International Journal of Technology Assessment in Health Care, 30:4 (2014), 423–429.

doi:10.1017/S0266462314000506

THE CONTRIBUTION OF DIFFERENT INFORMATION SOURCES TO IDENTIFY ADVERSE EFFECTS OF A MEDICAL DEVICE: A CASE STUDY USING A SYSTEMATIC REVIEW OF SPINAL FUSION

Su Golder, Kath Wright, Mark Rodgers
Centre for Reviews and Dissemination (CRD), University of York, York

Background: The most effective sources to search to identify adverse effects data for medical devices are currently unknown.
Methods: The included studies from a systematic review of the safety of recombinant human bone morphogenetic protein-2 (rhBMP-2) for spinal fusion were used for analysis. For each source searched, a record was made for each relevant publication of whether it was retrieved by the search strategy used and whether it was available in the database but not retrieved. To account for multiple publications of the same study, a record was made of the relevant studies identified. The sensitivity, precision, and number needed to read were calculated as well as the minimum combination of sources to identify all the publications or studies.
Results: There were eighty-two publications (forty-nine studies) included in the systematic review. Only one article was available in a database searched but not retrieved by our search strategy. Science Citation Index (SCI) and EMBASE both achieved the highest sensitivity (62 percent), followed closely by MEDLINE/PubMED (56 percent). With the search strategies used, the minimum combination of sources needed to identify all the publications was SCI, EMBASE, CENTRAL, and either MEDLINE or PubMED, in addition to reference checking, contacting authors and an automated current awareness service. In relation to identifying all the relevant studies, the minimum combination of studies was similar with the exclusion of CENTRAL.
Conclusions: To identify all the relevant publications or studies included in this case study systematic review, several different sources needed to be searched.

Keywords: Adverse effects, Systematic review, Meta-analysis, Information storage and retrieval, Bibliographic databases

There are a wide range of medical devices available, ranging from bandages to devices that can improve quality of life (e.g., hip replacements), or even life-saving items such as implantable cardiac defibrillators or life-support machines. However, medical devices such as breast implants and hip prostheses have had adverse publicity in the press recently due to major concerns regarding potential harm (1;2). Hence, comprehensive evaluations of evidence on the safety of medical devices are now an important priority for patients, healthcare professionals as well as policy makers. Although medical devices can have just as serious adverse effects as drugs, the availability of safety data is not on par with drug data (3). The regulatory requirements for research evidence on the safety of new devices are universally less stringent than those for medicines (4). Within Europe, regulations require manufacturers to obtain a CE mark for a new device from any of the many commercial agencies called "notified bodies" to which the European Union delegates the job of certifying medical devices. Unfortunately the amount of clinical evidence needed for CE marking is typically small with low levels of evidence on safety and can vary between "notified bodies" (5). In the United States, medical devices are approved by the FDA (Food and Drug Administration), the same body who approve drug interventions. However, the processes are very different. Unlike drugs, a new device can be approved for use if it is "substantially equivalent" to an existing product and again unlike drugs, with a new device evidence from trials is not necessary. In respect to postmarketing surveillance, a voluntary approach is adopted with medical devices, and device manufacturers can decide which serious side effects they choose to report. In contrast, for new drugs all serious adverse events must be reported by the manufacturer, regardless of their nature or presumed causality (6).

To produce an unbiased evaluation of these adverse effects, systematic reviews should aim to identify as many relevant studies as possible. The selection of sources searched to identify relevant articles will affect which studies are found and ultimately which studies are missed. With limited evidence published on adverse effects of medical devices, it is all the more important that searches for the evidence be as comprehensive as possible, obtaining as much of the scarce data as is feasible. Authors of systematic reviews of adverse effects have tended to focus on

We thank Professor Lesley Stewart and Dr. Yoon Loke for their useful comments and suggestions.

searching MEDLINE and reference checking to identify relevant information (7). However, it is unclear whether such a limited search would identify all the relevant articles with adverse effects data (8).

