Received: 2 July 2021 | Revised: 23 March 2022 | Accepted: 19 May 2022

DOI: 10.1111/hir.12441

ORIGINAL ARTICLE

Health Information and Libraries Journal | WILEY

Updated generic search filters for finding studies of adverse drug effects in Ovid MEDLINE and Embase may retrieve up to 90% of relevant studies

Su Golder PhD[1] | **Kelly Farrah MLIS**[2] | **Monika Mierzwinski-Urban MLIS**[2] | **Beth Barker BSc (hons)**[3] | **Anna Rama**[4]

[1]Department of Health Sciences, University of York, York, UK

[2]Research Information Services, Canadian Agency for Drugs and Technologies in Health (CADTH), Ottawa, Ontario, Canada

[3]Department of Social and Policy Sciences, University of Bath, Bath, UK

[4]Hull York Medical School (HYMS), University of York, York, UK

Correspondence
Su Golder, Department of Health Sciences, University of York, York, YO10 5DD, UK.
Email: su.golder@york.ac.uk

Abstract

Background: The most current objectively derived search filters for adverse *drug* effects are 15 years old and other strategies have not been developed and tested empirically.

Objective: To develop and validate search filters to retrieve evidence on adverse drug effects from Ovid MEDLINE and Ovid Embase.

Methods: We identified systematic reviews of adverse drug effects in Epistemonikos. From these reviews, we collated their included studies which we then randomly divided into three tests and one validation set of records. We constructed a search strategy to maximise relative recall using word frequency analysis with test set one. This search strategy was then refined using test sets two and three and validated on the final set of records.

Results: Of 107 systematic reviews which met our inclusion criteria, 1948 unique included studies were available from MEDLINE and 1980 from Embase. Generic adverse drug effects searches in MEDLINE and Embase achieved 90% and 89% relative recall, respectively. When specific adverse effects terms were added recall was improved.

Conclusion: We have derived and validated search filters that retrieve around 90% of records with adverse drug effects data in MEDLINE and Embase. The addition of specific adverse effects terms is required to achieve higher recall.

KEYWORDS

Embase, evaluation, information retrieval, literature searching, medical subject headings (MeSH), methodological filters, search strategies

INTRODUCTION

Information is required in clinical practice to generate appropriate advice to patients on the benefits versus the harms of specific medications. As clinicians generally rely on evidence-based resources such as clinical practice guidelines and systematic reviews in clinical decision making, these need to consider the evidence on adverse

drug effects (Peryer et al., 2021). Any properly conducted systematic review will employ highly sensitive search strategies that aim to identify as many relevant papers as possible, whilst minimising retrieval of irrelevant papers. However, retrieving a complete cohort of studies on adverse effects is challenging due to poor reporting, inconsistent terminology and lack of prior knowledge of the specific adverse effects to search for (Golder et al., 2006). Adverse effects are often overlooked in primary research and, even when they are considered, they are likely to be secondary or tertiary outcomes and subsequently poorly reported in the title, abstract and indexing of database records.

Whilst randomised controlled trials (RCTs) may be the most appropriate study design to study intervention effectiveness this is not always the case for adverse effects. In particular, RCTs are often not the most appropriate study design to assess long-term, rare or unexpected adverse effects or those that occur in vulnerable patients who are often excluded from RCTs. It is, therefore, common to include study designs beyond RCTs for identifying the adverse drug effects. Whilst search filters for RCTs have been proven to perform well, searching for non-RCT study designs is more problematic (Higgins et al., 2019).

One way to help enable efficient searching is through the development of search filters. Search filters are combinations of search terms which are designed to improve the efficiency and effectiveness of searching. Objectively (research-based) derived search filters for medical device adverse effects and complications in surgery were developed in 2018 and 2019 (Golder et al., 2018; Golder, Farrah, et al., 2019). Yet, the majority of reviews of adverse effects study drug interventions (Golder et al., 2016; Golde, Peryer, et al., 2019) and objectively derived search filters for adverse drug effects are now outdated (Badgett et al., 1999; Golder & Loke, 2012a, 2012b, 2012c; Wieland & Dickersin, 2005). Furthermore, search filters for adverse effects of medical devices have been proven to be ineffective for identifying adverse effects of drug interventions (Farrah et al., 2016). Expert derived but unvalidated search filters have been suggested more recently and testing of these filters is now required (BMJ, 2021; CADTH, 2021). Validation of search strategies requires the testing (and often retesting) of the search on a different set of references to those used in the development of the strategy (Hausner et al., 2012).

We aimed to create up-to-date validated adverse drug effects search filters with high recall for both OVID MEDLINE and OVID Embase and to compare these new search filters to previously available search strategies.

Key messages

- Generic adverse drug effects searches can achieve over 89% relative recall in either MEDLINE or Embase.
- The addition of specific named adverse drug effects search terms is likely to improve relative recall to over 90%.
- Despite improvements in the reporting of adverse drug effects—not all records indicate that the full paper contains adverse drug effects data.
- Our proposed search filter achieves a slightly higher relative recall than most other adverse drug effects filters.
- More research is required to measure the performance of adverse drug filters in terms of improving precision.

