



Cochrane
Library

Cochrane Database of Systematic Reviews

Methods for obtaining unpublished data (Review)

Young T, Hopewell S

Young T, Hopewell S.
Methods for obtaining unpublished data.
Cochrane Database of Systematic Reviews 2011, Issue 11. Art. No.: MR000027.
DOI: [10.1002/14651858.MR000027.pub2](https://doi.org/10.1002/14651858.MR000027.pub2).

www.cochranelibrary.com

TABLE OF CONTENTS

ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
BACKGROUND	3
OBJECTIVES	3
METHODS	3
RESULTS	4
Figure 1.	5
DISCUSSION	7
AUTHORS' CONCLUSIONS	8
ACKNOWLEDGEMENTS	8
REFERENCES	9
CHARACTERISTICS OF STUDIES	10
ADDITIONAL TABLES	15
APPENDICES	16
CONTRIBUTIONS OF AUTHORS	16
DECLARATIONS OF INTEREST	17
SOURCES OF SUPPORT	17
INDEX TERMS	17

[Methodology Review]

Methods for obtaining unpublished data

Taryn Young¹, Sally Hopewell²

¹Centre for Evidence-based Health Care, Faculty of Health Sciences, Stellenbosch University, Tygerberg, South Africa. ²UK Cochrane Centre, Oxford, UK

Contact: Taryn Young, Centre for Evidence-based Health Care, Faculty of Health Sciences, Stellenbosch University, PO Box 19063, Tygerberg, 7505, South Africa. tyoung@sun.ac.za.

Editorial group: Cochrane Methodology Review Group.

Publication status and date: New, published in Issue 11, 2011.

Citation: Young T, Hopewell S. Methods for obtaining unpublished data. *Cochrane Database of Systematic Reviews* 2011, Issue 11. Art. No.: MR000027. DOI: [10.1002/14651858.MR000027.pub2](https://doi.org/10.1002/14651858.MR000027.pub2).

Copyright © 2011 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

In order to minimise publication bias, authors of systematic reviews often spend considerable time trying to obtain unpublished data. These include data from studies conducted but not published (unpublished data), as either an abstract or full-text paper, as well as missing data (data available to original researchers but not reported) in published abstracts or full-text publications. The effectiveness of different methods used to obtain unpublished or missing data has not been systematically evaluated.

Objectives

To assess the effects of different methods for obtaining unpublished studies (data) and missing data from studies to be included in systematic reviews.

Search methods

We identified primary studies comparing different methods of obtaining unpublished studies (data) or missing data by searching the Cochrane Methodology Register (Issue 1, 2010), MEDLINE and EMBASE (1980 to 28 April 2010). We also checked references in relevant reports and contacted researchers who were known or who were thought likely to have carried out relevant studies. We used the Science Citation Index and PubMed 'related articles' feature to identify any additional studies identified by other sources (19 June 2009).

Selection criteria

Primary studies comparing different methods of obtaining unpublished studies (data) or missing data in the healthcare setting.

Data collection and analysis

The primary outcome measure was the proportion of unpublished studies (data) or missing data obtained, as defined and reported by the authors of the included studies. Two authors independently assessed the search results, extracted data and assessed risk of bias using a standardised data extraction form. We resolved any disagreements by discussion.

Main results

Six studies met the inclusion criteria; two were randomised studies and four were observational comparative studies evaluating different methods for obtaining missing data.

Methods to obtain missing data

Five studies, two randomised studies and three observational comparative studies, assessed methods for obtaining missing data (i.e. data available to the original researchers but not reported in the published study).

Methods for obtaining unpublished data (Review)

Copyright © 2011 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Two studies found that correspondence with study authors by e-mail resulted in the greatest response rate with the fewest attempts and shortest time to respond. The difference between the effect of a single request for missing information (by e-mail or surface mail) versus a multistage approach (pre-notification, request for missing information and active follow-up) was not significant for response rate and completeness of information retrieved (one study). Requests for clarification of methods (one study) resulted in a greater response than requests for missing data. A well-known signatory had no significant effect on the likelihood of authors responding to a request for unpublished information (one study). One study assessed the number of attempts made to obtain missing data and found that the number of items requested did not influence the probability of response. In addition, multiple attempts using the same methods did not increase the likelihood of response.

Methods to obtain unpublished studies

One observational comparative study assessed methods to obtain unpublished studies (i.e. data for studies that have never been published). Identifying unpublished studies ahead of time and then asking the drug industry to provide further specific detail proved to be more fruitful than sending of a non-specific request.

Authors' conclusions

Those carrying out systematic reviews should continue to contact authors for missing data, recognising that this might not always be successful, particularly for older studies. Contacting authors by e-mail results in the greatest response rate with the fewest number of attempts and the shortest time to respond.

PLAIN LANGUAGE SUMMARY

Methods for obtaining unpublished data

This methodology review was conducted to assess the effects of different methods for obtaining unpublished studies (data) and missing data from studies to be included in systematic reviews. Six studies met the inclusion criteria, two were randomised studies and four were observational comparative studies evaluating different methods for obtaining missing data.

Five studies assessed methods for obtaining missing data (i.e. data available to the original researchers but not reported in the published study). Two studies found that correspondence with study authors by e-mail resulted in the greatest response rate with the fewest attempts and shortest time to respond. The difference between the effect of a single request for missing information (by e-mail or surface mail) versus a multistage approach (pre-notification, request for missing information and active follow-up) was not significant for response rate and completeness of information retrieved (one study). Requests for clarification of methods (one study) resulted in a greater response than requests for missing data. A well-known signatory had no significant effect on the likelihood of authors responding to a request for unpublished information (one study). One study assessed the number of attempts made to obtain missing data and found that the number of items requested did not influence the probability of response. In addition, multiple attempts using the same methods did not increase the likelihood of response.

One study assessed methods to obtain unpublished studies (i.e. data for studies that have never been published). Identifying unpublished studies ahead of time and then asking the drug industry to provide further specific detail proved to be more fruitful than sending of a non-specific request.

Those carrying out systematic reviews should continue to contact authors for missing data, recognising that this might not always be successful, particularly for older studies. Contacting authors by e-mail results in the greatest response rate with the fewest number of attempts and the shortest time to respond.