Previous research on the contribution of different information sources for adverse effects data has focused on pharmaceutical interventions. A systematic review of comparative evaluations of data sources for adverse effects identified nineteen studies, the majority of which were concerned with adverse drug reactions, with none evaluating sources for safety data on medical devices (9). A more recent evaluation of sources also included only adverse drug effects (8). The most efficient combination of sources for information on adverse effects for medical devices is, therefore, currently unknown.

In view of the gap in knowledge, we aimed to evaluate the impact of searching different information sources for adverse effects of a medical device. A case study systematic review of the safety of recombinant human bone morphogenetic protein-2 (rhBMP-2) was selected for analysis (10;11). Here, the rhBMP-2 protein is delivered via a medical device (collagen sponge carrier within a titanium cage) that is widely used as an alternative to iliac crest bone graft to promote fusion in spinal surgery. RhBMP-2 is licensed as a medical device with a pharmaceutical component and as with many other medical devices requires a surgical procedure for implementation.

METHODS

The adverse effects of RhBMP-2 have received much publicity and a high quality systematic review encompassing all the evidence was required to answer the many ambiguities. The Centre for Reviews and Dissemination was commissioned by Yale University to conduct such a review. The systematic review of the safety of rhBMP-2 included a search of BIOSIS Previews (1969–2008 only), CENTRAL, EMBASE, MEDLINE, PubMED, Science Citation Index (SCI), ClinicalTrials.gov, DARE, HTA, and ToxFile. PubMED was selected in addition to MEDLINE as PubMED includes citations not included in MEDLINE as well as the MEDLINE database itself. In addition to database searching, reference checking was undertaken, authors of key papers contacted, a call for evidence was published in Spine Journal, The Back Letter newsletter and on the Internet, and automated "current awareness" searches were set up in Zetoc Alert from the British Library and in MEDLINE to notify us whenever new data were loaded onto the databases. The search strategy contained just two facets; recombinant human bone morphogenetic protein-2 (rhBMP-2) and spinal fusion. Multiple synonyms, text words, and indexing terms were used for each facet. The full search strategy is published elsewhere (10;11)

Inclusion Criteria

All studies (randomized controlled trials [RCTs] and observational studies) of more than ten adult participants that compared rhBMP-2 with any other spinal fusion technique and reported adverse effects were eligible for inclusion in the systematic review.

Analysis

The included references from this case study systematic review formed the basis of the analysis. A record was made of where each of the publications were available and where they were identified. For each publication available on a database but not retrieved by the search strategy, the bibliographic record was examined to determine the reason why it had not been identified. A record was also made of any relevant publications that were retrieved from only one data source.

The sensitivity, precision, and numbers needed to read (NNR) for the searches in each of the databases was calculated using the following definitions;

$$\text{Sensitivity}\,(\%) = \frac{\text{number of included records retrieved} \times 100}{\text{total number of included records}}$$

$$\text{Precision}\,(\%) = \frac{\text{number of included records retrieved} \times 100}{\text{total number of records retrieved}}$$

$$\text{Number Needed to Read (NNR)} = \frac{\text{total number of records retrieved}}{\text{number of included records retrieved}}$$

$$\text{OR Number Needed to Read (NNR)} = 1/\text{precision}$$

In addition, sensitivity*precision was calculated to allow equilibrium between sensitivity and precision to be assessed (12).

RCTs and Observational Studies

The analysis was conducted separately for the included clinical trial publications and observational study publications because certain databases might provide better access to specific study designs, for example, CENTRAL focuses on clinical trials.

Minimum Combination of Sources

The minimum combination of sources required to identify all the included publications using the original search strategies used was recorded. In addition, the minimum number of sources from which all the included publications were available (independent of the search strategy used) was recorded. This analysis was repeated with all RCT publications and all observational study publications separately.