METHODS

Systematic review identification

We conducted a search of Epistemonikos (https://www.epistemonikos.org/) to identify a current cohort of systematic reviews of adverse effects. Epistemonikos was chosen as it is currently the largest open source of systematic reviews available to the authors.

We were unable to simply sift the records available in Epistemonikos due to the large volume of systematic reviews published. A limit was placed of 'Publication Type: Systematic Reviews', and 'Publication Date: 2019 to Current' in order to retrieve a recent cohort of systematic reviews. Sensitive searches were conducted on 29 January 2020 with the terms;

safe* OR complication* OR adverse* OR side effect* OR harm* OR risk* OR tolerability OR teratogen* OR toxic* OR induced (title) OR pharmacotox* OR neurotox* OR cardiotox* OR nephrotox* OR immunotox* OR hepatox* OR cytotox* OR immunocytotox* (title or abstract).

These terms were identified from previous adverse drug effects search filters (Golder et al., 2006; Wolters Kluwer Health, 2021).

A systematic review was considered eligible for inclusion if:

- The intervention was a pharmaceutical drug medication (prescribed or over the counter).
- The outcome(s) included adverse effect(s) of the drug intervention.

- The search strategy was reported in full in the published paper or Supporting Information, and no adverse effects search filter or adverse effects search terms (either generic, such as 'adverse reactions' or 'harms' or named, such as 'weight gain' or 'rash') were used in any of the fields searched. Typically, reviews included relied solely on searching for terms for the population/condition and intervention. This enabled us to construct an unbiased cohort which did not include articles that had been retrieved because they already contained adverse effects terms.
- The search was required to have included either handsearching or reference checking in addition to database searches of MEDLINE and Embase. This was to compensate for potential deficiencies in the search strategies used in the reviews.
- Each review needed to include at least one study related to adverse effects. This was to avoid empty reviews which would be unable to contribute any primary studies to our cohort.
- Due to logistical constraints on obtaining translations reviews needed to be published in English or have an English language translation available, and a full-text copy of the paper needed to be accessible to the authors.

All the titles and abstracts of the records retrieved were independently screened by two researchers (SG, KF or MM-U) using DistillerSR (Evidence Partners, c2021, https://www.evidencepartners.com) and potentially relevant systematic reviews were selected. Any discrepancies between the researchers were resolved by discussion and consensus or by consulting a third reviewer. Due to constraints on time and resources, we randomly selected 200 of the potentially relevant systematic reviews using allocated numbers from RAMDOM.ORG (Haahr, 2021) and obtained full text of the systematic reviews selected for further screening. These full-text systematic reviews were again independently screened, with discrepancies resolved by discussion and consensus (SG, KF or MM-U).

Included primary studies

For each systematic review that met our inclusion criteria, we identified the included studies within the reviews that were purported to contain adverse effects data and then checked the full-text article. The use of included papers from systematic reviews has been shown to be an effective alternative to handsearching to identify a reference gold standard set of records for developing and evaluating search strategies (Sampson et al., 2006).

The next stage of the analysis was to ascertain whether each paper was contained in MEDLONE or Embase. We searched for each paper using several search iterations as necessary of the author names or words from the title or even abstract in an attempt to ascertain if the paper was contained in the database. The records available on MEDLINE and Embase were then divided into three test sets and one validation set of records for each respective database using random numbers generated by RANDOM.ORG.

Individual word and multiple-word frequency analysis on the first test set of records were undertaken using WriteWords to identify commonly occurring terms related to adverse effects. WriteWords is freely available on the internet and allows frequency counting of the usage of words or phrases (http://www.writewords.org.uk/phrase_count.asp). We entered the terms for each field (title, abstract, MeSH/EMTREE) separately. We noted any terms with similar stems (such as safety, safe, safely) so that we could test such terms separately and with appropriate truncation. We then noted the search terms identified via WriteWords along with the number of records retrieved using each term in the first test set of records. In order to provide a relative estimate of the precision of each of the search terms, we also recorded the total number of records retrieved in each database at the time of completing this research.

We calculated relative recall as a measure of the percentage of all known records retrieved using the search terms because it provides an estimate of sensitivity (Sampson et al., 2006). The relative recall of the relevant search terms was calculated using the following formula;

Relative Recall Calculation

$$\frac{\text{No of relevant records retrieved}}{\text{No of relevant records available}} \times 100 = \text{Relative recall as a percentage}\,(\%).$$

A first draft filter was created with the first test set of records. To do this we started with the search term that had the highest recall and then tested all other potentially relevant terms in turn to ascertain the term with the highest incremental increase in recall when searched with the first search term using the Boolean operator OR. This process continued until no more new records were being identified by any additional search terms being ORed with the preceding selected terms.

We then tested the combination of ORed search terms created with the first test set to ascertain the number of records identified in the second test set. We added additional search terms using the OR Boolean operator if they could further improve recall. After any additional modifications, the new version of the filter was applied to the third test set and again additional search terms were added using the OR Boolean operator if they improved

recall. The final version of the search filter was tested in the validation set (but no modifications were permitted).