BACKGROUND

Description of the problem or issue

Reporting bias arises when dissemination of research findings is influenced by the nature and direction of results. Publication bias (the selective publication of research studies as a result of the strength of study findings), time-lag bias (the rapid or delayed publication of results depending on the results) and language bias (the publication in a particular language depending on the nature and direction of the results) are typical types of reporting bias (Higgins 2009).

Publication bias, especially, is a major threat to the validity of systematic reviews (Song 2000; Sterne 2008). Hopewell et al examined the impact of grey literature (literature which has not formally been published) in meta-analysis of randomised trials of healthcare interventions and found that published trials tend to be larger and show an overall greater treatment effect than trials from grey literature (Hopewell 2007). Not making an attempt to include unpublished data in a systematic review can thus result in biased larger treatment effects (Higgins 2009).

In order to minimise publication bias, authors of systematic reviews often spend considerable time trying to obtain unpublished data. These include data from studies conducted but not published (unpublished data) as either an abstract or full-text paper, as well as missing data (data available to the original researchers but not reported) in published abstracts or full-text publications. Types of data commonly missing from published papers include details of allocation concealment and blinding, information about loss to follow-up and standard deviations. This is different from data that are 'missing' because the original researchers do not have them, but might be able to get them (e.g. a specific subgroup analysis not done by the original researchers, but which could be carried out retrospectively in response to a request from systematic review authors) or data that are missing because they were never collected by the original researchers and are not retrievable by other means (e.g. patient's quality of life at specific time points).

Often, the search for and retrieval of unpublished and missing data delays the time to review completion.

Description of the methods being investigated

Different methods are used to search for and obtain unpublished data or missing data from studies to be included in systematic reviews.

Authors of systematic reviews informally contact colleagues to find out if they know about unpublished studies (Greenhalgh 2005). In addition, formal requests for information on completed but unpublished studies, as well as ongoing studies, are sent to researchers (authors of identified included studies of the relevant review), experts in the field, research organisations and pharmaceutical companies (Lefebvre 2008; Song 2000). Some organisations might set up websites for systematic review projects, listing the studies identified to date and inviting submission of information on studies not already listed (Lefebvre 2008).

Prospective clinical trial registries, both national and international, are also searched to identify ongoing studies. Plus, registries of grey literature are searched to identify unpublished studies.

In order to obtain details about missing data (data available to the original researchers but not reported) authors of systematic reviews contact the authors of studies included in the review by telephone, e-mail or letters by post.

How these methods might work

Approaching researchers for information about completed but never published studies has had varied results, ranging from being willing to share information to no response (Greenhalgh 2005). The willingness of investigators of located unpublished studies to provide data may depend upon the findings of the study, where more favourable results may be shared more willingly (Smith 1998).

Why it is important to do this review

The effectiveness of the different methods used to obtain unpublished or missing data has not been systematically evaluated. This review will systematically evaluate these effects and will thus assist authors of reviews in improving their efficiency in conducting their reviews.

OBJECTIVES

To assess the effects of different methods for obtaining unpublished studies (data) and missing data from studies to be included in systematic reviews.

METHODS

Criteria for considering studies for this review

Types of studies

Primary studies comparing different methods of obtaining unpublished studies (data) or missing data. We excluded studies without a comparison of methods.

Types of data

All relevant studies in the healthcare setting.

Types of methods

Any method designed to obtain unpublished studies (data) or missing data (i.e. data available to researchers but not reported).

Types of outcome measures

Primary outcomes

Methods to obtain missing data (data available to researchers but not reported in the published study).

- Proportion of missing data obtained as defined and reported by authors.

Methods to obtain unpublished studies (data for studies that have never been published).

- Proportion of unpublished studies (data) obtained as defined and reported by authors.

Secondary outcomes

Methods to obtain missing data (data available to the original researchers but not reported in the published study).

- Completeness (extent to which data obtained answers to the questions posed by those seeking the data) of missing data obtained.
- Type (e.g. outcome data, baseline data) of missing data obtained.
- Time taken to obtain missing data (i.e. time from when efforts start until data are obtained).
- Number of attempts (as defined by the authors) made to obtain missing data.
- Resources required.

Methods to obtain unpublished studies (data for studies that have never been published).

- Time taken to obtain unpublished studies (i.e. time from when efforts start until data are obtained).
- Number of attempts (as defined by the authors) made to obtain unpublished studies (data).
- Resources required.

Search methods for identification of studies

To identify studies we carried out both electronic and manual searches. All languages were included.

Electronic searches

We searched the Cochrane Methodology Register (CMR) (Issue 1, 2009) using the search terms in [Appendix 1](#). We searched MEDLINE and Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations using OVID (1950 to 10 February 2009) ([Appendix 2](#)) and adapted these terms for use in EMBASE (1980 to 2009 Week 06) ([Appendix 3](#)). We conducted an updated search in EMBASE, MEDLINE and the Cochrane Methodology Register on 28 April 2010.

Searching other resources

We also checked references in relevant reports ([Horsley 2011](#)) and contacted researchers who were known or who were thought likely to have carried out relevant studies. We used the Science Citation Index and PubMed 'related articles' feature to identify any additional studies identified by the sources above (19 June 2009).

Data collection and analysis

Selection of studies

Two authors independently screened titles, abstracts and descriptor terms of the electronic search results for relevance based on the criteria for considering studies for this review. We obtained full-text articles (where available) of all selected abstracts and used an eligibility form to determine final study selection. We resolved any disagreements through discussion.

Data extraction and management

Two authors independently extracted data using a standardised data extraction form. We resolved any disagreements by discussion.

Data extracted included the following.

- Administrative details for the study - identification; author(s); published or unpublished; year of publication; year in which study was conducted; details of other relevant papers cited.

- Details of study - study design; inclusion and exclusion criteria; country and location of the study.
- Details of intervention - method(s) used to obtain unpublished or missing data.
- Details of outcomes and results - proportion, completeness and type of unpublished or missing data; time taken to obtain unpublished or missing data; number of attempts made and resources required.

Assessment of risk of bias in included studies

Two authors independently evaluated the risk of bias of included studies, which included both the inherent properties of the study and the adequacy of its reporting.