Individual Study Identification

To allow for multiple publications of the same study, the analysis was repeated with all relevant individual studies (as opposed to all relevant publications).

https://doi.org/10.1017/S0266462314000506 Published online by Cambridge University Press

RESULTS

Records Retrieved

A total of 6,807 references were identified from the database searches and 103 additional records were identified from bibliography hand searches and electronic update searches. In addition, the data for seventeen trials were provided by the manufacturer Medtronic Inc.

Included Studies

There were fifteen RCTs, four single arm studies and thirty-five observational studies that were eligible for inclusion in the analysis of the safety of RhBMP-2. Of these studies, three RCTs and three single arm studies were not publically available and were obtained directly from manufacturer Medtronic Inc.

The thirteen published RCTs and one published single arm study were identified in forty publications (eighteen conference abstracts and twenty-two journal publications). The thirty-five observational studies were identified in forty-two publications (thirteen conference abstracts and twenty-nine journal publications). These eighty-two publications (forty-nine studies) formed the basis of the analysis.

Non-database Sources

Seventeen of the eighty-two publications (21 percent) were not identified by any of the standard database searches and were found by either reference checking (fifteen publications), provided by the authors (one publication) or through an automated current awareness service (one publication).

Missing Records

There was only one publication in one database (SCI) that was available at the time of searching but not identified by our search strategy in that database. This article was a conference abstract in SCI and the electronic record did not have an abstract or any keywords assigned in SCI. However, this publication was identified in EMBASE where an abstract was available with terms for rhBMP-2 and cervical fusion, as well as the EMTREE indexing term "bone morphogenetic protein".

Database Sources

BIOSIS Previews was only available to be searched from 1969 to 2008 and so was excluded from the main analysis. No unique publications were identified from this database, so this decision did not impede the analysis of the results. The results from the other databases are presented in Table 1 and Figure 1.

The highest sensitivity for searching for all the publications was achieved in both SCI (62 percent) and EMBASE (62 percent), followed closely by MEDLINE or PubMED (56 percent) (Supplementary Table 1, which can be viewed online at http://dx.doi.org/10.1017/S0266462314000506). Although the sensitivity of searching CENTRAL or ToxFile was low when searching for all publications (26 percent and 21 percent, respectively), CENTRAL performed much better when the analysis was restricted to clinical trial publications (45 percent) and ToxFile (31 percent) performed better when the analysis was restricted to observational study publications (Figure 1 and Table 1).

Unique Publications

The highest number of unique relevant publications was identified through reference checking (ten clinical trial publications and five observational study publications). Unique relevant publications were also identified in SCI (six clinical trial publications and two observational study publications), EMBASE (seven observational study publications), CENTRAL (two clinical trial publications), provided by the authors (one observational study publication) or through automated current awareness service (one observational study publication) (Supplementary Table 1). In addition, if MEDLINE and PubMED are assumed to be one database, one unique observational study publication was identified in MEDLINE/PubMED.

There were, however, six unique journal publications identified by reference checking. If MEDLINE and PubMED are assumed to be one database, there was an additional unique journal publication identified in MEDLINE/PubMED. It is of especial interest that the majority of the uniquely identified publications were conference abstracts 28/34, 82 percent.

Precision and NNR

The highest precision (or lowest NNR) was achieved in CENTRAL 9.63 percent (NNR 10), followed by ToxFile 8.33 percent (NNR 12), and then MEDLINE 5.56 percent (NNR 18) (Supplementary Table 1). However, the precision of all the searches was relatively high for a systematic review (13).

Sensitivity*Precision

The highest sensitivity*precision was achieved by MEDLINE (3.12 percent), followed by CENTRAL (2.47 percent), and SCI (2.44 percent) (Supplementary Table 1).

Database Cost

The relative cost of searching each database is difficult to compare given the differing pricing mechanisms based on license agreements, types of access provision, and type of purchasing organisation (for example, public or private sector). However, PubMed and ToxFile are freely available to all and national provision of access to CENTRAL is available in countries such as the United Kingdom. SCI and EMBASE, both of which performed well, can be prohibitively expensive particularly for individuals or small organisations.