In order to ascertain whether any terms in the records of those not identified by our search filter could be indicative that the full text contained adverse effects data, we examined those records not retrieved by our generic search term filters. In particular, we recorded terms related to actual specific adverse effects (such as 'infection' or 'mortality', etc.) and the field (such as abstract or indexing) where they were found. We noted the number of database records with no indication that the full text contained information on adverse effects.

The process was first undertaken in MEDLINE and then repeated in Embase. The resultant performance of the new search filters was then compared to those published in the literature for adverse effects of other interventions, namely surgical procedures and medical devices. We also compared the performance in terms of relative recall of this new filter to other available search filters developed to retrieve adverse drug effects using our test and validation sets of records (Table S1) (BMJ, 2021; CADTH, 2021; Canadian Health Libraries Association, 2018; Golder et al., 2006; Wolters Kluwer Health, 2021).

In order to ascertain an approximation of the impact of the filter on precision when used with a search on a particular topic, the final search filters were used in two case examples in MEDLIE and Embase. The number of records retrieved with and without the filter were then recorded.

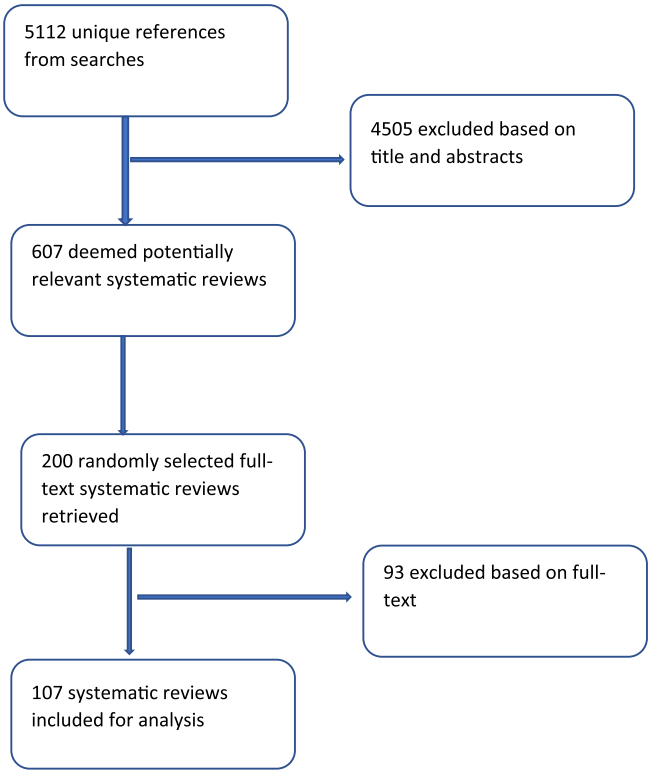

FIGURE 1 Flow diagram of systematic review selection process

RESULTS

From 5112 Epistemonikos records screened, 607 were deemed potentially relevant systematic reviews based on the title and abstract. We then selected a random sample of 200 of these 607 potentially relevant reviews for full-text screening and found that 107 met our inclusion criteria (Figure 1). These 107 reviews included 1989 primary studies (2314 studies before deduplication) and of these included primary studies, 1948 unique records were available on MEDLINE and 1980 on Embase. Results of testing are summarised in Table 1 and described in further detail below.

MEDLINE

The gold standard set of 1948 records in MEDLINE were randomly allocated into three test sets of 487 records each and one validation set of 487 records.

First test set for the development of the MEDLINE search filter

Of the search terms identified in the first test set, 'adverse effects (ae)' as a floating subheading had the highest

TABLE 1 Performance of the search strategies

	Search terms	Test set of records (1)	Test set of records (2)	Test set of records (3)	Validation set of records
MEDLINE	Box 1	(420/487) 86%	(435/487) 89%	(440/487) 90%	(439/487) 90%
	Box 1 with specific adverse effects terms	(437/487) 90%	(437/487) 90%	(444/487) 91%	(447/487) 92%
Embase	Box 2	(445/495) 90%	(416/495) 84%	(440/495) 89%	(443/495) 89%
	Box 2 with specific adverse effects terms	(460/495) 93%	(435/495) 88%	(454/495) 92%	(462/495) 93%

Note: >90% sensitivity is represented by green shading. <90% sensitivity is represented by red shading.

recall and was searched first. This was followed by the terms 'safe*' in the title and abstract which gave the highest incremental increase in recall when ORed with the floating subheading 'adverse effects (ae)'.

The addition of further terms resulted in a search strategy (Box 1) which retrieved 86% (420/487) of records. Of the 67 records not retrieved, 17 contained terms for specific adverse effects (Table S1) whereas 50 records gave no indication that the full paper contained information on adverse effects. The specific adverse effects terms (such as nausea or bleeding) were not added to the search as they tend to only apply to specific medications. A search strategy which incorporates both generic and specific adverse effects terms could potentially achieve a 90% (437/487) recall.