For randomised studies comparing different methods to obtain data we assessed the following criteria, based on The Cochrane Collaboration's 'Risk of bias' tool and classified as adequate, inadequate or unclear:

- generation of the allocation sequence;
- concealment of the allocation sequence;
- blinding of the participants, personnel and outcome assessor.

For non-randomised studies comparing different methods to obtain data we assessed the following criteria and reported whether they were adequate, inadequate or unclear:

- how allocation occurred;
- attempt to balance groups by design;
- use of blinding.

Based on these criteria, we assessed studies as being at 'high', 'low' or 'moderate' risk of bias.

Dealing with missing data

If any of the data were insufficient or missing, we sought data from the contact author of the empirical study using e-mail. This was successful for one of the two studies ([Higgins 1999](#)) for which we contacted the authors.

Data synthesis

Due to significant differences in study design, it was not possible to carry out a meta-analysis of the included studies. Therefore the results of the individual studies are presented descriptively, reporting individual study effect measures and 95% confidence intervals where available.

RESULTS

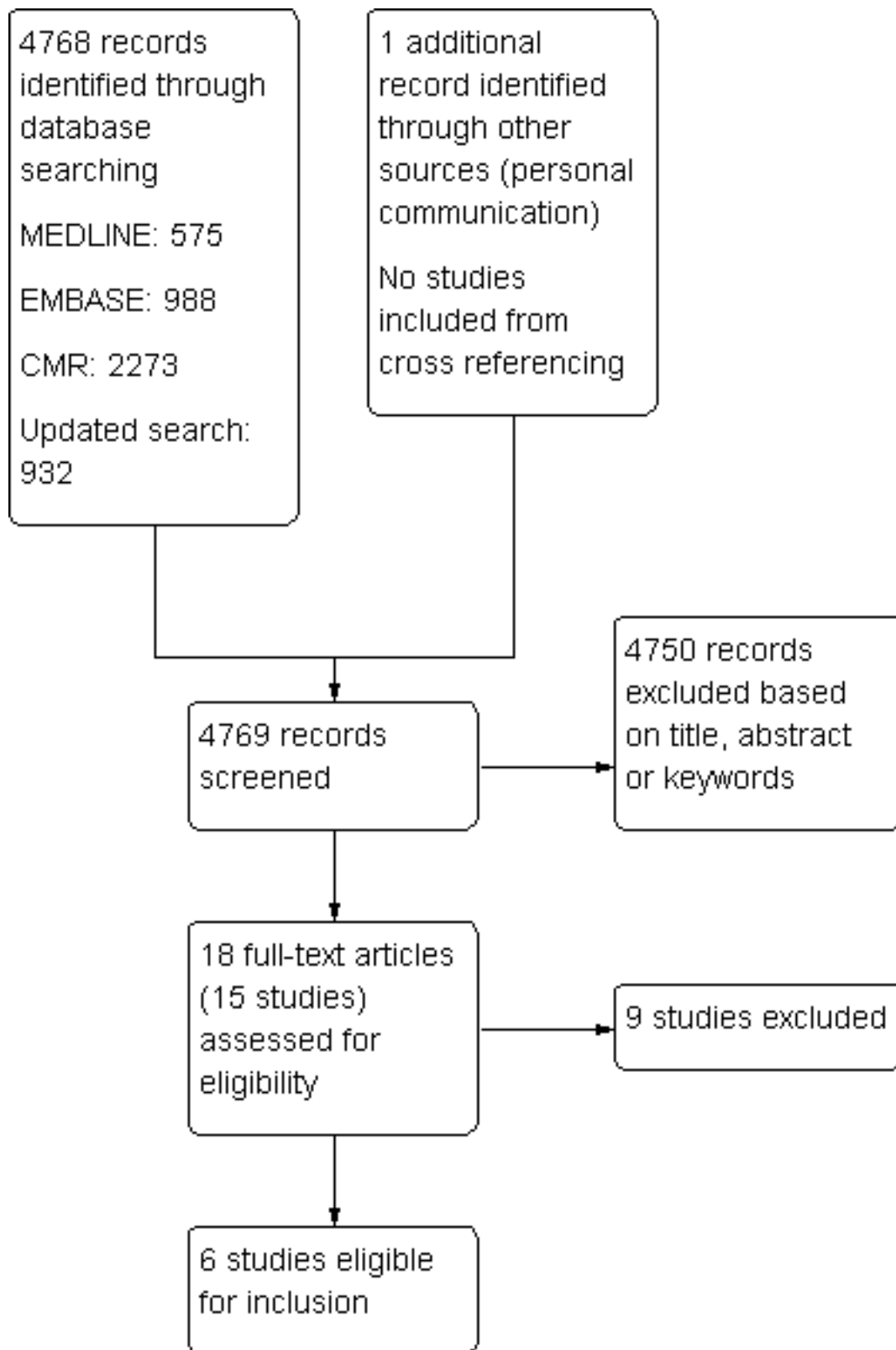
Description of studies

See [Characteristics of included studies](#) and [Characteristics of excluded studies](#).

Results of the search

Of 4768 identified abstracts and titles, we selected 18 potentially eligible publications, referring to 15 studies, for detailed independent eligibility assessment ([Figure 1](#)).

Figure 1. Study flow diagram.



Included studies

Six studies (Brown 2003; Gibson 2006; Guevara 2005; Higgins 1999; Milton 2001; Shukla 2003) met the inclusion criteria. Of these five were published as abstracts for Cochrane Colloquia and one as a full paper (Gibson 2006). Only one observational comparative study evaluated the effects of different methods to obtain unpublished data (Shukla 2003). The other five studies, two randomised studies (Higgins 1999; Milton 2001) and three observational comparative studies (Brown 2003; Gibson 2006; Guevara 2005), evaluated different methods for obtaining missing data. Table 1 provides a summary of the interventions studies and outcomes measured.

Excluded studies

Nine studies (Bohlius 2003; Eysenbach 2001; Hadhazy 1999; Hetherington 1987; Kelly 2002; Kelly 2004; McGrath 1998; Reveiz 2004; Wille-Jorgensen 2001) did not meet the inclusion criteria as there was no comparison of different methods of obtaining missing data.

Risk of bias in included studies

Brown 2003, Gibson 2006, Guevara 2005 and Shukla 2003, the four observational, comparative studies, did not report on the methodology used and therefore assessments of risk of bias for these studies are incomplete.

