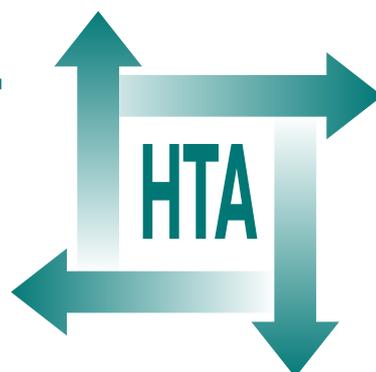# Dissemination and publication of research findings: an updated review of related biases
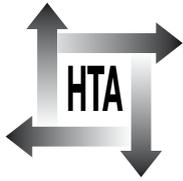
F Song, S Parekh, L Hooper,
YK Loke, J Ryder, AJ Sutton, C Hing,
CS Kwok, C Pang, and I Harvey

How to obtain copies of this and other HTA programme reports

An electronic version of this title, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (www.hta.ac.uk). A fully searchable DVD is also available (see below).

Printed copies of HTA journal series issues cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our despatch agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per issue and for the rest of the world £3 per issue.

How to order:

– fax (with **credit card details**)
– post (with **credit card details** or **cheque**)
– phone during office hours (**credit card** only).

Additionally the HTA website allows you to either print out your order or download a blank order form.

Contact details are as follows:

Synergie UK (HTA Department)
Digital House, The Loddon Centre
Wade Road
Basingstoke
Hants RG24 8QW

Email: orders@hta.ac.uk

Tel: 0845 812 4000 – ask for 'HTA Payment Services'
(out-of-hours answer-phone service)

Fax: 0845 812 4001 – put 'HTA Order' on the fax header

Payment methods

Paying by cheque
If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *University of Southampton* and drawn on a bank with a UK address.

Paying by credit card
You can order using your credit card by phone, fax or post.

Subscriptions

NHS libraries can subscribe free of charge. Public libraries can subscribe at a reduced cost of £100 for each volume (normally comprising 40–50 titles). The commercial subscription rate is £400 per volume (addresses within the UK) and £600 per volume (addresses outside the UK). Please see our website for details. Subscriptions can be purchased only for the current or forthcoming volume.

How do I get a copy of *HTA on DVD*?

Please use the form on the HTA website (www.hta.ac.uk/htacd/index.shtml). *HTA on DVD* is currently free of charge worldwide.

The website also provides information about the HTA programme and lists the membership of the various committees.

Dissemination and publication of research findings: an updated review of related biases

F Song,[1,2]* S Parekh,[1,2] L Hooper,[1]
YK Loke,[1] J Ryder,[1] AJ Sutton,[3] C Hing,[4]
CS Kwok,[1] C Pang,[1] and I Harvey[1]

[1]School of Medicine, Health Policy and Practice, University of East Anglia, Norwich, UK
[2]School of Allied Health Professions, University of East Anglia, Norwich, UK
[3]Department of Health Sciences, University of Leicester, UK
[4]Watford General Hospital, Hertfordshire, UK

*Corresponding author

Declared competing interests of authors: none

This report should be referenced as follows:

NIHR Health Technology Assessment programme

The Health Technology Assessment (HTA) programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The research findings from the HTA programme directly influence decision-making bodies such as the National Institute for Health and Clinical Excellence (NICE) and the National Screening Committee (NSC). HTA findings also help to improve the quality of clinical practice in the NHS indirectly in that they form a key component of the 'National Knowledge Service'.

The HTA programme is needs led in that it fills gaps in the evidence needed by the NHS. There are three routes to the start of projects.

First is the commissioned route. Suggestions for research are actively sought from people working in the NHS, from the public and consumer groups and from professional bodies such as royal colleges and NHS trusts. These suggestions are carefully prioritised by panels of independent experts (including NHS service users). The HTA programme then commissions the research by competitive tender.

Second, the HTA programme provides grants for clinical trials for researchers who identify research questions. These are assessed for importance to patients and the NHS, and scientific rigour.

Third, through its Technology Assessment Report (TAR) call-off contract, the HTA programme commissions bespoke reports, principally for NICE, but also for other policy-makers. TARs bring together evidence on the value of specific technologies.

Some HTA research projects, including TARs, may take only months, others need several years. They can cost from as little as £40,000 to over £1 million, and may involve synthesising existing evidence, undertaking a trial, or other research collecting new data to answer a research problem.

The final reports from HTA projects are peer reviewed by a number of independent expert referees before publication in the widely read journal series *Health Technology Assessment*.

Criteria for inclusion in the HTA journal series

Reports are published in the HTA journal series if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

Abstract

Dissemination and publication of research findings: an updated review of related biases

F Song,[1,2]* S Parekh,[1,2] L Hooper,[1] YK Loke,[1] J Ryder,[1] AJ Sutton,[3] C Hing,[4] CS Kwok,[1] C Pang,[1] and I Harvey[1]

[1]School of Medicine, Health Policy and Practice, University of East Anglia, Norwich, UK
[2]School of Allied Health Professions, University of East Anglia, Norwich, UK
[3]Department of Health Sciences, University of Leicester, UK
[4]Watford General Hospital, Hertfordshire, UK

*Corresponding author

Objectives: To identify and appraise empirical studies on publication and related biases published since 1998; to assess methods to deal with publication and related biases; and to examine, in a random sample of published systematic reviews, measures taken to prevent, reduce and detect dissemination bias.