BOX 1 MEDLINE **search strategy from first, second and third test set of records**

Test set 1 ($n = 487$)		Test set 2 ($n = 487$)		Test set 3 ($n = 487$)/final search strategy	
Search term	**No of records retrieved (percent recall)**	**Search term**	**No of records retrieved (percent recall)**	**Search term**	**No of records retrieved (percent recall)**
Ae.fs	255 (52%)	Ae.fs.	265 (54%)	OR Ae.fs.	242 (50%)
OR Safe*.ti,ab.	326 (67%)	OR Safe*.ti,ab.	339 (70%)	OR Safe*.ti,ab.	330 (68%)
OR De.fs.	362 (74%)	OR De.fs.	380 (78%)	OR De.fs.	374 (77%)
OR Adverse.ti,ab.	380 (78%)	OR Adverse.ti,ab.	396 (81%)	OR Adverse.ti,ab.	405 (83%)
OR Co.fs.	394 (81%)	OR Co.fs.	409 (84%)	OR Co.fs.	419 (86%)
OR Side effect*.ti,ab.	402 (83%)	OR Side effect*.ti,ab.	414 (85%)	OR Side effect*.ti,ab.	425 (87%)
OR Complication*.ti,ab.	407 (84%)	OR Complication*.ti,ab.	421 (86%)	OR Complication*.ti,ab.	427 (88%)
OR Ci.fs.	411 (84%)	OR Ci.fs.	421 (86%)	OR Ci.fs.	428 (88%)
OR Tolerated.ti,ab.	415 (85%)	OR Tolerated.ti,ab.	422 (87%)	OR Tolerated.ti,ab.	431 (89%)
OR Tolerance.ti,ab.	416 (85%)	OR Tolerance.ti,ab.	423 (87%)	OR Tolerance.ti,ab.	433 (89%)
OR Harm*.ti,ab.	417 (86%)	OR Harm*.ti,ab.	424 (87%)	OR Harm*.ti,ab.	433 (89%)
OR Toxicity.ti,ab.	418 (86%)	OR Toxicity.ti,ab.	425 (87%)	OR Toxicity.ti,ab.	433 (89%)
OR Risk.ti.	419 (86%)	OR Risk.ti.	429 (88%)	OR Risk.ti.	436 (90%)
OR Pregnancy complications/dt	420 (86%)	OR Pregnancy complications/dt]	429 (88%)	OR Pregnancy complications/dt	436 (90%)
		OR **Clinical trial phase IV.pt.**	431 (89%)	OR Clinical trial phase IV.pt.	437 (90%)
		OR **Drug hypersensitivity/**	432 (89%)	OR Drug hypersensitivity/	437 (90%)
		OR **Tolerability.ti,ab.**	433 (89%)	OR Tolerability.ti,ab.	437 (90%)
		OR **To.fs.**	434 (89%)	OR To.fs.	437 (90%)
		OR **Toxicology/**	435 (89%)	OR Toxicology/	437 (90%)
				OR **Drug induced.ti,ab.**	439 (90%)
				OR **Negative effects.ti,ab.**	440 (90%)

Note: Warning: de.fs may be too noisy, use with caution. .fs, floating subheading; .ti, title; ab, abstract; pt, publication type; /dt, drug therapy; /, MeSH term. *Terms in bold indicate new terms added within that test set.

Second test set for the development of the MEDLNE search filter

The search strategy from the first test set (Box 1) was tested on the second test set of records and retrieved 88% (429/487). There were six additional records that could have been retrieved if 'Clinical trial phase IV' as publication type, the MeSH term 'drug hypersensitivity/', the textword 'tolerability', the floating subheading 'to [toxicity]' and the MeSH term 'Toxicology/' were added to the search strategy. After adding these terms to the search strategy, the revised strategy retrieved 89% (435/487) of the records in this second test set.

Of the 52 records that had not been retrieved by this search strategy, 2 contained specific adverse effects terms (Table S1). A search strategy which incorporates both generic and specific adverse effects terms could potentially achieve 90% (437/487) recall in the second test set of records.

Third test set for the development of the MEDLINE search filter

The search strategy from the second test set (Box 1) was then tested on the third test set of records and retrieved 90% (437/487) of records. There were three additional records in this test set that could have been retrieved if 'drug induced' and 'negative effects' in the abstract were added to the search strategy. Hence, after adding these terms, the revised strategy retrieved 90% (440/487) of records in this third test set.

Of the 47 records not retrieved by the search strategy, four had terms related to specific adverse effects (Table S1). A search strategy which incorporates both generic and specific adverse effects terms could potentially achieve 91% (444/487) recall in the third test set of records.

Validation of the MEDLINE search filter

The revised search strategy (Box 1) performed consistently high on the validation set of records as the test sets and retrieved 90% (439/487) of records. The term 'toxic effects' in the abstract would have identified one additional record.

Of the remaining 47 records not retrieved, seven contained terms related to specific adverse effects (Table S1). A search strategy which incorporates both generic and specific adverse effects terms could potentially achieve 92% (447/487) recall in the validation set of records.

Comparative analysis

The newly derived MEDLINE search filter achieves marginally higher recall than previous search filters (89% vs. 87%, 86%, 82% or 70%) (Table 2) when tested against the complete gold standard set of relevant articles. Of the previous search filters, the CADTH search filter and Golder 2006 performed particularly well on our set of records (87% and 86% recall, respectively).