Types of Articles Identified

Conference abstracts formed only a small proportion of the relevant publications identified on MEDLINE or PubMED

Table 1. Database Performance for All Clinical Trial Publications and All Observational Study Publications

		All clinical trial publications			All observational study publications		
Database	Records retrieved	Relevant records retrieved	Sensitivity (%) ($n = 40$)	Precision (%)	Relevant records retrieved	Sensitivity (%) ($n = 42$)	Precision (%)
Science Citation Index (SCI)	1,302	26	65	2	25	62	2
EMBASE	1,542	20	50	1	31	60	2
MEDLINE	827	20	50	2	26	74	2
PubMED	1,176	20	50	2	26	62	3
CENTRAL	218	18	45	8	3	31	6
ToxFile	204	4	10	2	13	7	1

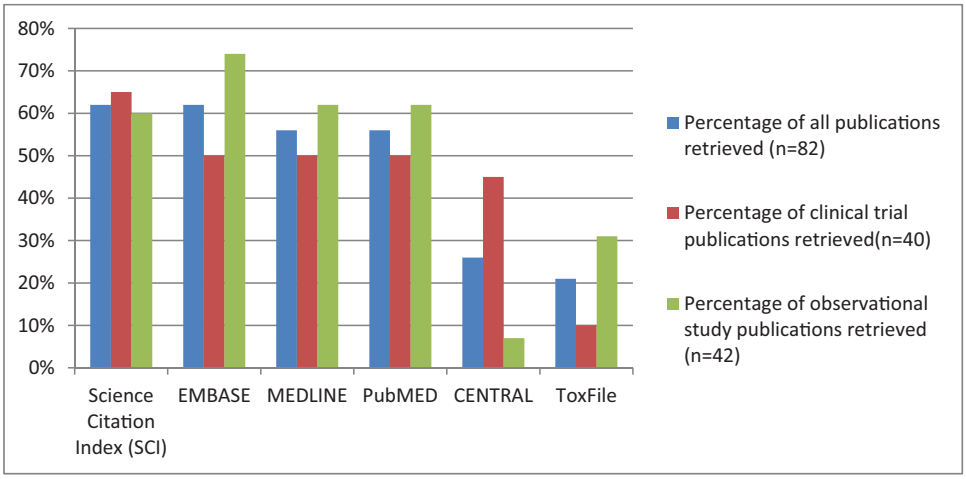

Figure 1. Percentage of relevant publications, clinical trial publications, and observational study publications retrieved by database searches.

as compared to the other databases (Supplementary Table 2, which can be viewed online at http://dx.doi.org/10.1017/S0266462314000506). Reference checking, on the other hand, identified a large proportion of conference abstracts.

Minimum Combination of Sources to Identify All Publications

The minimum combination of sources to identify all the publications was SCI, EMBASE, CENTRAL, and either MEDLINE or PubMED, in addition to reference checking, contacting authors and automated current awareness service.

Minimum Combination of Sources to Identify All RCT Publications

The minimum combination of sources to identify all the RCT publications was SCI, CENTRAL and either EMBASE, MEDLINE, or PubMED, in addition to reference checking.

Minimum Combination of Sources to Identify All Observational Study Publications

The minimum combination of sources to identify all the observational study publications was SCI, EMBASE, and either MEDLINE or PubMED, in addition to reference checking, contacting authors and automated current awareness service.

Individual Study Identification

EMBASE retrieved the highest number of all the relevant studies and the highest number of observational studies whereas SCI retrieved the highest number of clinical trial studies (Supplementary Figure 1, which can be viewed online at http://dx.doi.org/10.1017/S0266462314000506). The rank order of the databases, when the analysis is restricted to individual studies, remains fairly consistent with the analysis of publications identified (Figure 1 and Supplementary Figure 1). However, the sensitivity or recall improves for all the searches when the analysis is restricted to studies as opposed to publications, particularly for clinical trials in CENTRAL and TOXLINE.