We assessed risk of bias for the two randomised studies by looking at the methods used for allocation sequence generation, allocation concealment and blinding. Allocation concealment was adequate for Higgins 1999 and unclear for Milton 2001. Allocation sequence generation and blinding were not reported.

Effect of methods

Methods to obtain missing data

Five of the six studies assessed methods for obtaining missing data (i.e. data available to the original researchers but not reported in the published study).

Proportion of missing data obtained as defined and reported by authors

All five studies provided information on the proportion of missing data obtained.

Brown 2003 used a non-randomised design to compare contacting 112 authors (of 139 studies) via 39 e-mails and 73 letters. The study was designed as a comparative study but data per study arm were not reported. Twenty-one replies (19%) were received. One study published in the period 1980-1984 elicited no response, nine 1985-1989 studies elicited two responses, 41 1990-1994 studies elicited six responses, 38 1995-1999 studies elicited eight responses and 21 2000-2002 studies elicited four responses.

Gibson 2006 used a non-randomised design to compare contacting authors by e-mail, letter or both. Two hundred and forty-one studies (40%) had missing or incomplete data. They were unable to locate 95 authors (39%). Of the remaining 146 authors, 46 authors (32%) responded to information requests. The response rate differed by mode of contact - letter (24%), e-mail (47%) and both (73%). Response was significantly higher with e-mail compared to using letters (hazard ratio 2.5; 95% confidence interval (CI) 1.3 to 4.0). Combining letter and e-mail had a higher response

rate, however, it was not significantly different from using e-mail alone (reported $P = 0.36$). The combination of methods (letter plus e-mail follow-up) rather than multiple contacts using the same method was more effective for eliciting a response from the author. Response rates from US authors did not differ from those of other countries. The older the article, the less likely the response.

Guevara 2005 used a non-randomised design to compare e-mail versus letter versus fax. Fifteen authors (60%) responded to information requests. E-mail resulted in fewer attempts and a greater response rate than post or fax. Authors of studies published after 1990 were as likely to respond (67% versus 50%, reported $P = 0.45$) as authors of studies published earlier. Similarly, corresponding authors were no more likely to respond (58% versus 9%, reported $P = 0.44$) than secondary authors, although few secondary authors were contacted.

Higgins 1999 used a randomised comparison of single request for missing information (by e-mail or surface mail) ($n = 116$) versus a multistage approach involving pre-notification, request for missing information and active follow-up ($n = 117$) and found no significant difference between the two groups (risk ratio (RR) 1.04; 95% CI 0.74 to 1.45) in response rate.

Milton 2001 compared, using a randomised design, the response of clinical trial investigators to requests for information signed by either Richard Smith (RS), editor of the *British Medical Journal* ($n = 96$), or an unknown researcher ($n = 48$) and found no significant differences between signatory groups in response rates. By three weeks, 34% in the former and 27% in the unknown researcher's group had responded (odds ratio (OR) 1.35; 95% CI 0.59 to 3.11). No baseline data had been provided by three weeks. By the end of the study, at five weeks, 74% and 67% respectively had responded (OR 1.42; 95% CI 0.62 to 3.22) and 16 out of 53 studies in the RS group and five out of 27 authors in the unknown researcher's group had provided baseline data (OR 1.90; CI 0.55 to 6.94).

Completeness of data

One of the five studies assessed the extent to which data obtained answers to the questions posed by those seeking the data.

Higgins 1999 compared, using a randomised design, the completeness of information retrieved between study arms (single request for missing information (by e-mail or surface mail) ($n = 116$) versus multistage approach involving pre-notification, request for missing information and active follow-up ($n = 117$)) and found no significant difference between the two study methods.

Type of missing data obtained

Two of the five studies assessed the type of missing data obtained.

Brown 2003 used a non-randomised design to compare contacting 112 authors (of 139 studies) via 39 e-mails and 73 letters and received 21 replies (19%), of which nine provided relevant outcome and quality data, one provided additional data on study quality only and one provided information regarding duplicate publications. Eleven studies provided no useful information. Data per study arm were not reported.

Guevara 2005 used a non-randomised design to compare e-mail versus letter versus fax and reported that requests for clarification of methods resulted in a greater response (50% versus 32%, $P =$

Methods for obtaining unpublished data (Review)

0.03) than requests for missing data. Once again, data per study arm were not reported.

Time taken to obtain missing data

Two of the five studies assessed the time taken to obtain missing data (i.e. time from when efforts start until data are obtained).

[Gibson 2006](#) used a non-randomised design to compare e-mail versus letter versus fax and reported that the time to respond differed significantly by contact method ($P < 0.05$): e-mail (3 +/- 3 days; median one day), letter (27 +/- 30 days; median 10 days) and both (13 +/- 12 days; median nine days).

[Guevara 2005](#) used a non-randomised design to compare e-mail versus letter versus fax and reported that e-mail had a shorter response time than post or fax.

Number of attempts made to obtain missing data

One of the five studies assessed the number of attempts made to obtain missing data.

[Gibson 2006](#) used a non-randomised design to compare e-mail versus letter versus fax and reported that the number of items requested per authors averaged two or more. The number of items requested did not influence the probability of response. In addition, multiple attempts using the same methods did not increase the likelihood of response.

Resources required

One of the five studies assessed the resources required to obtain missing data.

[Brown 2003](#) used a non-randomised design to compare contacting 112 authors (of 139 studies) via 39 e-mails and 73 letters and reported total costs of 80 GBP for printing and postage. Cost was not reported per study arm.

Methods to obtain unpublished studies

One of the six included studies assessed methods to obtain unpublished studies (i.e. data for studies that have never been published).

Proportion of unpublished studies (data) obtained as defined and reported by authors

[Shukla 2003](#), using a non-randomised design, assessed two different approaches to seek unpublished information from the drug industry. The outcome of a general request letter was compared with efforts to identify unpublished data and then contacting the industry to provide further specific detail. With the first approach, no unpublished information was obtained. With the second approach, relevant unpublished information was obtained for four of five of the systematic reviews (in the form of manuscripts or oral/poster presentations).

No information was available for the following secondary outcome measures.

- Time taken to obtain unpublished studies (i.e. time from when efforts start until data are obtained).
- Number of attempts (as defined by the authors) made to obtain unpublished studies (data).