TABLE 2 Comparison of the new search strategy with previous search filters

	Golder et al. (2006)	CADTH	BMJ	CHLA	New filter
MEDLINE					
Test set 1	403/487 (83%)	405/487 (83%)	389/487 (80%)	327/487 (67%)	420/487 (86%)
Test set 2	414/487 (85%)	421/487 (86%)	403/487 (83%)	353/487 (72%)	435/487 (89%)
Test set 3	426/487 (87%)	430/487 (88%)	401/487 (82%)	337/487 (69%)	440/487 (90%)
Validation set	427/488 (88%)	434/488 (89%)	405/488 (83%)	350/487 (72%)	439/487 (90%)
All sets	86%	87%	82%	70%	89%

	Golder et al. (2006)	CADTH	BMJ	EMA	EMA—focused	New filter
Embase						
Test set 1	391/497 (79%)	409/497 (82%)	427/497 (86%)	442/497 (89%)	376/497 (76%)	(445/495) 90%
Test set 2	373/496 (75%)	390/496 (79%)	403/496 (81%)	441/496 (89%)	348/496 (70%)	(416/495) 84%
Test set 3	389/496 (78%)	415/496 (84%)	422/496 (85%)	446/496 (90%)	369/496 (74%)	(440/495) 89%
Validation set	399/497 (80%)	416/497 (84%)	422/497 (85%)	455/497 (92%)	388/497 (78%)	(443/495) 89%
All sets	78%	82%	84%	90%	75%	88%

Note: >90% sensitivity is represented by green shading. <90% sensitivity is represented by red shading.

TABLE 3 Comparison of the new adverse drug effects search strategy with search strategies for surgical procedures and medical devices

	Adverse effect filters	Surgical procedures (Golder et al., 2018)	Medical device (Golder et al., 2019)	Drug intervention (current study)
MEDLINE	Generic	87%	83%	90%
	Generic and specific	93%	92%	92%
Embase	Generic	92%	83%	89%
	Generic and specific	95%	90%	93%

Note: >90% sensitivity is represented by green shading. <90% sensitivity is represented by red shading.

The adverse drug effects search filter derived here using only generic adverse effects terms achieved a higher recall (90%) than either the search filter developed for the adverse effects of medical devices (83%) (Golder, Farrah, et al., 2019) or surgery (87%) (Golder et al., 2018) (Table 3).

Embase

The gold standard set of 1980 records in Embase were randomly divided into three test sets of 495 records, and a validation set of 495 records.

First test set of records for the development of the Embase search filter

The floating subheading 'adverse drug reaction (ae)' had the highest recall and was searched first. This was followed by 'safe*' in the title, abstract, keyword and other index terms word which gave the highest incremental increase in recall when added to the floating subheading 'adverse drug reaction (ae)'.

The addition of further terms resulted in a search strategy (Box 2) which retrieved 90% (445/495) records. Of the 50 records not retrieved by the search strategy, 15 had terms related to specific adverse effects (Table S1) whereas 35 gave no indication that the full paper contained information on adverse effects. A search strategy which incorporates both generic and specific adverse effects terms could therefore potentially achieve 93% (460/495) recall.

Second test set of records for the development of the Embase search filter

The search strategy from the first test set (Box 2) was tested on the second test set of records and retrieved 84% (415/495). There was an additional record that could have been retrieved if toxic effects in the title and abstract were added to the search strategy. After adding this term to the search strategy, the revised strategy retrieved 84% (416/495) of the records in this second test set.

Of the 79 records not retrieved, 19 had terms related to specific adverse effects (Table S1). A search strategy which incorporates both generic and specific adverse effects terms could therefore potentially achieve 88% (435/495) recall in the second test set of records.

Third test set of records for the development of the Embase search filter

The search strategy from the second test set (Box 2) was then tested on the third test set of records and retrieved 88% (438/495) of records. There were two additional records in this test set that could have been retrieved if toxicity and toxicities in the title and abstract were added to the search strategy. Hence, after adding these terms, the revised strategy retrieved 89% (440/495) of records in this third test set.

Of the 55 records not retrieved by the search strategy, 14 had terms related to specific adverse effects (Table S1). A search strategy which incorporates both generic and specific adverse effects terms could therefore potentially achieve 92% (454/495) recall in the third test set of records.

