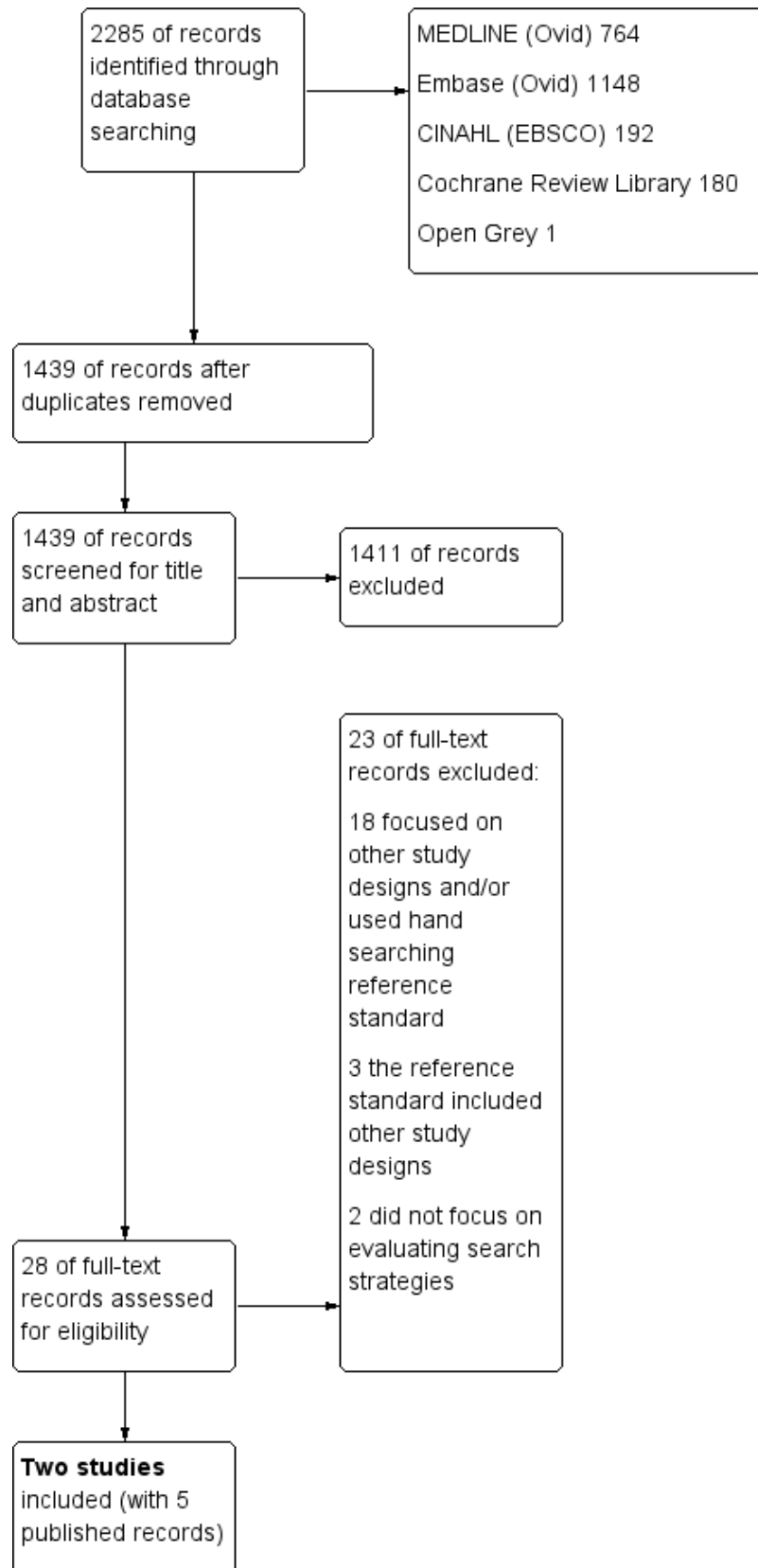
Included studies

The methodological filters in the included studies focused on observational studies of surgical interventions (Fraser 2006), and non-randomised comparison studies (cNRSs) (Furlan 2006) See Characteristics of included studies. The filters in Fraser 2006 retained the terms with highest precision and specificity while sensitivity was maximised. The filters in Furlan 2006 were developed with the aim of achieving a balance of sensitivity and precision. Both studies adopted a relative recall approach in which the reference standard is based on studies included in development systematic reviews (Beynon 2013; Glanville 2009; Sampson 2006). In addition, Fraser 2006 included additional observational studies retrieved during the development systematic review screening process which were not eligible for inclusion in the final review. Methodological filters in both studies were developed using search terms from the reference standard. Fraser 2006 developed two types of methodological filters: Precision Terms Filters and Specificity Terms Filters (Table 2). The study analysed performance of six methodological filters in total i.e. of Precision Terms Filters and Specificity Terms Filters within MEDLINE, Embase, and combined MEDLINE/Embase database searches. Furlan 2006 reported on four types of search filters: Fixed method A, Fixed method B, Progressive method (controlled vocabulary (CV)), and Progressive method (controlled vocabulary or text words (CV or TW)) (Table 2). This study reported on performance of 12 filters in total, i.e. performance of these four types of methodological filters as part of a MEDLINE, Embase, and combined MEDLINE/Embase database searches, respectively. In summary, a total of 18 methodological filters were developed, six filters for MEDLINE, six for Embase, and six for combined MEDLINE/Embase. External validation was only performed on the Specificity Terms Filters from Fraser 2006.

Excluded studies

Twenty-three records were excluded after full-text screening (Characteristics of excluded studies). Eighteen studies were excluded because they did not explicitly focus on search strategies for retrieval of observational studies, their reference standard was developed using handsearching, or both. Two studies were excluded because they did not evaluate a search strategy (Grimshaw 2007; Mowatt 1999), and three were excluded because their reference standard included other study designs (Waffenschmidt 2017; Wieland 2002; Wieland 2005).

Development systematic reviews and identification of reference standards

The two eligible studies reported on the use of five development systematic reviews, i.e. systematic reviews that provided a reference standard. Fraser 2006 employed one development systematic review assessing the effectiveness and safety of laser in-situ keratomileusis (LASIK). Furlan 2006 used four development systematic reviews evaluating the use of local versus general anaesthesia for carotid endarterectomy (Endarterectomy), male circumcision for prevention of heterosexual acquisition of HIV in men (Circumcision), methadone maintenance at different dosages for opioid dependence (Methadone), and quality of life after rectal resection for cancer (QofL) (Characteristics of included studies).

Both studies (Fraser 2006; Furlan 2006) evaluated the performance of search filters that retrieved articles in MEDLINE and Embase via the Ovid search interface. They adopted the relative recall approach, which uses included systematic reviews or articles to compile their reference standards (Beynon 2013; Glanville 2009; Sampson 2006). Fraser 2006 used one development systematic review with a topic-only search strategy to identify a reference standard set. In this study, 217 articles (published in 2000 to 2004) constituted a reference standard for combined MEDLINE/Embase databases. Fraser 2006 also included additional observational studies retrieved during the development systematic review screening process which were not eligible for inclusion in the final review. Of the total 217 articles, 206 were indexed in MEDLINE and served as the reference standard set for MEDLINE, and the 191 indexed in Embase served as the reference standard set for Embase (Table 3). Of the reference standard, 83.9% of the articles were case series and 56.6% were prospective, observational studies.

Furlan 2006 used the comparative non-randomised studies (cNRSs) included in each of the four systematic reviews (Endarterectomy, Circumcision, Methadone and QofL) as the reference standard for that topic. For the topic of Endarterectomy, 39 cRNSs served as the reference standard for the combined MEDLINE/Embase database; 34 cNRSs indexed in MEDLINE served as the reference standard for MEDLINE, and 37 cNRSs served as the reference standard for Embase. Similarly, for the topic of Cricumcision, 29 cNRSs included in the systematic review served as the reference standard for the combined MEDLINE/Embase database and for MEDLINE; 27 cNRSs indexed in Embase served as the reference standard for Embase. For the topic of Methadone, 10 cNRSs were used as reference standard for the combined MEDLINE/Embase database, for MEDLINE, and for Embase. For the topic of QofL, 11 cNRCs served as the reference standard for the combined MEDLINE/Embase database, and 10 cNRSs for MEDLINE and for Embase (Table 3). In summary, 89 cNRSs reported in four systematic reviews (Endarterectomy, Circumcision, Methadone and QofL) constituted four reference standard sets. The included cNRSs in these four systematic reviews were a mixture of prospective and retrospective cohort studies, case-control studies, and cross-sectional studies.

Topic-only search strategies in Furlan 2006 were at least in part developed retrospectively based on the studies included in the four development systematic reviews, which employed different search strategies. The authors reported that they either aimed to use the search strategies from the original reviews where possible or that various combinations of terms were tried and the search strategy that yielded the highest sensitivity was used as a topic-only search strategy.The topic-only search strategies contained only terms related to the population, intervention, or outcomes, but not the terms related to study design. Topic-only search strategies in Fraser 2006 were prospectively developed as they were used in the development systematic review to retrieve studies.

Development of methodological filters

Both studies (Fraser 2006; Furlan 2006) used reference standard records to construct the methodological filters and combined the search terms (using Boolean logic) (Lefebvre 2011) in a way that was likely to retrieve the studies of interest. Neither study adopted published search strategies nor asked experts for suggestion of relevant terms. Fraser 2006 extracted terms to form the reference standard from one development systematic review, LASIK. For their MEDLINE filter, 37 candidate terms were extracted from 206 articles indexed in MEDLINE, including controlled vocabulary terms (i.e. MeSH terms and publication type field terms) and text words from the titles and abstracts. For their Embase filter, 35 candidate terms

were identified from 191 records indexed in Embase, including controlled thesaurus terms (i.e. EMTREE terms) and text words from the titles and abstracts. The candidate terms were tested in reverse order of precision or specificity to see if removal affected sensitivity and re-instated if they did to form "Precision Terms Filters" and "Specificity Terms Filters" (Fraser 2006).