- Resources required.

DISCUSSION

Despite extensive searches we identified only six studies as eligible for inclusion in this review. Of these, five were published as abstracts and one as a full paper. Due to lack of high-quality studies the results should be interpreted with caution. Five studies, two randomised studies and three observational comparative studies evaluated different methods for obtaining missing data (e.g. data available to the original researchers but not reported in the published study). Two studies found that correspondence with study authors by e-mail resulted in the greatest response rate with the fewest number of attempts and the shortest time to respond, when compared with correspondence by fax or letter. Combining letter and e-mail had a higher response rate, however, it was not significantly different from using e-mail alone. Another study found that you were more likely to solicit a response from authors whose studies were published more recently. In addition, requests for clarification of the study methods appeared to result in a greater response rate than requests for missing data about the study results.

The effect of a single request for missing information (by e-mail or surface mail) versus a multistage approach (pre-notification, request for missing information and active follow-up) did not appear to affect the rate of response or the completeness of information retrieved; neither did the number of attempts made to obtain missing data or the number of items requested. Interestingly, the use of a well-known signatory also had no significant effect on the likelihood of authors responding to a request for unpublished information. Only one study evaluated the effects of different methods to obtain unpublished data (e.g. data for studies that have never been published). This found that leg-work ahead of time to clarify and request the specific unpublished study information required can prove to be more fruitful than sending of a non-specific request. The *Cochrane Handbook for Systematic Reviews of Interventions* ([Higgins 2009](#)) suggests that review authors also consider contacting colleagues to find out if they are aware of any unpublished studies; we did not find any studies addressing the effectiveness of this approach.

When considering the findings from this review it is important to consider the limitations in the completeness of the available data and how this weakens the strength of any recommendations we are able to draw. The general problem that a large proportion of conference abstracts do not get published in full has been shown by others ([Scherer 2007](#)) and it was recently found that about two-thirds of the research studies presented at Cochrane Colloquia do not get published in full ([Chapman 2010](#)). We encountered this problem in this review with five of the six studies being available only as abstracts at Colloquia. They lacked information about the study methodology and detailed results, and were never written up and published in full. Despite attempts to contact the authors of these studies we were only able to obtain additional information for one of the five studies. Ironically, our systematic review is subject to the same problems of obtaining missing data which our review is trying to address. Assessment of risk of bias was also hampered by incomplete data; the four observational studies did not report on the study methods and only one of the two randomised studies reported on the method of allocation concealment. The [Brown 2003](#) study was designed as a comparative study, however only

combined results were reported. The study is therefore reported in this review as though it was a non-comparative study report of the experience of contacting original authors.

Missing and incomplete data continue to be a major problem and potential source of bias for those carrying out systematic reviews. If data were missing from study reports at random then there would be less information around but that missing information would not necessarily be biased. The problem is that there is considerable evidence showing that studies are more likely to be published and published more quickly if they have significant findings (Scherer 2007). Even when study results are published, there is evidence to show that authors are more likely to report significant study outcomes as opposed to non-significant study outcomes (Kirkham 2010). The findings from our review support the current recommendations in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2009) that those carrying out systematic reviews should continue to contact authors for missing data, recognising that this might not always be successful, particularly for older studies. In the absence of being able to contact authors to obtain missing data, review authors should also consider the potential benefits of searching prospective clinical trial registries and trial results registers for missing data. For example, in 2007 the US government passed legislation that the

findings for all US government funded research should be included on www.clinicaltrials.gov within one year of study completion, thus making available previously unpublished information. The setting up of websites for systematic review projects, listing the studies identified to date and inviting submission of information on studies not already listed (Lefebvre 2008), has also been proposed as a way of identifying unpublished studies.

AUTHORS' CONCLUSIONS

Implication for methodological research

The strength of the evidence included in this review is limited by the completeness of the available data; five of the six studies included in this review lacked information about the study methodology and their results. Despite extensive searching only one study assessed methods for obtaining unpublished data. Further robust, comparative, well-conducted and reported studies are needed on strategies to obtain missing and unpublished data.

ACKNOWLEDGEMENTS

We are very grateful to Julian Higgins who provided us with additional information regarding the First Contact study.

REFERENCES

References to studies included in this review

Brown 2003 {published data only}

Brown T, Hooper L. Effectiveness of brief contact with authors. XI Cochrane Colloquium: Evidence, Health Care and Culture; 2003 Oct 26-31; Barcelona, Spain. 2003.

Gibson 2006 {published data only}

Gibson CA, Bailey BW, Carper MJ, Lecheminant JD, Kirk EP, Huang G, et al. Author contacts for retrieval of data for a meta-analysis on exercise and diet restriction. *International Journal of Technology Assessment in Health Care* 2006;**22**(2):267-70.

Guevara 2005 {published data only}

Guevara J, Keren R, Nihtianova S, Zorc J. How do authors respond to written requests for additional information?. XIII Cochrane Colloquium; 2005 Oct 22-26; Melbourne, Australia. 2005.

Higgins 1999 {published data only}

Higgins J, Soornro M, Roberts I, Clarke M. Collecting unpublished data for systematic reviews: a proposal for a randomised trial. 7th Annual Cochrane Colloquium Abstracts, October 1999 in Rome. 1999.

Milton 2001 {published data only}

Milton J, Logan S, Gilbert R. Well-known signatory does not affect response to a request for information from authors of clinical trials: a randomised controlled trial. 9th Annual Cochrane Colloquium Abstracts, October 2001 in Lyon. 2001.

Shukla 2003 {published data only}

Shukla V. The challenge of obtaining unpublished information from the drug industry. XI Cochrane Colloquium: Evidence, Health Care and Culture; 2003 Oct 26-31; Barcelona, Spain. 2003.

References to studies excluded from this review

Bohlius 2003 {published data only}

Bohlius J, Langensiepen S, Engert A. Data hunting: a case report. XI Cochrane Colloquium: Evidence, Health Care and Culture; 2003 Oct 26-31; Barcelona, Spain. 2003.

Eysenbach 2001 {published data only}

* Eysenbach G, Tuische J, Diepgen TL. Evaluation of the usefulness of Internet searches to identify unpublished clinical trials for systematic reviews. *Medical Informatics and the Internet in Medicine* 2001;**26**(3):203-18.