Validation of the Embase search filter

The revised search strategy in Box 2 was then tested on the validation set of records and retrieved 89% (443/495) of the records. We conducted post hoc analysis to identify factors that may have affected the recall. When we explored the records that had not been retrieved from the validation set, 'Adverse Reactions Titles' searched in the Embase Section Headings field (ec), 'pregnancy complication/dt' as an EMTREE and subheading and the EMTREE terms 'multidrug resistance/', 'drug dependence/', 'withdrawal syndrome/' would have retrieved a

BOX 2 Embase search strategy from first, second and third test set of records

Test set 1 ($n = 495$)		Test set 2 ($n = 495$)		Test set 3 ($n = 495$)/final search strategy	
Search term	**No of records retrieved (percent recall)**	**Search term**	**No of records retrieved (percent recall)**	**Search term**	**No of records retrieved (percent recall)**
Ae.fs	316 (64%)	Ae.fs	307 (62%)	Ae.fs	321 (65%)
OR Safe*.ti,ab,kw, ox	363 (73%)	OR Safe*.ti,ab,kw,ox	346 (70%	OR Safe*.ti,ab,kw	357 (72%)
OR Adverse.ti,ab, kw,ox	380 (77%)	OR Adverse.ti,ab,kw,ox	375 (76%)	OR Adverse.ti,ab, kw,ox	390 (79%)
OR Po.fs	393 (79%)	OR Po.fs	386 (78%)	OR Po.fs	401 (81%)
OR Co.fs	403 (81%)	OR Co.fs	397 (80%)	OR Co.fs	413 (83%)
OR exp adverse drug reaction/	413 (83%)	OR exp adverse drug reaction/	398 (80%)	OR exp adverse drug reaction/	413 (83%)
OR Complication*. ti, ab	422 (85%)	OR Complication*.ti,ab	401 (81%)	OR Complication*. ti,ab	417 (84%)
OR Drug safety/	427 (86%)	OR Drug safety/	402 (81%)	OR Drug safety/	423 (85%)
OR To.fs	431 (87%)	OR To.fs	402 (81%)	OR To.fs	423 (85%)
OR Side effect*. ti,ab	434 (88%)	OR Side effect*.ti,ab	405 (82%)	OR Side effect*. ti,ab	429 (87%)
OR Risk.ti	437 (88%)	OR Risk.ti	409 (83%)	OR Risk.ti	435 (88%)
OR Tolerance.ti,ab	439 (89%)	OR Tolerance.ti,ab	410 (83%)	OR Tolerance.ti,ab	436 (88%)
OR Tolerated.ti,ab	440 (89%)	OR Tolerated.ti,ab	414 (84%)	OR Tolerated.ti,ab	437 (88%)
OR Harm.ti,ab	441 (89%)	OR Harm.ti,ab	415 (84%)	OR Harm.ti,ab	437 (88%)
OR Side reaction*. ti,ab	442 (89%)	OR Side reaction*.ti,ab	415 (84%)	OR Side reaction*. ti,ab	437 (88%)
OR drug withdrawal/	443 (89%)	OR drug withdrawal/	415 (84%)	OR drug withdrawal/	438 (88%)
OR health risks. ti,ab	444 (90%)	OR health risks.ti,ab	415 (84%)	OR health risks. ti,ab	438 (88%)
OR potential risks. ti,ab	445 (90%)	OR potential risks.ti,ab	415 (84%)	OR potential risks. ti,ab	438 (88%)
		OR **toxic effects.ti,ab**	416 (84%)	OR toxic effects. ti,ab	438 (88%)
				OR **toxicity.ti,ab**	439 (89%)
				OR **toxicities. ti,ab**	440 (89%)

Note: .fs, floating subheading; .ti, title; ab, abstract; kw, keyword; ox, other index terms word; /, MeSH term. *Terms in bold indicate new terms added within that test set.

further four studies—these terms are indicative of generic or specific adverse effects.

However, adverse effects specific to the individual paper were present in 19 of the 52 records not captured (Table S1). A search strategy which incorporates both generic and specific adverse effects terms could therefore potentially achieve 93% (462/495) recall in the validation set of records.

Comparative analysis

The newly derived Embase search filter achieves marginally higher recall than most of the previous search filters (88% vs. 90%, 84%, 82%, 78% and 75%) (Table 2) when tested against the complete gold standard set of relevant articles. Of the previous search filters, the European Medicines Agency (EMA)—broad search filter performed particularly well on our set of records (90%) and achieved the highest recall. However, the EMA search filter is unlikely to perform as well for precision as any of the other filters as it contains some very broad search terms not necessarily directly related to adverse effects, such as 'exp pregnancy/', 'exp death/' or 'case report?.ti,ab,kw.' (Table S1).

The adverse drug effects search filter derived here using only generic adverse effects terms achieved a higher recall (89%) than the search filter developed for the adverse effects of medical devices (83%) (Golder, Farrah, et al., 2019) but not higher than the surgery complications filter (92%) (Golder et al., 2018) (Table 3).

Our approximate measurements of precision indicate that floating subheadings such as drug effects (de.fs) in MEDLINE and complications (co.fs) in either MEDLINE or Embase are likely to have the lowest precision (Tables S1, S1 and Box S1). Our case examples using the search filters suggest that the search filters reduce the sifting burden for reviewers by around half (Box 1).

Summary

In summary, therefore, the search filters (Box 1 and 2) retrieved 86%, 89%, 90% and 90% of the relevant records in MEDLINE and 90%, 84%, 89% and 89% of the relevant records in Embase (Table 1). In each case, the addition of specific adverse effects terms could have improved the recall of the searches to 90%, 90%, 91% and 92% in MEDLINE and 93%, 88%, 92% and 93% in Embase (Table 1).