Furlan 2006 reported on two different approaches for developing methodological filters: fixed method and progressive method. The progressive method was review specific and resulted in different filters being developed and tested for each review; whereas in the fixed method approach, the same filters were used across all reviews.The progressive method was based on random selection of eligible studies from the development systematic reviews and retrieval of study design related controlled vocabulary terms, and repeating the steps above until a minimum of four controlled vocabulary terms (called Progressive method (CV)) or both controlled vocabulary and text words from titles and abstracts obtained (called Progressive method (CV or TW)). Fixed methods used a fixed set of terms associated with study design, based on either controlled vocabulary terms alone (called Fixed method A) or both controlled vocabulary terms and text-words that were included in at least half of studies in all four reference standards (called Fixed method B).

Description of validation systematic reviews

Two additional systematic reviews, a systematic review of photorefractive keratomileusis (PRK) for myopia and a systematic review of electrosurgery for tonsillectomy (Tonsillectomy), served to validate methodological filters externally in one study (Fraser 2006). The original MEDLINE and Embase strategies used to find the evidence for these reviews were topic-only search strategies and did not include a methodological filter. The publications searched for PRK were published from 2000 to 2004, and those for tonsillectomy were published from 1990 to 2004. The total number of articles within the validation standards was 39 for the PRK review and 30 for the Tonsillectomy review (Table 3). Fraser 2006 reports the proportion of the validation reference standard that was retrieved and the reduction in number of retrieved records, respectively. This is to acknowledge the limitations of the relative recall approach used in validation and the potential for studies to be "missing" from the systematic reviews-derived validation set (Characteristics of included studies).

Risk of bias in included studies

Rather than assessing the risk of bias, we assessed (and report below) the quality of the methodological filters using a recognised appraisal checklist for search filters (Glanville 2009).

Quality of methodological filters in the included studies

Using the Information Specialists' Sub-Group (ISSG) Search Filter Appraisal Checklist (Glanville 2009), we evaluated the quality of the methodological filters in the two included studies (Fraser 2006; Furlan 2006). The assessment findings are presented in Table 4. Due to the overlap in information between Table 4 and Description of studies section above, we only present focused information in this section.

The methodological filters in Fraser 2006 were developed to maximise the sensitivity, while reducing the number of irrelevant recodes as much as possible, and the filters in Furlan 2006 were developed with the aim of achieving a balance of sensitivity and precision.

The reference standards from both included studies comprised of studies published before 2006. This might have had an impact on the relevance of the findings because database indexing and study labelling may have changed since then, e.g. MEDLINE updates its index annually. The number of studies included in the reference standards in Furlan 2006 was below the suggested 100 articles (Jenkins 2004). These 89 reference standard studies were heterogeneous, consisting of prospective and retrospective cohort studies, case-control studies, and cross-sectional studies. Conversely, the reference standard set in Fraser 2006 was larger (around 200 articles) and more homogenous, stemming from only one development systematic review, and 83.9% of the set were case series. Both studies used reference standards to develop methodological filters, which might have led to biased exaggeration of the methodological filters' performance (Beynon 2013).

Topic-only search strategies used in Furlan 2006 were different from the search strategy used in the development systematic reviews and were developed retrospectively. While most had a sensitivity of 100%, one topic-only search strategy achieved sensitivity of only 85% in MEDLINE, omitting five studies due to a missing topic-related search term. This might have affected the sensitivity of the pertinent methodological filter. However, when using the combined MEDLINE/Embase topic-only search strategies, all eligible studies were retrieved (sensitivity of 100%). In addition, because the reference standard in these reviews may have been retrieved using study design-related terms, this could have introduced bias and lead to overestimation of the filter performance.

Fraser 2006 reported additional performance measures: specificity, proportion of the reference standard retrieved and reduction in retrieval. Search filters were not compared with other relevant published filters. Study limitations presented in Fraser 2006 focused on common relative recall approach-related issues, i.e. potential omission of eligible studies from reference-standards derived from systematic reviews. Furlan 2006 did not include a comment on potential limitations.

Specificity Terms Filters from Fraser 2006 were validated externally using two additional systematic reviews. The validation standards were pragmatically derived and included all non-randomised studies that met the review inclusion criteria that were retrieved by the original topic search. As the validation systematic review may have omitted studies eligible for inclusion in the reference standard, the researchers measured the proportion of the validation reference standard that was retrieved and the reduction in number of retrieved records instead of sensitivity and precision, respectively. The numbers of eligible studies included in the two validation systematic reviews were small: 39 studies and 30 studies, respectively.

Effect of methods

The sensitivity and precision of the methodological filters developed from the two eligible studies (Fraser 2006; Furlan 2006) were extracted, calculated, and tabulated in Table 5.

Perfomance of MEDLINE filters in development systematic reviews

Sensitivity of the six MEDLINE methodological filters developed from five systematic reviews ranged from 48% to 100% with a mean of 80.94% (SD = 14.59%, or median 80%). Precision ranged from 0.12% to 21.10% with a mean of 3.55% (SD = 6.54%, or median 0.92%).

Sensitivity of the methodological filters developed in Fraser 2006 was 99.5% (PrecisionTerms Filters and Specificity Terms Filters). Precision was 20.9 % (PrecisionTerms Filters) and 21.10% (Specificity Terms Filters), respectively. In Furlan 2006, the sensitivity of the search strategies ranged from 48% (Fixed method A for the topic of Circumcision) to 100% (Progressive Methods (CV) and Progressive Methods (CV or TW) for the topic of Circumcision). Precision ranged from 0.12% (Fixed Method B for the topic of Circumcision) to 4.47% (Fixed Method A for the topic of Endarterectomy).

In summary, Specificity Terms Filters for MEDLINE from Fraser 2006 demonstrated better overall performance in terms of sensitivity and precision than PrecisionTerms Filters.

Perfomance of Embase filters in development systematic reviews

Sensitivity of six Embase filters developed from the five systematic reviews ranged from 81% to 100% with a mean of 93.67% (SD = 5.59%, or median 93%). Precision ranged from 0.08% to 18.8% with a mean of 2.50% (SD = 5.78%, or median 0.85%).

In Fraser 2006, sensitivity for both search strategies was 100%. Precision was 17.9% (Precision Terms Filters) and 18.8% (Specificity Terms Filters). In Furlan 2006, sensitivity ranged from 81% (Fixed method A for Endarterectomy topic) to 100% (Fixed method B, Progressive Methods (CV) for QofL topic, and Progressive Methods (CV or TW) for both QofL and Circumcision topics). Precision ranged from 0.08% (Fixed method B and Progressive method (CV) for Circumcision topic) to 1.04% (Progressive method (CV) for Endarterectomy topic).

In summary, Embase filters in Fraser 2006 showed the highest sensitivity and precision, and those in Furlan 2006 demonstrated the lowest sensitivity and precision. In comparisons to MEDLINE search strategies, Embase search strategies showed higher sensitivity and precision.

Perfomance of combined MEDLINE/Embase filters in development systematic reviews

Similarly, six combined MEDLINE/Embase filters were developed from five systematic reviews. Sensitivity ranged from 76% to 100% with a mean of 94.25% (SD = 7.32%, or median 97.5%). Precision ranged from 0.10% to 17.1% with a mean of 1.64% (SD = 4.17%, or median 0.66%).

In Fraser 2006, sensitivity was 100% for both types of filters and precision was 16.7% (Precision Terms Filters) and 17.1% (Specificity Terms Filters), respectively. In Furlan 2006, sensitivity ranged from 76% (Fixed method A for Circumcision topic) to 100% (several Fixed Method B and Progressive method search strategies). Precision ranged from 0.10% (Fixed method A for Methadone) to 1.13% (Progressive method (CV) for Endarterectomy). In summary, similarly to other filters, the combined MEDLINE/Embase

filters from Fraser 2006 demonstrated the highest sensitivity and precision (Specificity Terms Filters), and the filters from Furlan 2006 showed the lowest sensitivity and precision (Fixed method A).

Overall, Specificity Terms Filters from Fraser 2006 demonstrated the highest precision (21.1% for MEDLINE, 18.8% for Embase, and 17.1% for combined MEDLINE/Embase) with very high sensitivity (99.5% to 100%).

Perfomance of methodological filters in validation systematic reviews

In Fraser 2006, Specificity Terms Filters, were validated externally by using two independent additional validation systematic reviews, reporting 85.2 % to 100% of reference standard retrieved and 21.9% to 39.6% reduction in retrievals.

Perfomance and frequency of individual search term used in the methodological filters

Fraser 2006 reported significant variation in the performance of the employed search terms. Sensitivity varied from 1.0% to 65.5%. The most sensitive terms were (compare\$ or compara \$).mp (65.5%); (postoperat\$ or post operat\$).mp and (preoperat\$ or pre operat\$).mp (63.1% and 54.4%, respectively); (preoperat\$ or pre operat\$).tw and (postoperat\$ or post operat\$).tw (53.4.0% and 51.0%, respectively); and the MeSH term Comparative studies (51.5%). Precision varied from 1.3% to 41.7% with highest values for cohort.tw (41.7%), cohort.mp (37.9%) and (non random\$ or nonrandom\$).tw (37.2%). The specificity, was overall high (54.2% to 99.5%), with the lowest scores found for the terms (postoperat\$ or post operat\$).mp (54.2%), (postoperat\$ or post operat\$).tw (68.1%) and evaluat\$.tw (70.6%).

In Furlan 2006, four controlled vocabulary terms in MEDLINE and four in Embase were common to all topic areas: Cohort studies, Comparative study, Follow-up studies, and Prospective studies in MEDLINE and Clinical article, Controlled study, Major clinical study, and Prospective study in Embase. Only two text words, "compared" and "multivariate", were common to all topics. The following controlled vocabulary terms had frequencies of 50% or higher: Comparative study and Risk Factors in MEDLINE and Controlled study and Major clinical study in Embase. There were two text words with frequencies of 50% or higher: "cohort" and "groups".