Eysenbach G, Tuische J, Diepgen TL. Evaluation of the usefulness of internet searches to identify unpublished clinical trials for systematic reviews. *Chinese Journal of Evidence-Based Medicine* 2002;**2**(3):196-200.

Hadhazy 1999 {published data only}

Hadhazy V, Ezzo J, Berman B. How valuable is effort to contact authors to obtain missing data in systematic reviews. 7th

Annual Cochrane Colloquium Abstracts, October 1999 in Rome. 1999.

Hetherington 1987 {published data only}

* Hetherington J. An international survey to identify unpublished and ongoing perinatal trials [abstract]. *Controlled Clinical Trials* 1987;**8**:287.

Hetherington J, Dickersin K, Chalmers I, Meinert CL. Retrospective and prospective identification of unpublished controlled trials: lessons from a survey of obstetricians and pediatricians. *Pediatrics* 1989;**84**(2):374-80.

Kelly 2002 {published data only}

Kelley GA, Kelley KS, Tran ZV. Retrieval of individual patient data for an exercise-related meta-analysis. *Medicine & Science in Sports & Exercise* 2002;**34**(5 (Suppl 1)):S225.

Kelly 2004 {published data only}

Kelley GA, Kelley KS, Tran ZV. Retrieval of missing data for meta-analysis: a practical example. *International Journal of Technology Assessment in Health Care* 2004;**20**(3):296-9.

McGrath 1998 {published data only}

McGrath J, Davies G, Soares K. Writing to authors of systematic reviews elicited further data in 17% of cases. *BMJ* 1998;**316**:631.

Revez 2004 {published data only}

* Revez L, Andres Felipe C, Egdar Guillermo O. Using e-mail for identifying unpublished and ongoing clinical trials and those published in non-indexed journals. 12th Cochrane Colloquium: Bridging the Gaps; 2004 Oct 2-6; Ottawa, Ontario, Canada. 2004.

Revez L, Cardona AF, Ospina EG, de Agular S. An e-mail survey identified unpublished studies for systematic reviews. *Journal of Clinical Epidemiology* 2006;**59**(7):755-8.

Wille-Jorgensen 2001 {published data only}

Wille-Jorgensen. Problems with retrieving original data: is it a selection bias?. 9th Annual Cochrane Colloquium, Lyon. October 2001.

Additional references

Chapman 2010

Chapman S, Eisinga A, Clarke MJ, Hopewell S. Passport to publication? Do methodologists publish after Cochrane Colloquia?. Joint Cochrane and Campbell Colloquium. 2010 Oct 18-22; Keystone, Colorado, USA. Cochrane Database of Systematic Reviews, Supplement 2010; Suppl: 14.

Greenhalgh 2005

Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *BMJ* 2005;**331**(7524):1064-5.

Higgins 2009

Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions. Version 5.0.2

[updated September 2009]. The Cochrane Collaboration, 2009. Available from www.cochrane-handbook.org. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.

Hopewell 2007

Hopewell S, McDonald S, Clarke M, Egger M. Grey literature in meta-analyses of randomized trials of health care interventions. *Cochrane Database of Systematic Reviews* 2007, Issue 2. [DOI: [10.1002/14651858.MR000010.pub3](https://doi.org/10.1002/14651858.MR000010.pub3)]

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](https://doi.org/10.1002/14651858.MR000026.pub2)]

Kirkham 2010

Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smith R, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ* 2010;**340**:c365.

Lefebvre 2008

Lefebvre C, Manheimer E, Glanville J on behalf of the Cochrane Information Retrieval Methods Group. Chapter 6: Searching for studies. In: Higgins JPT, Green S (eds). *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.0.0 [updated

February 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.

Scherer 2007

Scherer RW, Langenberg P, von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database of Systematic Reviews* 2007, Issue 2. [DOI: [10.1002/14651858.MR000005.pub3](https://doi.org/10.1002/14651858.MR000005.pub3)]

Smith 1998

Smith GD, Egger M. Meta-analysis. Unresolved issues and future developments. *BMJ* 1998;**316**(7126):221-5.

Song 2000

Song F, Eastwood AJ, Gilbody S, Duley L, Sutton AJ. Publication and related biases. *Health Technology Assessment* 2000;**4**(10):1-115.

Sterne 2008

Sterne JAC, Egger M, Moher D (editors). Chapter 10: Addressing reporting biases. In: Higgins JPT, Green S (editors). *Cochrane Handbook of Systematic Reviews of Interventions*. Version 5.0.0 [updated February 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies [ordered by study ID]

Brown 2003

Methods	This was a non-randomised comparative study. Within the context of 4 systematic reviews on the prevention of NSAID-induced gastro-intestinal toxicity trial authors were contacted by e-mail (preferential-ly) or letter, providing a semi-personalised information retrieval sheet.
Data	Context of 4 systematic reviews on the prevention of NSAID-induced gastro-intestinal toxicity; 112 authors (of 139 studies) were contacted
Comparisons	E-mail (n = 39) versus letter (n = 73) providing a semi-personalised information retrieval sheet. However, the results were not presented separately for each approach.
Outcomes	Additional information retrieved through contact with trial authors Costs incurred
Notes	This study was published as an abstract

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	Non-randomised comparison therefore not applicable
Allocation sequence generation?	Unclear	Non-randomised comparison therefore not applicable

Methods for obtaining unpublished data (Review)

Copyright © 2011 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Brown 2003 (Continued)

How allocation occurred?	Unclear	Not reported
Attempts to balance groups?	Unclear	Not reported
Use of blinding?	Unclear	Not reported

Gibson 2006

Methods	This was a non-randomised comparative study. The mode of contact and response levels of authors who had been asked to provide missing or incomplete data for a systematic review on diet and exercise interventions for weight loss was examined.
Data	A systematic review on diet and exercise interventions for weight loss
Comparisons	E-mail versus letter versus both (total n = 146; sample size per study arm not reported)
Outcomes	Proportion of responders over time among the different modes of contact Response rates from United States compared to other countries
Notes	This study was published as a full-text paper

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	Non-randomised comparison therefore not applicable
Allocation sequence generation?	Unclear	Non-randomised comparison therefore not applicable
How allocation occurred?	Unclear	Not reported
Attempts to balance groups?	Unclear	Not reported
Use of blinding?	Unclear	Not reported