Minimum Combination of Sources to Identify All Studies

The minimum combination of sources to identify all the studies was EMBASE, SCI, and either MEDLINE or PubMED, in

https://doi.org/10.1017/S0266462314000506 Published online by Cambridge University Press

addition to reference checking, contacting authors, and automated current awareness service.

DISCUSSION

This case study demonstrates the value of searching multiple sources to identify all the relevant studies or publications with safety data for a medical device. Although this is, to our knowledge, the only evaluation of safety data sources for a medical device, this evaluation can be compared with other studies of search strategies for retrieving adverse effects data (8;9). Previous comparative evaluations on safety data sources have focused on the adverse effects of drug interventions (9). However, there is some similarity in the databases evaluated. Ten evaluations have been carried out which include both MEDLINE and EMBASE (8;14–22) and five of these have also included ToxFile (8;14–17). However, only one comparative evaluation has included SCI (8).

In the ten evaluations which included MEDLINE and EMBASE, there were only two searches in which MEDLINE retrieved more relevant records than EMBASE. Both of these searches were for the only non-drug intervention queries searched (one searched on tooth extraction and the other natural products—aromatherapy/colloidal silver) (19;22).

The rank order of databases in the evaluations in the literature which included at least MEDLINE, EMBASE, and ToxFile and the rank order in the current study are presented in the Supplementary Table 3, which can be viewed online at http://dx.doi.org/10.1017/S0266462314000506. The current study shows a similar pattern to the published literature in that either SCI or EMBASE consistently retrieved the highest number of relevant records and ToxFile the lowest number with the exception of the Biarez et al. 1991 study (15). MEDLINE did not achieve the highest sensitivity in any of these evaluations.

Although ToxFile identified the lowest number of relevant references in almost all the evaluations (Supplementary Table 3), unique relevant references were identified in two of the four evaluations that recorded unique records (15;16).

In line with previous research on retrieving adverse effects data, no single source identified all the relevant publications or studies, and a combination of sources was required to obtain either all the relevant publications or all the relevant studies (8;9) This finding has also been reported previously in evaluations of subject areas other than adverse effects.

A more detailed comparison to a similar evaluation of the contribution of different information sources for adverse drug effects (glitazone related fractures) reveals many similarities (Figure 2) (8). SCI and EMBASE both retrieved a higher yield of relevant publications than MEDLINE, ToxFile, and CENTRAL in this review of spinal fusion and in the review on glitazones. Unsurprisingly, CENTRAL performed much better when the analysis was limited to clinical trials in both reviews.

The percentage of relevant studies retrieved by each database in this evaluation on spinal fusion follows a similar pattern to the glitazone review (Supplementary Figure 2, which can be viewed online at http://dx.doi.org/10.1017/S0266462314000506). In the glitazone review, however, the percentage of all relevant studies identified in each database is much lower. This may reflect the higher number of sources searched in the glitazone review than the spinal fusion review (hence, more relevant studies included outside these core databases) or it may reflect that more studies were missed by the search strategies in the glitazone review than in the spinal fusion review.

Non-bibliographic sources such as reference checking and contacting authors proved useful in this review as they have in other evaluations of information sources for adverse effects (8;9), or evaluations of shoulder pain (23). This is particularly interesting, given that reference checking will almost certainly retrieve more records than recorded as only those references not already retrieved by the searches will be noted by reviewers.

The precision of searches is important in terms of the time and resources that sifting a large number of records entails. In this case study the precision of all the searches was relatively high reflecting the focused nature of the review topic. Other studies have indicated levels of precision to be around 3 percent (7;13).

LIMITATIONS

A major limitation of this analysis is the relatively few sources searched in the original systematic review. This meant that only a few resources could be compared for their relative value in providing relevant data.

This analysis is also only based on one case study systematic review. This limits the generalizability of the results. This was an unusual review in that the authors were able to obtain unpublished data directly from the manufacturer. The included studies also included an unusually high number of conference abstracts and multiple publications for the same study (particularly for the clinical trials).