DISCUSSION

Summary of main results

We identified two eligible studies (Fraser 2006; Furlan 2006) reporting on 18 methodological filters, including six MEDLINE filters, six Embase filters and six combined MEDLINE/Embase filters. All methodological filters were developed using terms from their reference standard records. Fraser 2006 reported on six filters encompassing two types of filters for MEDLINE, for Embase, and for combined MEDLINE/Embase. Furlan 2006 reported on 12 filters, which were developed using four different approaches for MEDLINE, for Embase, and for combined MEDLINE/Embase. The reference standard in Fraser 2006 consisted of 217 articles from one systematic review, while in Furlan 2006 it consisted of 89 comparative non-randomised studies (cNRSs) from four systematic reviews. Six methodological filters developed from Fraser 2006 reported sensitivity of 99.5% to 100% and precision of 16.7% to 21.1%. Twelve filters from Furlan 2006 reported lower sensitivity

(≥ 90%) and precision (0.09% to 4.47%) compared to the search filters in Fraser 2006. Only the Specificity Terms Filters from Fraser 2006 went through external validation using two independent systematic reviews with sensitivity of 85.2% to 100%.

Overall completeness and applicability of evidence

We found only two eligible studies (Fraser 2006; Furlan 2006) that focused on different types of observational studies. Fraser 2006 targeted observational studies of surgical interventions and the methodological filters included terms such as (preoperat$ or pre operat$).mp. The reference standard included mostly case-series studies (nearly 90%). Conversely, Furlan 2006 focused on cNRSs which encompassed cohort, case-control and cross-sectional studies. Furthermore, the two included studies used different approaches to develop their methodological filters. The Fixed method used in Furlan 2006 corresponded to the approach used in Fraser 2006. Furlan 2006 also proposed a Progressive method where study design-related terms, both controlled vocabulary terms and text words, were randomly collated from the reference standard records, meaning that the progressively developed methodological filter would differ in different systematic reviews.

External validation was performed on only one type of filter in Fraser 2006. Fraser 2006 included only one development systematic review, while Furlan 2006 included four different systematic reviews, which may increase the generalizability of the latter's findings. The date of the reference standard and evaluation performance (before 2006) might also have had an impact on the findings and limit the generalizability of the findings. Both studies used the Ovid interface and it is not clear how these filters would perform if used in PubMed and other interfaces.

Our searches have also revealed a number of search strategies for observational studies which are currently in use, but which lack publicly accessible and peer-reviewed performance assessment and validation (BMJ Best Practice 2018; ISSG 2017; SIGN 2018; UTHealth 2018).

Quality of the evidence

Fraser 2006 included more than 200 observational studies retrieved by topic-only search strategies in one systematic review focused on surgical interventions to construct the reference standard. Furlan 2006 used four systematic reviews (Endarterectomy, Circumcision, Methadone, and QofL) to constitute four sets of reference standards, which potentially increased generalizability. However, the total number of included studies across these four reference standards was under 100, the sample size required to estimate the confidence interval around the sensitivity estimate, which might have reduced the accuracy of their findings (Sampson 2006). Topic-only search strategies in Furlan 2006 developed at least in part retrospectively, which led to a sensitivity of 85% for one topic-only search strategy because it failed to retrieve all relevant studies, might also have influenced the performance of pertinent methodological filters. In both Fraser 2006 and Furlan 2006, search strategies were derived from reference standards, which might have led to an overestimation of the performance of the search strategies. External validation was performed only for the Specificity Terms Filter in Fraser 2006. The scope of the identified search strategies in Fraser 2006 and Furlan 2006 differed: Fraser 2006 retrieved mostly case studies and Furlan 2006 focused on cNRSs, although they overlapped partially. The dates

of the searches used to compile the reference standard in Furlan 2006 were older than those in Fraser 2006, which might have contributed to lower sensitivity, due to poorer database indexing of observational study at that time. Finally, the use of the progressive method in Furlan 2006 implies that researchers need to devise their own filter for each review, which may prove a challenge for those with limited research experience.

Potential biases in the review process

As in most systematic reviews, it is possible that we missed some eligible studies. However, we have performed extensive and sensitive searches of six electronic databases and checked the reference lists of published filters and filter evaluations. Both eligible studies used a relative recall approach, i.e. a reference standard which was taken from a published systematic review. The use of the relative recall approach is economic because it allows for a quick inclusion of studies from a range of journals and publication years and may be more generalisable than a comparison to a reference standard retrieved by handsearching (Sampson 2006). However, it can have limitations, such as being highly specific and confounded by the topic-only searches and time when the searches were performed (Lefebvre 2017).

Agreements and disagreements with other studies or reviews

We have not encountered another review assessing the performance of search strategies for the retrieval of observational studies in MEDLINE and Embase. We found several studies focusing on evaluating search strategies for other types of studies and using handsearching reference standards. In general, these studies had lower precision and at times also lower sensitivity compared to Fraser 2006. There are studies on the development of search strategies as part of the McMaster Hedge's project for prognostic, causation and health services research studies using a handsearching-based reference standard. These search strategies have sensitivity of 80% to 100% and precision generally below 10% (Haynes 1994; Haynes 2004; Haynes 2005A; Haynes 2005B; McMaster Hedges Team 2017a; McMaster Hedges Team 2017b). A conference abstract evaluating performance of a Cochrane MEDLINE search strategy for identifying interrupted time series (ITS) and controlled before-and-after (CBA) designs using a handsearching reference standard, reported that the sensitivity was 70.2% for CBAs and 55.5% for ITS with a precision of 7.0% (Fraser 1998).

We also found several reviews collating studies evaluating search strategies for study designs other than observational studies. A review of studies evaluating search strategies for identifying information on adverse effects reported on highly sensitive search strategies capable of retrieving over 95%, but with low precision of 0.8% to 2.8%. The limitations reported in that review correspond to those in our review and include a relatively small number of records, absence of a validation set of records for testing, and limited evaluation of precision (Golder 2009). A study evaluating diverse search strategies for retrieval of epidemiological studies using a combination of relative recall and handsearching, reported on the performance properties of 27 filters with sensitivity ranging from 24.38% to 95.94%, and precision between 0.36% and 3.07% (Waffenschmidt 2017) A Cochrane Review of search strategies to identify diagnostic accuracy studies in MEDLINE and Embase reported that none of the included 70 methodological filters

had sufficiently high sensitivity for systematic reviews, with a reasonable degree of precision (Beynon 2013). It also reported that performance levels of the search filters typically declined when the filters were validated in evaluation studies.

In comparison to the body of research discussed above, our review presents studies with overall higher sensitivity and precision and employs an alternative to the use of handsearching for the reference standard, i.e. a relative recall approach. However, validation was only performed for one of the 18 filters presented in the included studies for this review, which limits the reliability and generalizability of our findings.

AUTHORS' CONCLUSIONS

Implication for methodological research

There is a need to re-evaluate the performance of the reported methodological filters because the two included studies were published more than 10 years ago and the reporting and database indexing of observational studies may have changed since then. Evaluations should be external and robust, using large validation sets from a variety of topics. Future research should aim to evaluate existing search strategies that are currently in use but which lack rigorous evaluation (BMJ Best Practice 2018; ISSG 2017; UTHealth 2018). There is also scope for the development and validation of new search strategies. Future studies should aim to use large reference standards from a range of developmental systematic reviews. They should also include external validation to help determine the generalizability of the methodological filters.

REFERENCES

References to studies included in this review

Fraser 2006 *{published and unpublished data}*

Fraser C, Burr J. Identifying observational studies of surgical interventions in Medline and Embase: validation of a search filter [abstract]. IV Annual Meeting of Health Technology Assessment International. 2007 June:297.

Fraser C, Murray A, Burr J. Identifying non-randomised studies in MEDLINE and EMBASE [abstract]. *Italian Journal of Public Health [Internet]* 2005;**2**(Suppl 1):286.

* Fraser C, Murray A, Burr J. Identifying observational studies of surgical interventions in MEDLINE and EMBASE. *BMC Medical Research Methodology* 2006;**6**:41-9.

Furlan 2006 *{published and unpublished data}*

Furlan AD, Bombardier C. Controlled vocabulary associated with non-randomized studies of interventions for low back pain [abstract]. XI Cochrane Colloquium: Evidence, Health Care and Culture. 2003 Oct:69.

* Furlan AD, Irvin E, Bombardier C. Limited search strategies were effective in finding relevant non-randomized studies. *Journal of Clinical Epidemiology* 2006;**59**(12):1303-11.

References to studies excluded from this review

Brown 2006 *{published data only}*

Brown ML. New search strategies optimize MEDLINE retrieval of sound studies on treatment or prevention of health disorders. A review of: Haynes, R. Brian, K. Ann McKibbon, Nancy L. Wilczynski, Stephen D. Walter, and Stephen R. Were. "Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey." BMJ 330.7501 (21 May 2005): 1179. *Evidence Based Library and Information Practice* 2006;**1**(2):23-6.

Fraser 1998 *{published data only (unpublished sought but not used)}*

Fraser C, Thomson MA. Identifying non-randomised studies in Medline [abstract]. Sixth International Cochrane Colloquium; Baltimore, MD, USA. 1998 Oct 22-26.

Geersing 2012 *{published data only}*

Geersing GJ, Bouwmeester W, Zuithoff P, Spijker R, Leeflang M, Moons KG. Search filters for finding prognostic and diagnostic prediction studies in Medline to enhance systematic reviews. PLOS One 2012; Vol. 7, issue 2:e32844.

Grimshaw 2007 *{published data only}*

Grimshaw J, McGowan J, Salzwedel D. Update of the EPOC methodological search filter [abstract]. XV Cochrane Colloquium. 2007 Oct 23-27:77-78.

Haynes 1994 *{published data only}*

Haynes RB, Wilczynski N, McKibbon KA, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. *Journal of the American Medical Informatics Association* 1994;**1**(6):447-58.

Haynes 2004 *{published data only}*

Haynes RB, Wilczynski NL. Optimal search strategies for retrieving scientifically strong studies of diagnosis from Medline: analytical survey. *BMJ* 2004;**328**(7447):1040-5.

Haynes 2005A *{published data only}*

Haynes RB, Kastner M, Wilczynski NL. Developing optimal search strategies for detecting clinically sound and relevant causation studies in EMBASE. *BMC Medical Informatics and Decision Making* 2005;**5**(1):8-14.

Haynes 2005B *{published data only}*

Haynes RB, McKibbon KA, Wilczynski NL, Walter SD, Werre SR. Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. *BMJ* 2005;**330**(7501):1179-84.