DISCUSSION

We have derived and validated search filters for Ovid MEDLINE and Ovid Embase to identify primary studies on adverse drug effects using a cohort of included studies from systematic reviews. This provides the most recent objectively derived search filters to date for adverse effects of drug interventions. We also present an indication of the performance in terms of relative recall of individual search terms and their combinations. The filters will also inevitably increase the precision of searches for adverse drug effects when combined with other terms. Although we were unable to quantify improvement in precision we were able to provide an approximation. Preliminary results from case studies suggest that precision for these filters may be low, but that the filters could still reduce the number needed to screen, in some cases reducing results retrieved by more than half. Further testing of precision using a validated methodology is required to obtain an accurate estimate of precision.

We also identified that some articles only report specific named adverse effects terms without generic adverse effects terms reported in the title, abstract or indexing. This means that searchers should continue to augment any of the generic search filters suggested with terms for specific named adverse effects where appropriate. Our (Table S1) contains a list of some of the specific terms commonly used in the records in our test sets although many others may be applied.

There is always a trade off when searching in terms of recall and precision that can be achieved. Whilst systematic review searches tend to favour high recall and an ideal of 100% this can rarely if ever be achieved. Thus, generally lower levels of recall are deemed acceptable and in reporting our results we adopted a 90% or above threshold as our target (Beynon et al., 2013). We demonstrate that 100% recall is unachievable because nearly 10% of relevant records still do not contain any terms in the title, abstract or indexing to indicate they contain relevant adverse drug effects data thus sometimes examination of the full text will be required. In addition, search filters have always varied in the level of recall that can be achieved. With our newly developed search filters the relative recall of searches using solely generic adverse effects terms was 90% in MEDLINE and 89% in Embase and with the addition of specific adverse effects terms (to the generic adverse effects terms) the recall could be raised to 92% in MEDLINE and 93% in Embase.

The majority of research on adverse effects is in relation to drug interventions and this research is more established (Cohen & Billingsley, 2011; Golder et al., 2016). It may, therefore, be hypothesised that the terminology and indexing of research on the adverse effects of drugs are more consistent and well applied than for the adverse effects of medical devices and surgical procedures. However, our research does not indicate this to necessarily be the case, with all filters performing consistently well achieving between 83% and 92% relative recall irrespective of the type of intervention (drug, device or surgery) (Golder, Farrah, et al., 2019; Golder et al., 2018).

We anticipate that these new search filters will help searchers when developing search strategies to identify relevant studies on adverse drug effects for a systematic review. In addition, we demonstrate the value of

additional named specific adverse effects terms when maximising relative recall. However, we do not recommend these adverse drug effects filters be used without due consideration, particularly as some of the search terms may only apply more to certain medications or have particularly low precision. For example the floating subheading—drug effects in MEDLINE and the textword 'safe*' are likely to give a high number of irrelevant results when conducting adverse effects searches.

LIMITATIONS

The biggest limitation of the present study is the lack of a validated measurement of precision of either the search terms or the combination of search terms. In order to measure precision with the same rigour as relative recall, we would have required a large set of non-relevant records to compare to our set of relevant records. This would have enabled us to identify not just the most frequently occurring relevant terms but also the most discriminating terms.

Another limitation of our study is that our sample of systematic reviews was obtained using search terms for safety in just one database, Epistemonikos. Although we included many synonyms and this is a relatively large database, this may have limited the generalisability of our findings.

The next steps with regard to adverse drug effects search filter development would mostly helpfully consist of testing and validation on systematic review case studies (in which precision could be measured) and further research with larger sample sizes of relevant papers and non-relevant papers.

CONCLUSIONS

This is the first objectively derived search filter for adverse drugs effects in over 15 years. The filter will be most helpful in assisting with searches where unmanageable numbers of records would otherwise be retrieved and additional named specific adverse effects terms can be added to the filter to further increase its relative recall.

Further research on larger datasets is required in order to measure the precision of searching for adverse drug effects and to test the suggested search filters with real case examples.

CONFLICT OF INTEREST

None.

ORCID

Su Golder https://orcid.org/0000-0002-8987-5211

TWITTER

Su Golder @SuGolder

REFERENCES

Badgett, R., Chiquette, E., Anagnostelis, B., & Mulrow, C. (1999). Locating reports of serious adverse drug reactions. In *The best evidence for health care: The role of the Cochrane Collaboration*. Abstracts of the 7th Cochrane Colloquium, 5–9 October 1999, Rome, Italy.