More research is required with more case study systematic reviews which include more databases and sources searched within each review. Nevertheless, our evaluation is an important first step to guide systematic reviewers in an area which is likely to gain greater prominence, given the heightened public interest in safety of devices following recent scares.

CONCLUSIONS

Several sources need to be searched to identify information on adverse effects of medical devices. The minimum combination of sources to identify all the publications in this case study was SCI, EMBASE, CENTRAL, and either MEDLINE or PubMED, in addition to reference checking, contacting authors, and automated current awareness service.

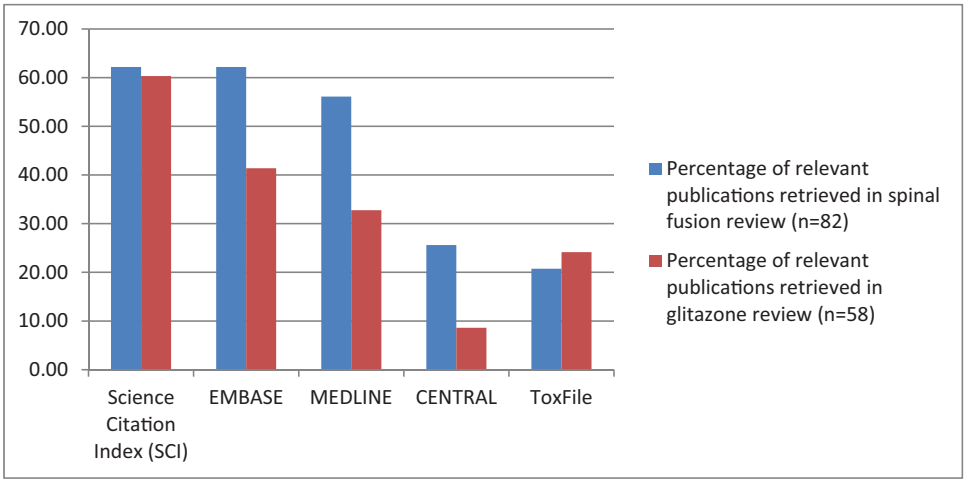

Figure 2. Comparison of percentage of relevant publications retrieved by database searches in spinal fusion review and glitazone review.

Similarities between the core sources required for searching for adverse drug effects and for the adverse effects of medical devices are demonstrated here. The relative value of EMBASE and SCI over MEDLINE, the value of MEDLINE over ToxFile and CENTRAL and the unique contribution of reference checking remains consistent when evaluating the adverse effects of a drug intervention or the adverse effects of a medical device.

SUPPLEMENTARY MATERIAL

Supplementary Table 1:
http://dx.doi.org/10.1017/S0266462314000506
Supplementary Table 2:
http://dx.doi.org/10.1017/S0266462314000506
Supplementary Figure 1:
http://dx.doi.org/10.1017/S0266462314000506
Supplementary Table 3:
http://dx.doi.org/10.1017/S0266462314000506
Supplementary Figure 2:
http://dx.doi.org/10.1017/S0266462314000506

CONTACT INFORMATION

Su Golder (su.golder@york.ac.uk) **Kath Wright, Mark Rodgers,** Centre for Reviews and Dissemination (CRD), University of York, York

CONFLICTS OF INTEREST

None of the authors have any competing interests. S.G. conceived the idea, wrote the protocol, carried out the analysis, and wrote the first draft paper. K.W. conducted the original searches in the systematic review, helped carry out the analysis, and commented and contributed to all stages of the paper. M.R. carried out all stages in the original review, from sifting records, data extraction, and analysis; M.R. also helped ascertain the included studies for this study, and commented and contributed to all versions of the study.