Holland 2005 *{published data only}*

Holland JL, Wilczynski NL, Haynes RB, Hedges T. Optimal search strategies for identifying sound clinical prediction studies in EMBASE. *BMC Medical Informatics and Decision Making* 2005;**5**:11-6.

Montori 2005 *{published data only}*

Montori VM, Wilczynski NL, Morgan D, Haynes RB, and for the Hedges Team. Optimal search strategies for retrieving systematic reviews from Medline: analytical survey. *BMJ* 2005;**330**(7482):68-73.

Mowatt 1999 *{published data only}*

Mowatt G, Hero L, Campbell M, Fraser C, Grilli R, Grimshaw J, et al. Use of non-randomised evidence in Cochrane Effective Practice and Organisation of Care (EPOC) group's reviews [abstract]. 7th Annual Cochrane Colloquium; Rome Italy [Internet]. 1999 Oct 5-9.

Waffenschmidt 2017 *{published data only}*

Waffenschmidt S, Hermanns T, Gerber-Grote A, Mostardt S. No suitable precise or optimized epidemiologic search filters were available for bibliographic databases. *Journal of Clinical Epidemiology* 2017;**82**:112-8.

Wieland 2002 *{published data only}*

Wieland S, Brodney S, Dickersin K. Designing an efficient and precise search strategy for observational studies. 10th Cochrane Colloquium Abstracts, Stavanger. 2002 31 July - 3 August:23-24.

Wieland 2005 *{published data only}*

Wieland S, Dickersin K. Selective exposure reporting and Medline indexing limited the search sensitivity for observational studies of the adverse effects of oral contraceptives. *Journal of Clinical Epidemiology* 2005;**58**(6):560-7.

Wilczynski 1993 *{published data only}*

Wilczynski NL, Walker CJ, McKibbon KA, Haynes RB. Assessment of methodologic search filters in MEDLINE. *Proceedings of the*

Annual Symposium on Computer Application in Medical Care 1993;**NA**:601-5.

Wilczynski 2003 *{published data only}*

Wilczynski NL, Haynes RB, Hedges T. [Developing optimal search strategies for detecting clinically sound causation studies in MEDLINE]. AMIA 2003 Symposium Proceedings. 2003:719-23.

Wilczynski 2004A *{published data only}*

Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound prognostic studies in MEDLINE: an analytic survey. *BMC Medicine* 2004;**2**:23-7.

Wilczynski 2004B *{published data only}*

Wilczynski N, Haynes B. Optimal search strategies for retrieving scientifically strong studies of treatment and diagnosis from MEDLINE: an analytical survey [abstract]. 12th Cochrane Colloquium: Bridging the Gaps; Ottawa, Ontario, Canada [Internet]. 2004 Oct 2-6:70-1.

Wilczynski 2004C *{published data only}*

Wilczynski NL, Haynes RB, Lavis JN, Ramkissoonsingh R, Arnold-Oatley AE. Optimal search strategies for detecting health services research studies in MEDLINE. *Canadian Medical Association Journal* 2004;**171**(10):1179-85.

Wilczynski 2005A *{published data only}*

Wilczynski NL, Haynes RB. Optimal search strategies for detecting clinically sound prognostic studies in EMBASE: an analytic survey. *Journal of the American Medical Informatics Association* 2005;**12**(4):481-5.

Wilczynski 2005B *{published data only}*

Wilczynski NL, Haynes RB. EMBASE search strategies for identifying methodologically sound diagnostic studies for use by clinicians and researchers. *BMC Medicine* 2005;**3**:7-12.

Wilczynski 2010 *{published data only}*

Wilczynski NL, Haynes RB, for the QI Hedges Team. Optimal search filters for detecting quality improvement studies in Medline. *Quality & Safety in Health Care* 2010;**19**(6):e31-5.

Wong 2003 *{published data only}*

Wong SSL, Wilczynski NL, Haynes RB, Ramkissoonsingh R. Developing optimal search strategies for detecting sound clinical prediction studies in MEDLINE. *AMIA Annual Symposium Proceedings* 2003;**2003**:728-32.

Additional references

Beynon 2013

Beynon R, Leeflang MM, McDonald S, Eisinga A, Mitchell RL, Whiting P, et al. Search strategies to identify diagnostic accuracy studies in MEDLINE and EMBASE. *Cochrane Database of Systematic Reviews* 2013, Issue 9. [DOI: 10.1002/14651858.MR000022.pub3]

BMJ Best Practice 2018

BMJ Best Practice. Study design filter. https://bestpractice.bmj.com/info/toolkit/learn-ebm/study-design-search-filters/ 2018.

Doust 2005

Doust JA, Pietrzak E, Sanders S, Glasziou PP. Identifying studies for systematic reviews of diagnostic tests was difficult due to the poor sensitivity and precision of methodologic filters and the lack of information in the abstract. *Journal of Clinical Epidemiology* 2005;**58**:444-9.

Durão 2015

Durão S, Kredo T, Volmink J. Validation of a search strategy to identify nutrition trials in PubMed using the relative recall method. *Journal of Clinical Epidemiology* 2015;**68**(6):610-6.

Edwards 2002

Edwards P, Clarke M, DiGuiseppi C, Pratap S, Roberts I, Wentz R. Identification of randomized controlled trials in systematic reviews: accuracy and reliability of screening records. *Statistics in Medicine* 2002;**21**(11):1635-40.

EndNote X7 [Computer program]

Thomson Reuters. EndNote X7. New York, USA: Thomson Reuters Corporation, 2015.

Glanville 2008

Glanville J, Bayliss S, Booth A, Dundar Y, Fernandes H, Fleeman ND, et al. So many filters, so little time: the development of a search filter appraisal checklist. *Journal of the Medical Library Association* 2008;**96**(4):356-61.

Glanville 2009

Glanville J, Kaunelis D, Mensinkai S. How well do search filters perform in identifying economic evaluations in MEDLINE and EMBASE. *International Journal of Technology Assessment in Health Care* 2009;**25**(4):522-9.

Glasziou 2001

Glasziou P, Irwig L, Bain C, Colditz G. Systematic Reviews in Health Care: a Practical Guide. Cambridge: Cambridge University Press, 2001.

Golder 2009

Golder S, Loke YK. Search strategies to identify information on adverse effects: a systematic review. *Journal of the Medical Library Association* 2009;**97**(2):84-92.

Golder 2018

Golder S, Wright K, Loke YK. The development of search filters for adverse effects of surgical interventions in Medline and Embase. *Health Information and Libraries Journal* 2018;**35**(2):121-9.

Gough 2012a

Gough D, Oliver S, Thomas J. An Introduction to Systematic Reviews. London: SAGE, 2012.

Gough 2012b

Gough D, Thomas J, Oliver S. Clarifying differences between review designs and methods. *Systematic Reviews* 2012;**1**:28.

Hayward 1997

Hayward S, Brunton G, Thomas K, Ciliska D. Searching for the evidence:source, time and yield. 2nd International Conference Scientific Basis of Health Services & 5th Annual Cochrane Colloquium. 1997.

Higgins 2011

Higgins JP, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Horsley 2011

Horsley T, Dingwall O, Sampson M. Checking reference lists to find additional studies for systematic reviews. *Cochrane Database of Systematic Reviews* 2011, Issue 8. [DOI: 10.1002/14651858.MR000026.pub2]

ISSG 2017

Glanville J, Lefebvre C, Wright K, editors. ISSG Search Filter Resource [Internet]. The InterTASC Information Specialists' Sub-Group; 2008 [updated 2017 Aug 14]. https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home 2017.

Jenkins 2004

Jenkins M. Evaluation of methodological search filters-a review. *Health Information and Libraries Journal* 2004;**21**(3):148-63.

Lefebvre 2011

Lefebvre C, Manheimer E, Glanville J. Chapter 6: Searching for studies. In: Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration. Available from handbook.cochrane.org.

Lefebvre 2013

Lefebvre C, Glanville J, Wieland LS, Coles B, Weightman AL. Methodological developments in searching for studies for systematic reviews: past, present and future?. *Systematic Reviews* 2013;**2**:78.

Lefebvre 2017

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 Technology Assessment 2017; Vol. 21, issue 69.

Lemeshow 2005

Lemeshow AR, Blum RE, Berlin JA, Stoto MA, Colditz GA. Searching one or two databases was insufficient for meta-analysis of observational studies. *Journal of Clinical Epidemiology* 2005;**58**:867-73.

Ligthelm 2007

Ligthelm RJ, Borzi V, Gumprecht J, Kawamori R, Wenying Y, Valensi P. Importance of observational studies in clinical practice. *Clinical Therapeutics* 2007;**29**:1284-92.

Littleton 2004

Littleton D, Marsalis S, Bliss DZ. Searching the literature by design. *Western Journal of Nursing Research* 2004;**26**(8):891-908.

Manchikanti 2009

Manchikanti L, Datta S, Smith HS, Hirsch JA. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: part 6. Systematic reviews and meta analyses of observational studies. *Pain Physician* 2009;**12**:819-50.

McMaster Hedges Team 2017a

McMaster Hedges Team. Search Filters for MEDLINE in Ovid Syntax and the PubMed translation. https://hiru.mcmaster.ca/hiru/HIRU_Hedges_MEDLINE_Strategies.aspx 2017.

McMaster Hedges Team 2017b

McMaster Hedges Team. Search Strategies for EMBASE in Ovid Syntax. https://hiru.mcmaster.ca/hiru/HIRU_Hedges_EMBASE_Strategies.aspx 2017.

Robinson 2002

Robinson KA, Dickersin K. Development of a highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed. *International Journal of Epidemiology* 2002;**31**(1):150-3.

Sampson 2006

Sampson M, Zhang L, Morrison A, Barrowman NJ, Clifford TJ, Platt RW, et al. An alternative to the hand searching gold standard: validating methodological search filters using relative recall. *BMC Medical Research Methodology* 2006;**6**:33.

SIGN 2018

Scottish Intercollegiate Guidelines Network. Search filter for observational studies. https://www.sign.ac.uk/search-filters.html.