Guevara 2005

Methods	<p>This was a non-randomised comparative study. As part of a Cochrane Review comparing the effects of inhaled corticosteroids to cromolyn, authors of all included trials were contacted to clarify methods and/or to obtain missing outcome data. Authors listed as corresponding authors were contacted by</p> <ul style="list-style-type: none"> • e-mail if available • by post if e-mail was unavailable or if there was no response to e-mail, and • by fax if e-mail and post were both unavailable or there was no response to either <p>Remaining authors were contacted if there was no response by the corresponding author</p>
Data	Cochrane Review comparing the effects of inhaled corticosteroids to cromolyn; study authors of all 25 included trials were contacted to clarify methods and/or to obtain missing outcome data

Methods for obtaining unpublished data (Review)

Guevara 2005 (Continued)

Comparisons	E-mail versus letter versus fax (total n = 25; sample size per study arm not reported)
Outcomes	Response rate Time to response
Notes	This study was published as an abstract

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	Non-randomised comparison therefore not applicable
Allocation sequence generation?	Unclear	Non-randomised comparison therefore not applicable
How allocation occurred?	Unclear	Unclear
Attempts to balance groups?	Unclear	E-mail was the preferred method of contact
Use of blinding?	No	No blinding

Higgins 1999

Methods	This was a randomised comparison. Contact persons or authors (primary investigators) of published studies were eligible for the study if (i) the study had been identified as probably or definitely fulfilling the criteria for inclusion in a Cochrane Review, (ii) any information needed to complete the systematic review was missing from the published report, and (iii) a postal, or e-mail, address was available for them. The reviewers should have completed assessment of studies for inclusion in the review and any data extraction.
Data	Randomised trial of Cochrane Review authors where the reviewer was uncertain how first contact should be made with the investigator of a primary study which was included in the Cochrane Review in order to obtain missing information
Comparisons	Single request for missing information (by e-mail or surface mail) (n = 116) versus multistage approach involving pre-notification, request for missing information and active follow-up (n = 117)
Outcomes	<p>Primary outcome: Amount of missing information retrieved from the investigator within 12 weeks of sending the original letter. A 4-point ordinal scale was used:</p> <ul style="list-style-type: none"> • No response: no requested information retrieved • Inadequate response: some, but not all, requested information retrieved, and insufficient information to include the trial in the review with no important gaps • Satisfactory response: some, but not all, requested information retrieved, and sufficient information to include the trial in the review with no important gaps • Perfect response: all requested information retrieved <p>Secondary outcomes:</p> <p>The time taken to receive some or all of the requested information</p> <p>Any response or acknowledgement from the investigator or someone else involved with the study or its data</p>

Methods for obtaining unpublished data (Review)

Higgins 1999 (Continued)

Cost, in terms of postage and telephone call time

Notes

 This study was called 'First Contact' and was published as an abstract. The study's website is available at <http://www.mrc-bsu.cam.ac.uk/firstcontact/index.html>

Additional information was obtained from study authors

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Yes	Central randomisation with minimisation to attempt to balance confounders
Allocation sequence generation?	Unclear	Central randomisation
How allocation occurred?	Unclear	Randomised comparison therefore not applicable
Attempts to balance groups?	Unclear	Randomised comparison therefore not applicable
Use of blinding?	Unclear	Not reported

Milton 2001

Methods

This was a randomised comparison. Authors of eligible RCTs of interventions for essential hypertension published since 1996 and forming part of a methodological systematic review were randomised to receive a mailed questionnaire with a cover letter signed by Richard Smith (RS) or Julie Milton (JM), on stationery appropriate to each. After 3 weeks non-responders were sent a questionnaire by recorded mail, with the same signatory. After a further 5 weeks, JM attempted to telephone non-responders, telling authors randomised to RS as signatory that she was calling on his behalf.

Data

Authors of 144 eligible RCTs of interventions for essential hypertension published since 1996 and forming part of a methodological systematic review

Comparisons

Using a well know signatory (n = 96) versus an unknown researcher (n = 48) on the cover letter of a mailed questionnaire

Outcomes

Response of clinical trial investigators to requests for information

Notes

This study was published as an abstract

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	Allocation was performed by an independent statistician
Allocation sequence generation?	Unclear	Not reported
How allocation occurred?	Unclear	Randomised comparison therefore not applicable
Attempts to balance groups?	Unclear	Randomised comparison therefore not applicable

Methods for obtaining unpublished data (Review)

Copyright © 2011 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Milton 2001 (Continued)

Use of blinding?	Unclear	Not reported
------------------	---------	--------------

Shukla 2003

Methods	This was a non-randomised comparative study. Over 4 years, for each of 5 systematic reviews of drugs at CCOHTA, 2 different approaches were used to seek unpublished information from the drug industry. With the first approach, a general request letter was sent. With the second approach, unpublished studies were identified ahead of time via handsearching of conference abstracts, review articles and bibliographies of included studies plus electronic searches of BIOSIS Previews. A Google search was also run. Unpublished studies were identified and industry was asked to provide further specific detail.
Data	Five systematic reviews of drugs at CCOHTA. Number of trials not reported.
Comparisons	Identifying unpublished studies ahead of time and then asking industry to provide further specific detail versus general request by letter for unpublished studies
Outcomes	Unpublished information obtained from the drug industry
Notes	This study was published as an abstract

Risk of bias

Item	Authors' judgement	Description
Allocation concealment?	Unclear	Non-randomised comparison therefore not applicable
Allocation sequence generation?	Unclear	Non-randomised comparison therefore not applicable
How allocation occurred?	Unclear	Not reported
Attempts to balance groups?	Unclear	Not reported
Use of blinding?	Unclear	Not reported

CCOHTA: Canadian Co-ordinating Office for Health Technology Assessment
 NSAID: non-steroidal anti-inflammatory drug
 RCT: randomised controlled trial

Characteristics of excluded studies [ordered by study ID]

Study	Reason for exclusion
Bohlius 2003	Study that looked at contacting authors to obtain missing data. It was excluded because there was no comparison of different methods of obtaining missing data.
Eysenbach 2001	Study that looked at use of the internet to identify unpublished studies. It was excluded because there was no comparison of different methods of obtaining unpublished studies.
Hadhazy 1999	Study that looked at contacting authors to obtain missing data. It was excluded because there was no comparison of different methods of obtaining missing data.