REFERENCES

1. BBC News. *Breast implants: UK to review risk assessment data*. 2011 [updated 2011]. http://www.bbc.co.uk/news/health-16373140 (accessed August 2, 2013).
2. BBC News. *MHRA: Metal hip implant patients need life-long checks*. 2012 [updated 2012]. http://www.bbc.co.uk/news/health-17192520 (accessed August 2, 2013).
3. Cohen D. Europeans are left to their own devices. *BMJ*. 2011;342:d2748.
4. Campbell B, Stainthorpe AC, Longson CM. How can we get high quality routine data to monitor the safety of devices and procedures? *BMJ*. 2013;346:f2782.
5. Cohen D. Faulty hip implant shows up failings of EU regulation. *BMJ*. 2012;345:e7163.
6. Avorn J. Regulation of devices. *BMJ*. 2010;341:c5730.
7. Golder S, Loke YK, Zorzela L. Some improvements are apparent in identifying adverse effects in systematic reviews from 1994 to 2011. *J Clin Epidemiol*. 2013;66:253-260.
8. Golder S, Loke YK. The contribution of different information sources for adverse effects data. *Int J Technol Assess Health Care*. 2012;28:133-137.
9. Golder S, Loke YK. Sources of information on adverse effects: A systematic review. *Health Info Libr J*. 2010;27:176-190.
10. Rodgers MA, Brown JV, Heirs MK, et al. Reporting of industry funded study outcome data: Comparison of confidential and published data on the safety and effectiveness of rhBMP-2 for spinal fusion. *BMJ*. 2013;346:f3981.
11. Simmonds MC, Brown JV, Heirs MK, et al. Safety and effectiveness of recombinant human bone morphogenetic protein-2 for spinal fusion: A meta-analysis of individual-participant data. *Ann Intern Med*. 2013;158:877-889.
12. Tanon A, Champagne F, Contandriopoulos A-P, et al Patient safety and systematic reviews: Finding papers indexed in MEDLINE, EMBASE and CINAHL. *Qual Saf Health Care*. 2010;19:452-461.
13. Sampson M, Tetzlaff J, Urquhart C. Precision of healthcare systematic review searches in a cross-sectional sample. *Res Synth Methods*. 2011;2:119-125.
14. Bagnall AM, Jones L, Glanville J, Kleijnen J, eds. *Assessing adverse events in a systematic review of atypical antipsychotics for schizophrenia*. 4th Symposium on Systematic Reviews: Pushing the Boundaries, July 2002; Oxford, UK.
15. Biarez O, Sarrut B, Doreau CG, Etienne J. Comparison and evaluation of nine bibliographic databases concerning adverse drug reactions. *DICP*. 1991;25:1062-1065.

16. Golder S, McIntosh HM, Duffy S, Glanville J. Developing efficient search strategies to identify reports of adverse effects in MEDLINE and EMBASE. *Health Info Libr J*. 2006;23:3-12.
17. Madden M, MacDonald A. An evaluation and comparison of nine drug information retrieval services. *Drug Inf J*. 1977;11:47-59.
18. Sodha RV, Van Amelsvoort T. Multi-database searches in biomedicine: Citation duplication and novelty assessment using carbamazepine as an example. *J Info Sci*. 1994;20:139-141.
19. Stone VL, Fishman DL, Frese DB. Searching online and Web-based resources for information on natural products used as drugs. *Bull Med Libr Assoc*. 1998;86:523-527.
20. Scientific. T. *Derwent Drug File: Definitive drug journal and conference information.* Thomson Scientific; 2004 [updated 2004]. http://scientific.thomson.com/media/dw/productpdfs/ddf-compare.pdf (accessed September 20, 2008).
21. Van Putte N. A comparison of four biomedical databases for the retrieval of drug literature. *Health Inf lib*. 1991;3:119-127.
22. Verheijen-Voogd C, Mathijsen A. A contribution to the comparison of the usefulness of the data bases of Excerpta Medica and MEDLARS in biomedical literature retrieval. *Aslib Proc*. 1974;26:136-151.
23. Beyer F, Wright K. Can we prioritise which databases to search? A case study using a systematic review of frozen shoulder management. *Health Info Libr J*. 2012;30:49-58.

https://doi.org/10.1017/S0266462314000506 Published online by Cambridge University Press