UTHealth 2018

University of Texas School of Public Health. Search Filters for Various Databases: Ovid PsycINFO. http://libguides.sph.uth.tmc.edu/search_filters/ovid_psycinfo_filters 2018.

Vandenbroucke 2014

Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *International Journal of Surgery* 2014;**12**(12):1500-24.

Viswanathan 2013

Viswanathan M, Berkman ND, Dryden DM, Hartling L. Assessing Risk of Bias and Confounding in Observational Studies of Interventions or Exposures: Further Development of the RTI Item Bank. Methods Research Report.. Vol. **AHRQ Publication**

No. 13-EHC106-EF, Rockville, MD: Agency for Healthcare Research and Quality, 2013.

Wilczynski 2007

Wilczynski NL, Haynes RB. EMBASE search strategies achieved high sensitivity and specificity for retrieving methodologically sound systematic reviews. *Journal of Clinical Epidemiology* 2007;**60**:29-33.

Young 2011

Young T, Hopewell S. Methods for obtaining unpublished data. *Cochrane Database of Systematic Reviews* 2011, Issue 11. [DOI: 10.1002/14651858.MR000027.pub2]

Yousefi-Nooraie 2013

Yousefi-Nooraie R, Irani S, Mortaz-Hedjri S, Shakiba B. Comparison of the efficacy of three PubMed search

filters in finding randomized controlled trials to answer clinical questions. *Journal of Evaluation in Clinical Practice* 2013;**19**(5):723-6.

References to other published versions of this review

Marcano Belisario 2013

Marcano Belisario JS, Tudor Car L, Reeves TJ, Gunn LH, Car J. Search strategies to identify observational studies in MEDLINE and EMBASE. *Cochrane Database of Systematic Reviews* 2013, Issue 12. [DOI: 10.1002/14651858.MR000041]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Fraser 2006

Methods	**Method of identification of reference standards:** articles included in the development systematic reviews
	Method of deriving methodological filter terms: analysis of reference standard: the terms related to study design from reference standard were assessed and identified as candidate terms for the MEDLINE and Embase methodology filters. The candidate terms were tested in reverse order of precision or specificity to see if removal affected sensitivity and re-instated if they did. Therefore, the reference standard consisted of the term with highest precision and specificity while sensitivity was maximised.
Data	**Reference standard year:** development set 2000-2004; validation set 2000-2004 and 1990-2004
	Number of reference standards: 217 articles for combined MEDLINE/Embase in development systematic reviews
	Number of validation standard records: 39 and 30 for combined MEDLINE/Embase in validation systematic reviews, PRK and tonsillectomy, respectively
Comparisons	**Reference standard also contained non-gold standard records:** yes. There was no gold standard
	Description of non-gold standard records if used in reference standard: reference standard, which used articles retrieved by electronic search in the included development systematic reviews
Outcomes	**Performance of filters:** sensitivity and precision
	Number of methodological filters developed: 6: Precision Terms Filters and Specificity Terms Filters for MEDLINE, for Embase, and for combined MEDLINE/Embase, respectively
Notes	• MEDLINE and Embase development and validation systematic reviews
	• Validation standards and filters in validation systematic reviews were constructed the same way as the reference standards in development systematic reviews
	• Only Specificity Terms Filters were validated externally by two additional systematic reviews

Furlan 2006

Methods	**Method of identification of reference standards:** cNRSs included in the four development systematic reviews
	Method of deriving methodological filter terms: analysis of reference standards: two approaches (a fixed method and a progressive method) were used to develop four sets of filters for four development systematic reviews: Fixed method A, Fixed method B, Progressive method (CV) and Progressive method (CV or TW). A quote: "fixed search strategy" was a fixed set of CVs associated with study design. A progressive method (PM) for selection of study design terms was developed and tested: (a) Start with the list of included studies from each systematic review; (b) Randomly select one study from this list and identify it in the database under study; (c) Retrieve the CV (related to study design) associated with the study identified in step (b); (d) Repeat (b) and (c) until a minimum of four CVs was reached; (e) Devise a quote: ''limited search'' in which the ''topic-only'' results were limited to the search terms identified above. The same methodology was applied to identify and test search strategies limited to text words (titles and abstracts)
Data	**Years searched for the reference:** development set 1966-2005 (MEDLINE) or 1980-2005 (Embase)
	Number of reference standard for each topic:: for combined MEDLINE/Embase in development systematic reviews, 39 cNRSs in endarterectomy, 29 cNRSs in circumcision, 10 cNRSs in methadone, 11 cNRSs in QofL.
Comparisons	**Reference standard also contained non-gold standard records:** no
	Description of non-gold standard records if used in reference standard: NA
Outcomes	**Performance of filters:** Sensitivity and Precision (since PPV was calculated in the same way as Precision so that we use Precision instead of PPV)
	Number of methodological filters developed: 12: Fixed method A, Fixed method B, Progressive method((CV), and Progressive method(CV or TW) for MEDLINE, Embase, and combined MEDLINE/Embase, respectively
Notes	• A ''topic-only'' search strategy, included only terms related to the population, intervention, or outcomes, but not the terms related to study design. Though it was called "reference standard", it did not serve as the benchmark for comparison but a search strategy without methodological filter
	• Development and validation of methodological filters for identifying comparative non-randomised studies (cNRSs) in MEDLINE and Embase. Four updated systematic reviews were selected to generate and validate internally the gold standard, topic-only reference standards, and methodological filters
	• The methodological filters were not validated externally

cNRSs: comparative non-randomised studies; **CV:** controlled vocabulary; **PM:** progressive method; **PPV:** positive predictive value; **PRK:** photorefractive keratomileusis; **QoL**: quality of life; **TW:** text word

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Brown 2006	The authors searched for treatment studies.The gold standard was handsearching of journals
Fraser 1998	The study was focused on retrieval of ITS and CBAs exclusively. The gold standard was handsearching of journals
Geersing 2012	The study was focused on retrieval of prognostic and diagnostic prediction studies exclusively.The gold standard was handsearching of journals
Grimshaw 2007	The study was focused on the revision of the EPOC methodological filter used to retrieve citations from MEDLINE for retrieval of ITS and CBAs exclusively.

Study	Reason for exclusion
Haynes 1994	The study was focused on retrieval of clinically sound studies.The gold standard was handsearching of journals
Haynes 2004	The study was focused on retrieval of diagnostic accuracy studies. The gold standard was handsearching of journals
Haynes 2005A	The study was focused on retrieval of causation studies.The reference standard was handsearching of journals.
Haynes 2005B	The study was focused on retrieval of retrieval of treatment studies.The reference standard was handsearching of journals.
Holland 2005	The study focused on retrieval of clinical prediction studies. The reference standard was handsearching of journals.
Montori 2005	The study was focused on retrieval of systematic reviews. The reference standard was handsearching of journals.
Mowatt 1999	The study was focused on describing the use of non-randomised evidence in EPOC reviews.
Waffenschmidt 2017	Focused on retrieval of epidemiological studies, not explicitly observational studies. The reference standard included non-epidemiologic articles as well as studies retrieved using handsearching.
Wieland 2002	This is a conference abstract for Wieland 2005.
Wieland 2005	The reference standard included a randomised controlled trial. The search strategy does not seem to explicitly focus on observational studies.
Wilczynski 2004B	The study focused on retrieval of treatment and diagnostic studies.The reference standard was handsearching of journals.
Wilczynski 1993	The study was focused on retrieval of aetiology, prognosis, diagnosis, and prevention and treatment studies. The reference standard was handsearching of journals.
Wilczynski 2003	The study was focused on retrieval of causation studies. The reference standard was handsearching of journals.
Wilczynski 2004A	The study was focused on retrieval of prognosis studies.The reference standard was handsearching of journals.
Wilczynski 2004C	The study was focused on retrieval of health services research studies. The reference standard was handsearching of journals.
Wilczynski 2005A	The study focused on retrieval of prognostic studies. The reference standard was handsearching of journals.
Wilczynski 2005B	The study was focused on retrieval of diagnostic studies. The reference standard was handsearching of journals.
Wilczynski 2010	The study was focused on retrieval of quality improvement studies. The reference standard was handsearching of journals.
Wong 2003	The study was focused on retrieval of clinical prediction studies. The reference standard was handsearching of journals.

CBAs: controlled before-and-after trials; **EPOC:** Cochrane Effective Practice and Organisation of Care; **ITS:** interrupted time series

ADDITIONAL TABLES

Table 1. Calculation of sensitivity and precision

		Reference standard	Non-reference standard
Methodological filters	**Retrieved**	A	B
	Not retrieved	C	D

Sensitivity = A/(A+C) x100
Precision or positive predictive value (PPV) = A/(A+B) x 100

Table 2. Summary of developed methodological filters for MEDLINE, Embase, and combined MEDLINE/Embase, respectively

Study ID	Methodological filters			
	Filter 1	**Filter 2**	**Filter 3**	**Filter 4**
Fraser 2006	Precision Terms Filters	Specificity Terms Filters	NA	NA
Furlan 2006	Fixed method A	Fixed method B	Progressive method (CV)	Progressive method (CV or TW)

CV: controlled vocabulary; **TW:** text words

Table 3. Summary of reference standards

Study ID	Systematic reviews	Number of studies per reference standard		
		MEDLINE	**Embase**	**MEDLINE/Embase**
Fraser 2006	LASIK	206	191	217
	PRK	39	33	39
	Tonsillectomy	27	27	30
Furlan 2006	Endarterectomy	34	37	39
	Circumcision	29	27	29
	Methadone	10	10	10
	QofL	10	10	11

LASIK: a systematic review to assess the effectiveness and safety of laser in-situ keratomileusis
PRK: a systematic review of photorefractive keratomileusis for myopia (used for external validation)
Tonsillectomy: a systematic review of electrosurgery for tonsillectomy (used for external validation)
Endarterectomy: a systematic review of local versus general anaesthesia for carotid endarterectomy
Circumcision: a systematic review of male circumcision for prevention of heterosexual acquisition of HIV in men
Methadone: a systematic review of methadone maintenance at different dosages for opioid dependence