Methods for obtaining unpublished data (Review)

Copyright © 2011 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Study	Reason for exclusion
Hetherington 1987	Study that looked at surveying content experts to identify unpublished studies. It was excluded because there was no comparison of different methods of obtaining unpublished studies.
Kelly 2002	This study contacted authors of studies for individual patient data using letters. It was excluded because there was no comparison of different methods of obtaining missing data.
Kelly 2004	Study that looked at contacting authors to obtain missing data. It was excluded because there was no comparison of different methods of obtaining missing data.
McGrath 1998	Study that looked at contacting authors to obtain missing data. It was excluded because there was no comparison of different methods of obtaining missing data.
Reveziz 2004	Study that looked at surveying content experts to identify unpublished studies. It was excluded because there was no comparison of different methods of obtaining unpublished studies.
Wille-Jorgensen 2001	Primary authors and/or sponsoring pharmaceutical companies of studies in general surgery which might contain colorectal patients were contacted per mail, e-mail and/or personal contacts. The responses from the 3 methods were not compared.

ADDITIONAL TABLES

Table 1. Summary of interventions and outcomes

Study ID	Comparison	Outcome measures
Methods for obtaining missing data (data available to researchers but not reported in the published study)		
Brown 2003	E-mail versus letter providing a semi-personalised information retrieval sheet	Additional information retrieved through contact with trial authors Costs incurred
Gibson 2006	E-mail versus letter versus both	Proportion of responders over time Response rates United States compared to other countries
Guevara 2005	E-mail versus letter versus fax	Response rate Time to response
Higgins 1999	Single request for missing information (by e-mail or surface mail) versus multistage approach involving pre-notification, request for missing information and active follow-up	Whether contact is established Whether appropriate information is obtained
Milton 2001	Using a well-known signatory versus an unknown researcher on the cover letter of a mailed questionnaire	Response of clinical trial investigators to requests for information
Methods for obtaining unpublished data (data for studies that have never been published)		
Shukla 2003	Identifying unpublished studies ahead of time and then asking industry to provide further specific detail versus general request by letter for unpublished studies	Unpublished information obtained from the drug industry

Methods for obtaining unpublished data (Review)

APPENDICES

Appendix 1. Cochrane Methodology Register search strategy

#1 ("study identification" next general) or ("study identification" next "publication bias") or ("study identification" next "prospective registration") or ("study identification" next internet) or ("data collection") or ("missing data") or ("information retrieval" next general) or ("information retrieval" next "retrieval techniques") or ("information retrieval" next "comparisons of methods"):kw in Methods Studies

#2 (request* or obtain* or identify* or locat* or find* or detect* or search or "ask for") NEAR/3 (grey or unpublished or "un published" or "not published"):ti or (request* or obtain* or identify* or locat* or find* or detect* or search or "ask for") NEAR/3 (grey or unpublished or "un published" or "not published"):ab

#3 (request* or obtain* or identify* or locat* or find* or detect* or search or "ask for") NEAR/3 (missing or missed or insufficient or incomplete or lack* or addition*):ti or (request* or obtain* or identify* or locat* or find* or detect* or search or "ask for") NEAR/3 (missing or missed or insufficient or incomplete or lack* or addition*):ab

#4 (missing or incomplete or unpublished or "un published" or "not published") NEAR/3 (data or information or study or studies or evidence or trial or trials):ti or (missing or incomplete or unpublished or "un published" or "not published") NEAR/3 (data or information or study or studies or evidence or trial or trials):ab

#5 (bad or ambiguous or insufficient or incomplete) NEAR/6 report*:ti or (bad or ambiguous or insufficient or incomplete) NEAR/3 report*:ab

#6 (#1 OR #2 OR #3 OR #4 OR #5)

Appendix 2. MEDLINE search strategy

1. ((request\$ or obtain\$ or identify\$ or locat\$ or find\$ or detect\$ or search or ask for) adj3 (grey or unpublished or "un published" or "not published")) adj3 (data or information or evidence or study or studies or trial? or paper? or article? or report? or literature or work)).tw.

2. ((request\$ or obtain\$ or identify\$ or locat\$ or find\$ or detect\$ or search or ask for) adj3 (missing or insufficient or incomplete or lack\$ or addition\$) adj3 (data or information or evidence)).tw.

3. ((bad or ambiguous or insufficient or incomplete) adj6 reporting).tw.

4. 1 or 2 or 3

5. (2000\$ or 2001\$ or 2002\$ or 2003\$ or 2004\$ or 2005\$ or 2006\$ or 2007\$ or 2008\$ or 2009\$).ep.

6. 4 and 5

Appendix 3. EMBASE search strategy

1. ((request\$ or obtain\$ or identify\$ or locat\$ or find\$ or detect\$ or search or ask for) adj3 (grey or unpublished or "un published" or "not published")) adj3 (data or information or evidence or study or studies or trial? or paper? or article? or report? or literature or work)).tw.

2. ((request\$ or obtain\$ or identify\$ or locat\$ or find\$ or detect\$ or search or ask for) adj3 (missing or insufficient or incomplete or lack\$ or addition\$) adj3 (data or information or evidence)).tw.

3. ((bad or ambiguous or insufficient or incomplete) adj6 reporting).tw.

4. 1 or 2 or 3

5. (2004\$ or 2005\$ or 2006\$ or 2007\$ or 2008\$ or 2009\$).em.

6. 4 and 5

CONTRIBUTIONS OF AUTHORS

Taryn Young (TY) developed and Sally Hopewell (SH) provided comments on the protocol. Both authors reviewed the search results, selected potential studies for inclusion, worked independently to do a formal eligibility assessment and then extracted data from included studies. TY drafted the review with input from SH.

DECLARATIONS OF INTEREST

None known.

SOURCES OF SUPPORT

Internal sources

- South African Cochrane Centre, South Africa.
- UK Cochrane Centre, NHS Research & Development Programme, UK.

External sources

- No sources of support supplied

INDEX TERMS

Medical Subject Headings (MeSH)

*Access to Information; *Electronic Mail; Documentation [*methods]; Publishing; Randomized Controlled Trials as Topic