Beynon, R., Leeflang, M. M., McDonald, S., Eisinga, A., Mitchell, R. L., Whiting, P., & Glanville, J. M. (2013). Search strategies to identify diagnostic accuracy studies in MEDLINE and EMBASE. *Cochrane Database of Systematic Reviews*, *2013*(9), Mr000022. https://doi.org/10.1002/14651858.MR000022.pub3

BMJ. (2021). *BMJ best practice: Study design search filters*. BMJ Publishing Group Limited. https://bestpractice.bmj.com/info/toolkit/learn-ebm/study-design-search-filters/

CADTH. (2021). *Strings attached: CADTH's database search filters*. https://www.cadth.ca/strings-attached-cadths-database-search-filters

Canadian Health Libraries Association. (2018). *Canadian Health Libraries Association: MEDLINE Ovid adverse effects search filter*. https://extranet.santecom.qc.ca/wiki/!biblio3s/doku.php?id=concepts:effets-secondaires

Cohen, D., & Billingsley, M. (2011). Europeans are left to their own devices. *BMJ*, *342*, d2748. https://doi.org/10.1136/bmj.d2748

Farrah, K., Mierzwinski-Urban, M., & Cimon, K. (2016). Effectiveness of adverse effects search filters: Drugs versus medical devices. *Journal of the Medical Library Association*, *104*(3), 221–225. https://doi.org/10.3163/1536-5050.104.3.007

Golder, S., Farrah, K., Mierzwinski-Urban, M., Wright, K., & Loke, Y. K. (2019). The development of search filters for adverse effects of medical devices in MEDLINE and EMBASE. *Health Information & Libraries Journal*, *36*(3), 244–263. https://doi.org/10.1111/hir.12260

Golder, S., & Loke, Y. K. (2012a). Failure or success of electronic search strategies to identify adverse effects data. *Journal of the Medical Library Association*, *100*(2), 130–134. https://doi.org/10.3163/1536-5050.100.2.012

Golder, S., & Loke, Y. K. (2012b). The performance of adverse effects search filters in MEDLINE and EMBASE. *Health Information and Libraries Journal*, *29*(2), 141–151. https://doi.org/10.1111/j.1471-1842.2012.00980.x

Golder, S., & Loke, Y. K. (2012c). Sensitivity and precision of adverse effects search filters in MEDLINE and EMBASE: A case study of fractures with thiazolidinediones. *Health Information and Libraries Journal*, *29*(1), 28–38. https://doi.org/10.1111/j.1471-1842.2011.00972.x

Golder, S., Loke, Y. K., Wright, K., & Sterrantino, C. (2016). Most systematic reviews of adverse effects did not include unpublished data. *Journal of Clinical Epidemiology*, *77*, 125–133. https://doi.org/10.1016/j.jclinepi.2016.05.003

Golder, S., McIntosh, H. M., Duffy, S., & Glanville, J. (2006). Developing efficient search strategies to identify reports of adverse

effects in MEDLINE and EMBASE. *Health Information and Libraries Journal*, *23*(1), 3–12. https://doi.org/10.1111/j.1471-1842.2006.00634.x

Golder, S., Peryer, G., & Loke, Y. K. (2019). Overview: Comprehensive and carefully constructed strategies are required when conducting searches for adverse effects data. *Journal of Clinical Epidemiology*, *113*, 36–43. https://doi.org/10.1016/j.jclinepi.2019.05.019

Golder, S., Wright, K., & Loke, Y. K. (2018). The development of search filters for adverse effects of surgical interventions in MEDLINE and Embase. *Health Information & Libraries Journal*, *35*(2), 121–129. https://doi.org/10.1111/hir.12213

Haahr, M. (2021). *True random integer generator*. RANDOM.ORG: True Random Number Service. Randomness and Integrity Services Ltd. https://www.random.org

Hausner, E., Waffenschmidt, S., Kaiser, T., & Simon, M. (2012). Routine development of objectively derived search strategies. *Systematic Reviews*, *1*(1), 19. https://doi.org/10.1186/2046-4053-1-19

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. E. (2019). *Cochrane handbook for systematic reviews of interventions* (2nd ed.). John Wiley & Sons.

Peryer, G., Golder, S., Junqueira, D., Vohra, S., & Loke, Y. K. (2021). Chapter 19: Adverse effects. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (Eds.), *Cochrane Handbook for Systematic Reviews of interventions version 6.2*. Cochrane. www.training.cochrane.org/handbook

Sampson, M., Zhang, L., Morrison, A., Barrowman, N. J., Clifford, T. J., Platt, R. W., Klassen, T. P., & Moher, D. (2006). An alternative to the hand searching gold standard: Validating methodological search filters using relative recall. *BMC Medical Research Methodology*, *6*, 33. https://doi.org/10.1186/1471-2288-6-33

Wieland, S., & Dickersin, K. (2005). Selective exposure reporting and MEDLINE indexing limited the search sensitivity for observational studies of the adverse effects of oral contraceptives. *Journal of Clinical Epidemiology*, *58*(6), 560–567. https://doi.org/10.1016/j.jclinepi.2004.11.018

Wolters Kluwer Health. (2021). *EMA: Embase adverse effects search filter*. Wolters Kluwer Health. https://ospguides.ovid.com/OSPguides/embase.htm, https://www.ema.europa.eu/en/documents/other/monitoring-medical-literature-entry-relevant-information-eudravigilance-database-european-medicines_en.pdf

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Golder, S., Farrah, K., Mierzwinski-Urban, M., Barker, B., & Rama, A. (2022). Updated generic search filters for finding studies of adverse drug effects in Ovid MEDLINE and Embase may retrieve up to 90% of relevant studies. *Health Information & Libraries Journal*, 1–11. https://doi.org/10.1111/hir.12441