QofL: a systematic review of quality of life after rectal resection for cancer

Table 4. Assessment of quality of evidence according to ISSG appraisal checklist

Information and methodological issues	Fraser 2006	Furlan 2006
A. Information		
A.1. State the author's objective	Quote: "To develop effective MEDLINE and EMBASE filters, to identify non-randomised evidence for surgical interventions, to be used in conjunction (using Boolean operator AND) with a subject search strategy"	Quote: "To develop and evaluate a method to limit search strategies according to study design of comparative non-randomised studies (cNRSs)"
A.2. State the focus of the research [ ] Sensitivity-maximising [ ] Precision-maximising [ ] Specificity-maximising [ ] Balance of sensitivity and precision [ ] Others	Maximising sensitivity while consisting of terms with high specificity (Specificity Terms Filters) and with high precision (Precision Terms Filters)	Balance of sensitivity and precision.
A.3. Databases and search interfaces	MEDLINE and Embase using Ovid interface	MEDLINE and Embase using OVID interface
A.4. Describe the methodological focus of the filter	Observational studies of surgical interventions	Non-randomised comparative studies (cNRSs)
A.5. Describe any other topic that forms an additional focus of the filter	Surgical interventions	NA
B. Identification of reference standard of known relevant records		
B.1. Did the authors identify one or more reference standard?	One reference standard was identified during the screening, which consisted of 217 articles from one systematic review on effectiveness and safety of laser in-situ keratomileusis (LASIK).It also included studies retrieved during the development systematic review screening process which were not eligible for inclusion in the final review.	Four reference standards were included from four development systematic reviews assessing effectiveness of • local vs. general anaesthesia for carotid endarterectomy (Endarterectomy) (39 articles); • male circumcision for prevention of heterosexual acquisition of HIV in men (Circumcision) (29 articles); • methadone maintenance at different dosages for opioid dependence (Methadone) (10 articles); • quality of life after rectal resection for cancer review (QofL) (11 articles).
B.2. How did the authors identify the records in each reference standard?	The authors used a systematic review to assess the effectiveness and safety of LASIK with the MEDLINE and Embase search strategies that had incorporated terms that pertained only to the intervention and	The reference standards were retrieved from systematic reviews, which met the inclusion criteria:

![Cochrane Library logo] Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

Table 4. Assessment of quality of evidence according to ISSG appraisal checklist *(Continued)*

	medical conditions of interest to the review and were restricted to the publication years 2000–2004. The search strategy included only topic-related terms	• published in the Cochrane Database of Systematic Reviews (CDSR), therefore only reviews of effectiveness of health-care interventions are included; • being up-to-date (published in 2003 or later); • including a minimum of 10 cNRSs; • having at least 90% of the included studies indexed in MEDLINE or Embase; • displaying the search terms used for the topic area to allow us to devise the search strategy; • having a search strategy not limited to study design.
B.3. Report the dates of the records in each reference standard.	Articles that constituted the reference standard were retrieved from systematic reviews published in the years 2000 - 2004	The topic searches in the Cochrane systematic review to retrieve the cNRS studies were preformed between 1966 - Jul 2005 (MEDLINE) and 1980 - Jul 2005 (Embase)
B.4. What are the inclusion criteria for reference standard?	Both prospective and retrospective non-randomised comparative studies and case series constituted the reference standard	Eligible studies included prospective and retrospective non-randomised comparative studies including cohort, case-control, cross-sectional studies
B.5. Describe the size of each reference standard and the authors' justification if provided	• 206 articles for MEDLINE • 191 articles for Embase • 217 articles for combined MEDLINE/Embase	For MEDLINE: • Endarterectomy: 34 • Circumcision: 29 • Methadone: 10 • QofL:10 For Embase: • Endarterectomy: 37 • Circumcision: 27 • Methadone: 10 • QofL: 10 For combined MEDLINE/Embase: • Endarterectomy: 39 • Circumcision: 29 • Methadone: 10 • QofL: 11 Total number of studies from four reference standards was 89
B.6. Are there limitations to the reference standard?	There was only one set of reference standards from one topic and included studies published before 2005	The reference standards included a small number of studies published before 2006
B.7. How was each reference standard used?	Reference standards were used to develop search filters and assess their performance. This may have led to an overestimation of the search strategy's performance	Reference standards were used to develop topic-only search strategy, and to validate their performance. This may have lead to an overestimation of the search strategies' performance

Table 4.   Assessment of quality of evidence according to ISSG appraisal checklist *(Continued)*

B.8. Other observations.	NA	NA

C. How did the authors identify the search terms in their filters?

C.1. Adapted a published search strategy.	No	No
C.2. Asked experts for suggestions of relevant terms.	No	No
C.3. Used a database thesaurus.	Yes. • For MEDLINE filters, 37 candidate terms were identified from 206 MEDLINE records, including 11 controlled thesaurus terms (8 MeSH terms and three terms from the publication type field) • For Embase filters, out of 35 candidate terms, 10 EMTREE terms, 18 text words or phrases from the titles and abstracts and seven words or phrases that occurred in either the EMTREE terms, title or abstract fields were included	Yes • A set of controlled vocabulary and text words of study design from reference standard records were chosen for methodological filters • Controlled vocabularies were displayed as MeSH in MEDLINE and Emtree in Embase, which were database thesauruses
C.4. Performed statistical analysis of terms in a reference standard of records.	Descriptive analysis of retrieval parameters was performed for individual search terms	Descriptive analysis of retrieval parameters was performed for individual search terms
C.5. Extracted terms from the reference standard of records.	Yes. Quote from the study: "The titles, abstracts, thesaurus controlled subject headings (for MEDLINE and EMBASE) and the publication type field (for MEDLINE) were subjectively assessed by the information specialist for all the MEDLINE and EMBASE records of the reference standard. Terms that explained or gave an indication of methodology employed or systematic assessment of postoperative sequelae were identified and considered as candidate terms for the MEDLINE and EMBASE methodology filters." For example, 37 candidate terms for MEDLINE filter were identified from reference standard for MEDLINE in the development systematic review, LASIK	Yes. This study reported on two different approaches to methodological filter development: fixed and progressive method. Progressive methods based on random selection of eligible studies from the development reviews and retrieval of study design related controlled vocabulary terms, and repeating the above steps until a minimum of four controlled vocabulary terms (called Progressive method (CV)) or both controlled vocabulary and text words from titles and abstracts (called Progressive method (CV or TW)) Fixed methods used a fixed set of terms associated with study design, based on either only controlled vocabulary terms (called Fixed method A) or both controlled vocabulary and text-words that were included in at least half of studies in all four reference standard (called Fixed method B)
C.6. Extracted terms from some relevant records (but not a reference standard)	No	No
C.7. Tick all types of search terms tested. [ ] Subject headings [ ] Text words (e.g. in title, abstract)	Subject headings, Text words and Publication types field terms MEDLINE: MeSH terms, Publication type field terms, text words Embase: EMTREE terms, text words	Subject headings, Text words **Fixed method A -** Controlled Vocabulary Terms **Fixed method B -** Controlled Vocabulary Terms; Textwords

Table 4. Assessment of quality of evidence according to ISSG appraisal checklist *(Continued)*

[ ] Publication types		**Progressive method (CV) -** Controlled Vocabulary Terms
[ ] Subheadings		**Progressive method (CV or TW) -** Controlled Vocabulary Terms; Text words
[ ] Check tags		
[ ] Other, please specify		
C.8. Include the citation of any adapted strategies.	NA	Unknown

C.9. How were the (final) combination(S) of search terms selected?	• Initially all the candidate terms for each database were combined, using the Boolean operator OR, to form the separate MEDLINE and Embase filters. These were run in combination with the subject-only search strategies (using the Boolean operator AND)
	
	• Each candidate term was then tested to establish if its removal from the filter reduced overall sensitivity. If sensitivity was unaffected the term was considered redundant and was excluded from further analysis, while if sensitivity decreased the term was re-instated
	
	• To minimise the number of irrelevant records retrieved, two approaches were explored:
	
	  • The terms were tested in order of precision, beginning with the lowest so that preference was given to retaining the terms with higher precision. The resulting MEDLINE and Embase filters are referred to as the Precision Terms Filters
	
	  • The terms were tested in order of specificity, beginning with the lowest so that preference was given to retaining the terms with higher specificity. The resulting MEDLINE and Embase filters are referred to as the Specificity Terms Filters
	
	• By this process of elimination, redundant terms were removed and the combination of retained terms aimed to minimise the number of retrieved irrelevant records. six filters were thus developed: Precision Terms Filters and Specificity Terms Filters for MEDLINE, and Embase, and for both
		Fixed method A - all Controlled Vocabulary Terms associated with study design of cNRSs identified by all the topics areas
		
		Fixed method B - Controlled Vocabulary Terms associated with study design of cNRSs and text words that were common to all topic areas, plus those that had an incidence of 50% of more in any topic area
		
		Progressive methods - A progressive method (PM) for selection of study design terms was developed and tested: (a) Start with the list of included studies from each systematic review; (b) Randomly select one study from this list and identify it in the database under study; (c) Retrieve the Controlled Vocabulary Terms (related to study design) associated with the study identified in step (b); (d) Repeat (b) and (c) until a minimum of four CVs is reached; (e) Devise a ''limited search'' in which the ''topic-only'' results are limited to the search terms identified above. (f) The same methodology was applied to identify and test search strategies limited to text words (titles and abstract)

C.10. Were the search terms combined (using Boolean logic) in a way that is likely to retrieve the studies of interest?	Yes. Using the Boolean operator "OR" and "AND"	Yes
C.11. Other observations.	NA	NA

D. Internal validity testing

(this type of testing is possible when the search filter terms were developed from a known reference standard of records)

D.1. How many filters were tested for internal validity?	All filters were tested internally	All filters were tested internally
D.2. Was the performance of the search filter tested on the reference standard	Yes. All six methodological filters were tested on the reference standard	Yes. All 12 filters were tested on the reference standard

Table 4.   Assessment of quality of evidence according to ISSG appraisal checklist *(Continued)*

from which it was derived? (Yes/No/Unclear (please describe))		
D.3. Report sensitivity data (a single value, a range, ''Unclear,''* or ''Not reported,'' as appropriate).	. Please refer to Table 5 Performance of methodological filters in development systematic reviews.	Please refer to Table 5 Performance of methodological filters in development systematic reviews.
D.4. Report precision data (a single value, a range, ''Unclear,''* or ''Not reported,'' as appropriate).	Please refer to Table 5 Performance of methodological filters in development systematic reviews.	Please refer to Table 5 Performance of methodological filters in development systematic reviews.
D.5. Report specificity data (a single value, a range, ''Unclear,''* or ''Not reported,'' as appropriate).	NA	NA
D.6. Other performance measures reported.	NA	NA
D.7. Other observations	NA	PPV in Furlan 2006 was calculated in the same way as precision in Fraser 2006. Therefore, we reported it as precision in our review.

E. External validity testing

(this section relates to testing the search filter on records that are different from the records used to identify the search terms)

E.1. How many filters were tested for external validity on records different from those used to identify the search terms?	Specificity Terms Filters for MEDLINE, for Embase, and for combined MEDLINE/Embase, respectively, were tested externally	No filter was validated externally
E.2. Describe the validation set(S) of records, including the interface.	There were two validation standard sets derived from two additional systematic reviews, photorefractive keratomileusis (PRK) for myopia and electrosurgery for tonsillectomy (Tonsillectomy). For PRK, 39 articles were for MEDLINE and for combined MEDLINE/Embase, 33 for Embase. The publication years were 1990-2004. For Tonsillectomy, 27 articles were for MEDLINE and for Embase, 30 for combined MEDLINE/Embase. The publication years were 1990-2004 The interface was not clearly described	NA
E.3. On which validation set(s) was the filter tested?	"Specificity Terms Filters" was validated externally on two independent validation systematic reviews, PRK and Tonsillectomy, respectively	NA
E.4. Report sensitivity data for each validation set (a single value, a range, ''Unclear,'' or ''Not reported,'' as appropriate).	Sensitivity was not reported	NA

Table 4. Assessment of quality of evidence according to ISSG appraisal checklist *(Continued)*

E.5. Report precision data for each validation set (report a single value, a range, ''Unclear,'' or ''Not reported,'' as appropriate).	Precision was not reported	NA
E.6. Report specificity data for each validation set (a single value, a range, ''Unclear,'' or ''Not reported,'' as appropriate).	Specificity was not reported	NA
E.7. Other performance measures reported.	Reduction in retrievals was reported: 21.9% to 39.6%	NA
E.8. Other observations.	For the external validation of "Specificity Terms Filters", reduction in retrievals and proportion of the reference standard retrieved was reported:instead of the sensitivity and precision as the systematic reviews-derived validation set may have omitted records eligible for inclusion in the reference standard.	NA

F. Limitations and comparisons

F.1. Did the authors discuss any limitations to their research?	Yes. • The reference standards were derived from screening titles and abstracts from a subject-only search, but not a handsearching. This might have failed to pick up some relevant records • The reference standards were limited to the publication years 2000-2004 when the indexing was better. They might not efficient to retrieve earlier publications.	No
F.2. Are there other potential limitations to this research that you have noticed?	• Although both validation sets contained a small amount of records, the author justified that they were independent of the reference standard. • Only one development systematic review was used to develop the search strategy	• No external validation was performed • The reference standards were small • The reference standard formation and evaluation of the search strategies' performance took place more than 10 years ago
F.3. Report any comparisons of the performance of the filter against other relevant published filters (sensitivity, precision, specificity, or other measures).	No.	No.
F.4. Include the citations of any compared filters.	No.	No.
F.5. Other observations and/or comments.	NA	NA

G. Other comments

Table 4.   Assessment of quality of evidence according to ISSG appraisal checklist *(Continued)*

G.1. Have you noticed any errors in the document that might impact on the usability of the filter?	No.	No
G.2 Are there any published errata or comments (e.g., in the MEDLINE record)?	No.	No.
G.3. Is there public access to prepublication history and/or correspondence?	Yes.	Yes.
G.4. Are further data available on a linked site or from the authors?	No.	No.
G.5. Include references to related papers and/or other relevant material.	Fraser 2006	Furlan 2006
G.6. Other comments.	No	No

LASIK: a systematic review to assess the effectiveness and safety of laser in-situ keratomileusis
PRK: a systematic review of photorefractive keratomileusis for myopia
Tonsillectomy: a systematic review of electrosurgery for tonsillectomy
Endarterectomy : a systematic review of local vs. general anaesthesia for carotid endarterectomy
Circumcision: a systematic review of male circumcision for prevention of heterosexual acquisition of HIV in men
Methadone: a systematic review of methadone maintenance at different dosages for opioid dependence
QofL: a systematic review of quality of life after rectal resection for cancer
CV: controlled vocabulary
TW: text words
MP: Term from MeSH and/or title and/or abstract fields
NA: not applicable
PPV: positive predictive values
CI: confidence interval

Table 5.   Performance of methodological filters in development systematic reviews

Study ID	Development systematic reviews	Methodological filters	MEDLINE		EMBASE		MEDLINE/EMBASE	
			Sensitivity (%)	Precision (%)	Sensitivity %	Precision (%)	Sensitivity %	Precision (%)
Fraser 2006	LASIK	Topic-only (no filter)	100	13.2	100	12.6	100	11.1
		Precision Terms Filters	99.5	20.9	100	17.9	100	16.7
		Specificity Terms Filters	99.5	21.1	100	18.8	100	17.1
Furlan 2006	Endarterectomy	Topic-only (no filter)	85	1.81	100	0.62	100	0.58
		Fixed method A	74	4.47	81	0.94	92	1.09
		Fixed method B	76	3.77	95	0.97	95	0.98
		Progressive method (CV)	65	4.02	86	1.04	92	1.13
		Progressive method (CV or TW)	82	4.20	95	1.03	95	1.03
	Circumcision	Topic-only (no filter)	100	0.06	100	0.06	100	0.04
		Fixed method A	48	0.37	93	0.10	76	0.1
		Fixed method B	93	0.12	93	0.08	100	0.07
		Progressive method (CV)	100	0.30	93	0.08	100	0.1
		Progressive method (CV or TW)	100	0.29	100	0.09	100	0.1
	Methadone	Topic-only (no filter)	100	0.06	100	0.06	100	0.04
		Fixed method A	70	0.14	90	0.15	90	0.1
		Fixed method B	80	0.15	90	0.12	100	0.1
		Progressive method (CV)	70	0.18	90	0.12	90	0.1
		Progressive method (CV or TW)	80	0.16	90	0.11	90	0.09

QofL	Topic-only (no filter)	100	0.54	100	0.54	100	0.38
	Fixed method A	70	0.85	90	0.89	82	0.63
	Fixed method B	90	0.98	100	0.97	100	0.73
	Progressive method (CV)	70	1.03	100	0.87	91	0.7
	Progressive method (CV or TW)	90	1.06	100	0.84	100	0.69

LASIK: a systematic review to assess the effectiveness and safety of laser in-situ keratomileusis
PRK: a systematic review of photorefractive keratomileusis for myopia
Tonsillectomy: a systematic review of electrosurgery for tonsillectomy
Endarterectomy : a systematic review of local vs. general anaesthesia for carotid endarterectomy
Circumcision: a systematic review of male circumcision for prevention of heterosexual acquisition of HIV in men
Methadone: a systematic review of methadone maintenance at different dosages for opioid dependence
QofL: a systematic review of quality of life after rectal resection for cancer

APPENDICES

Appendix 1. Search strategies for MEDLINE (Ovid)

1. Epidemiologic Studies/

2. exp Case-Control Studies/

3. exp Cohort Studies/

4. Cross-Sectional Studies/

5. (epidemiologic adj (study or studies)).ab,ti.

6. case control.ab,ti.

7. (cohort adj (study or studies)).ab,ti.

8. cross sectional.ab,ti.

9. cohort analy$.ab,ti.

10. (follow up adj (study or studies)).ab,ti.

11. longitudinal.ab,ti.

12. retrospective$.ab,ti.

13. prospective$.ab,ti.

14. (observ$ adj3 (study or studies)).ab,ti.

15. adverse effect?.ab,ti.

16. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15

17. medline.ti.

18. embase.ti.

19. pubmed.ti.

20. (database? and searching).ti.

21. *MEDLINE/

22. *PubMed/

23. *Databases, Bibliographic/

24. 17 or 18 or 19 or 20 or 21 or 22 or 23

25. 16 and 24

26. ((identify$ or develop$ or design$ or test$ or assess$ or evaluat$ or robust$ or optim$ or effic$ or effect$ or sensitiv$ or simpl$ or specific$ or precis$) adj3 ("search strat$" or "search filter?")).ab,ti.

27. 16 and 26

28. 25 or 27

Appendix 2. Search strategies for Embase (Ovid)

1. epidemiology/

2. exp case control study/

3. cohort analysis/

4. cross-sectional study/

5. case study/

6. follow up/

7. longitudinal study/

8. retrospective study/

9. prospective study/

10. observational study/

11. correlational study/

12. epidemiologic.ab,ti.

13. case control.ab,ti.

14. case referent.ab,ti.

15. case stud$.ab,ti.

16. case series.ab,ti.

17. cohort?.ab,ti.

18. cross sectional.ab,ti.

19. follow up.ab,ti.

20. longitudinal.ab,ti.

21. retrospective$.ab,ti.

22. prospective$.ab,ti.

23. observational.ab,ti.

24. adverse effect?.ab,ti.

25. (Controlled before and after).ab,ti.

26. Interrupted time series.ab,ti.

27. Correlational.ab,ti.

28. ecological stud$.ab,ti.

29. Descriptive stud$.ab,ti.

30. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29

31. medline.ti.

32. embase.ti.

33. pubmed.ti.

34. (database? and searching).ti.

35. *medline/

36. *embase/

37. *bibliographic database/

38. 31 or 32 or 33 or 34 or 35 or 36 or 37

39. 30 and 38

40. ((identify$ or develop$ or design$ or test$ or assess$ or evaluat$ or robust$ or optim$ or effic$ or effect$ or sensitiv$ or simpl$ or specific$ or precis$) adj3 ("search strat$" or "search filter?")).ti,ab.

41. 30 and 40

42. 39 or 41

Appendix 3. Search strategies for CINAHL

S1 (MH "Case Control Studies+")

S2 (MH "Case Studies")

S3 (MH "Cross Sectional Studies")

S4 (MH "Prospective Studies+")

S5 (MH "Retrospective Panel Studies")

S6 (MH "Correlational Studies")

S7 (MH "Ecological Research")

S8 (MH "Descriptive Research")

S9 TI epidemiologic OR AB epidemiologic

S10 TI "case control" OR

S11 TI "case referent" OR AB "case referent*"

S12 TI "case stud*" OR AB "case stud*"

S13 TI "case series" OR AB "case series"

S14 TI cohort* OR AB cohort*

S15 TI "cross sectional" OR AB "cross sectional"

S16 TI "follow up" OR AB "follow up"

S17 TI longitudinal OR AB longitudinal

S18 TI retrospective* OR AB retrospective*

S19 TI prospective* OR AB prospective*

S20 TI observational OR AB observational

S21 TI "adverse effect*" OR AB "adverse effect*"

S22 TI "Controlled before and after" OR AB "Controlled before and after"

S23 TI "Interrupted time series" OR AB "Interrupted time series"

S24 TI Correlational OR AB Correlational

S25 TI "ecological stud*" OR AB "ecological stud*"

S26 TI "Descriptive stud*" OR AB "Descriptive stud*"

S27 S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26

S28 TI medline

S29 TI embase

S30 TI pubmed

S31 TI (database* and searching)

S32 (MM "*MEDLINE")

S33 (MM "*PUBMED")

S34 (MM "*EMBASE")

S35 S28 OR S29 OR S30 OR S31 OR S32 OR S33 OR S34

S36 S27 AND S35

S37 TI (((identify* or develop* or design* or test* or assess* or evaluat* or robust* or optim* or effic* or effect* or sensitiv* or simpl* or specific* or precis*) N3 ("search strat*" or "search filter*"))) OR AB (((identify* or develop* or design* or test* or assess* or evaluat* or robust* or optim* or effic* or effect* or sensitiv* or simpl* or specific* or precis*) N3 ("search strat*" or "search filter*")))

S38 S27 AND S37

Note: S is the search ID for CINAHL

Appendix 4. Search strategies for Cochrane Library

We used the Cochrane Library listed in NTU databases, selected 'Advanced Search', followed by 'Search Manager'.

1. epidemiologic:ab,ti

2. "case control":ab,ti

3. "case referent":ab,ti

4. "case stud*":ab,ti

5. "case series":ab,ti

6. cohort*:ab,ti

7. "cross sectional":ab,ti

8. "follow up":ab,ti

9. longitudinal:ab,ti

10. retrospective*:ab,ti

11. prospective*:ab,ti

12. observational:ab,ti

13. "adverse effect*":ab,ti

14. "Controlled before and after":ab,ti

15. "Interrupted time series":ab,ti

16. Correlational:ab,ti

17. "ecological stud*":ab,ti

18. "Descriptive stud*":ab,ti

19. #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18

20. medline:ti

21. embase:ti

22. pubmed:ti

23. (database* and searching):ti

24. #20 or #21 or #22 or #23

25. #19 and #24

26. ((identify* or develop* or design* or test* or assess* or evaluat* or robust* or optim* or effic* or effect* or sensitiv* or simpl* or specific* or precis*) NEAR/3 ("search strat*" or "search filter*")):ab,ti

27. #19 and #26

28. #25 or #27

Appendix 5. Search Strategies for Google Scholar

With all of the words: Search Strategy

With at least one of the words: Medline Embase PubMed

Where my words occur: anywhere in the article

Return articles: 3,000,000

Appendix 6. Search strategies for OpenGrey

((subject:Epidemiologic Studies OR subject:Case-Control Studies OR subject:Cohort Studies OR subject:Cross-Sectional Studies OR abstract:(epidemiologic NEAR/1 (study or studies)) OR title:(epidemiologic NEAR/1 (study or studies)) OR abstract:(case control) OR title:(case control) OR abstract:(cohort adj (study or studies)) OR title:(cohort adj (study or studies)) OR abstract:(cross sectional) OR title:(cross sectional) OR abstract:(cohort analy*) OR title:(cohort analy*) OR abstract:(follow up NEAR/1 (study or studies)) OR title: (follow up NEAR/1 (study or studies)) OR abstract:longitudical OR title:longitudinal OR abstract:retrospective* OR title:retrospective* OR abstract:prospective* OR title:prospective* OR abstract:(observ* NEAR/3 (study or studies)) OR title:(observ* NEAR/3 (study or studies)) OR abstract:(adverse effect*) OR title:(adverse effect*)) AND (title:medline OR title:embase OR title:pubmed OR title:(database* and searching) OR subject:MEDLINE OR subject:PubMed OR subject:(Databases, Bibliographic))) OR ((subject:Epidemiologic Studies OR subject:Case-Control Studies OR subject:Cohort Studies OR subject:Cross-Sectional Studies OR abstract:(epidemiologic NEAR/1 (study or studies)) OR title:(epidemiologic NEAR/1 (study or studies)) OR abstract:(case control) OR title:(case control) OR abstract:(cohort adj (study or studies)) OR title:(cohort adj (study or studies)) OR abstract:(cross sectional) OR title:(cross sectional) OR abstract:(cohort analy*) OR title: (cohort analy*) OR abstract:(follow up NEAR/1 (study or studies)) OR title:(follow up NEAR/1 (study or studies)) OR abstract:longitudical OR title:longitudinal OR abstract:retrospective* OR title:retrospective* OR abstract:prospective* OR title:prospective* OR abstract:(observ* NEAR/3 (study or studies)) OR title:(observ* NEAR/3 (study or studies)) OR abstract:(adverse effect*) OR title:(adverse effect*)) AND (abstract:((identify* or develop* or design* or test* or assess* or evaluat* or robust* or optim* or effic* or effect* or sensitive* or simpl* or specific* or precis*) NEAR/3 ("search strat*" or "search filter*")) OR title:((identify* or develop* or design* or test* or assess* or evaluat* or robust* or optim* or effic* or effect* or sensitive* or simpl* or specific* or precis*) NEAR/3 ("search strat*" or "search filter*"))))

F E E D B A C K

Recently published eligible study, 12 April 2019

Summary

We want to congratulate the authors of this Cochrane Methodology Review for publishing summarized evidence on the important topic of search filter evaluation for observational studies.

As authors of an extensive validation study on the same topic, published in December 2018, we would like to alert the interested reader that the implications for research identified by the authors of this Review have been answered comprehensively in our study.

Our research focused on MEDLINE search filters, but the findings are highly likely to apply to filters developed for other databases as well. There was an unfortunate time lapse between the search strategy run to inform this Cochrane Methodology Review (last run on April 2018) and the publication of our study in December 2018. But even if the Cochrane Methodology Review´s strategy had been run again before publication, it would not have identified our study, as the terminology used by us ("non-randomized studies") is not included in the search strategies of this review. The conclusion of our study is that the performance of the currently available non-randomized study (NRS) filters is insufficient for effective use in daily practice. Therefore, it is necessary to develop new strategies allowing a sound identification of NRS in the future (e.g. new NRS filters in combination with other search techniques). Details of our validation study can be found in the following publication: Hausner E, Metzendorf MI, Richter B, Lotz F, Waffenschmidt S. Study filters for non-randomized studies

of interventions consistently lacked sensitivity upon external validation. BMC Medical Research Methodology 2018;18(1):171. doi: 10.1186/s12874-018-0625-4. PMID: 30563471.

Reply

Thank you for your comment and for bringing your validation study to our attention. We will be pleased to consider it in the future update of our Cochrane Methodology Review. We fully agree with you that there is an urgent need for more research on methodological filters that will help people to find a wide range of observational studies, including those that are purely observational as well as experimental studies such as non-randomised trials of the effects of interventions.

Contributors

Feedback submitted by Maria-Inti Metzendorf (Information Scientist at the Cochrane Metabolic and Endocrine Disorders Group). Reply prepared by Lorainne Tudor Car (lead author) and Mike Clarke (Co-ordinating Editor, Cochrane Methodology Review Group) and agreed with Maria-Inti.

WHAT'S NEW

Date	Event	Description
14 October 2019	Feedback has been incorporated	Feedback and reply added in relation to a recently published eligible study

CONTRIBUTIONS OF AUTHORS

LL ran the searches, screened the literature, extracted the data, analysed the data and drafted the manuscript. LTC helped design the study, screened literature searches, extracted data, analysed the data and helped draft the manuscript. HS and RA provided comments and edited the review.

DECLARATIONS OF INTEREST

None.

SOURCES OF SUPPORT

Internal sources

- Lee Kong Chian School of Medicine, Nanyang Technological University Start Up Grant, Singapore.

  Salary for LTC, LL & HES

External sources

- No sources of support supplied

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

There are some differences between the protocol and the review.

- Changes in searched databases
  - Cochrane Methodology Register (CMR): we searched the Cochrane Library through the Nanyang Technology University library website instead of CMR because the last Issue that contained an up-to-date CMR was in July 2012. We performed our searches in April 2017.
  - Google Scholar: we changed Google Scholar search from "Electronic searches" to "Searching other resources". Three search strategies (Appendix 5) retrieved more than three million citations, so we screened the "titles and abstracts" for the first 500 citations from each search and found zero additional eligible studies.
- Changes in "Assessment of risk of bias in included studies": we did not assess item (5) requirements of the intended application; nor did we assess item (7) ability to reproduce accurately historical searches when included studies from a review are the reference standard. Instead, we used the UK InterTasc Information Specialists' Sub-Group (ISSG) Search Filter Appraisal Checklist (Glanville 2008) and assessed (1) information and objective of the search filter; (2) identification of a reference standard; (3) search filter development; (4) performance assessment (or internal validation as it is referred to in the checklist), (5) external validation and (6) potential limitations and comparisons.

- We provided additional details about our exclusion criteria in the Methods, Types of studies section.

INDEX TERMS

Medical Subject Headings (MeSH)

*Databases, Bibliographic; *Non-Randomized Controlled Trials as Topic; *Observational Studies as Topic; Abstracting and Indexing [standards]; Information Storage and Retrieval [*methods]; MEDLINE; Sensitivity and Specificity; Systematic Reviews as Topic