Data sources: The main literature search, in August 2008, covered the Cochrane Methodology Register Database, MEDLINE, EMBASE, AMED and CINAHL. In May 2009, PubMed, PsycINFO and OpenSIGLE were also searched. Reference lists of retrieved studies were also examined.

Review methods: In Part I, studies were classified as *evidence* or *method* studies and data were extracted according to types of dissemination bias or methods for dealing with it. Evidence from empirical studies was summarised narratively. In Part II, 300 systematic reviews were randomly selected from MEDLINE and the methods used to deal with publication and related biases were assessed.

Results: Studies with significant or positive results were more likely to be published than those with non-significant or negative results, thereby confirming findings from a previous *HTA* report. There was convincing evidence that outcome reporting bias exists and has an impact on the pooled summary in systematic reviews. Studies with significant results tended to be published earlier than studies with non-significant results, and empirical evidence suggests that published studies tended to report a greater treatment effect than those from the grey literature. Exclusion of non-English-language studies appeared to result in a high risk of bias in some areas of research such as complementary and alternative medicine. In a few cases, publication and related biases had a potentially detrimental impact on patients or resource use. Publication bias can be prevented before a literature review (e.g. by prospective registration of trials), or detected during a literature review (e.g. by locating unpublished studies, funnel plot and related tests, sensitivity analysis modelling), or its impact can be minimised after a literature review (e.g. by confirmatory large-scale trials, updating the systematic review). The interpretation of funnel plot and related statistical tests, often used to assess publication bias, was often too simplistic and likely misleading. More sophisticated modelling methods have not been widely used. Compared with systematic reviews published in 1996, recent reviews of health-care interventions were more likely to locate and include non-English-language studies and grey literature or unpublished studies, and to test for publication bias.

Conclusions: Dissemination of research findings is likely to be a biased process, although the actual impact of such bias depends on specific circumstances. The prospective registration of clinical trials and the endorsement of reporting guidelines may reduce research dissemination bias in clinical research. In systematic reviews, measures can be taken to minimise the impact of dissemination bias by systematically searching for and including relevant studies that are difficult to access. Statistical methods can be useful for sensitivity analyses. Further research is needed to develop methods for qualitatively assessing the risk of publication bias in systematic reviews, and to evaluate the effect of prospective registration of studies, open access policy and improved publication guidelines.

Contents

List of abbreviations

AIDS	acquired immunodeficiency syndrome		ITT	intention-to-treat
CDUS	Clinical Data Update System		JIF	journal impact factor
CI	confidence interval		LILACS	Latin American and Caribbean Health Sciences Literature
CINAHL	Cumulative Index to Nursing and Allied Health Literature		NIH	National Institutes of Health
CMRD	Cochrane Methodology Register Database		NSAID	non-steroidal anti-inflammatory drug
CONSORT	Consolidated Standards of Reporting Trials		OR	odds ratio
CSR	Cochrane Systematic Review		QUOROM	Quality of Reporting of Meta-analyses
CTSP	Clinical Trials Search Portal (WHO clinical trial register)		R&D	research and development
DARE	Database of Abstracts of Reviews of Effectiveness		RCT	randomised controlled trial
			REC	Research Ethics Committee
EQUATOR	Enhancing the Quality and Transparency of Health Research		ROR	ratio of odds ratios
			RR	relative risk or rate ratio
FDA	United States Food and Drug Administration		RTOG	Radiation Therapy Oncology Group
HR	hazard ratio		SIGLE	System for Information on Grey Literature in Europe
HRHR	hazard ratio of hazard ratios		SSRI	selective serotonin reuptake inhibitor
HRT	hormone replacement therapy		STARD	Statement for Reporting Studies of Diagnostic Accuracy
ICH	International Conference on Harmonisation		STROBE	Standards for the Reporting of Observational Studies in Epidemiology
ICMJE	International Committee of Medical Journal Editors		TSA	trial sequential analysis
IPD	individual patient data		WHO	World Health Organization
IQR	interquartile range			
IRB	Institutional Review Board			
ISRCTN	International Standard Randomised Controlled Trial Number			

All abbreviations that have been used in this report are listed here unless the abbreviation is well known (e.g. NHS), or it has been used only once, or it is a non-standard abbreviation used only in figures/tables/appendices, in which case the abbreviation is defined in the figure legend or in the notes at the end of the table.

Executive summary

Background

The validity of research synthesis is threatened if studies with significant or striking findings are more likely to be published than those with non-significant results. A previous *Health Technology Assessment* (*HTA*) monograph published in 2000 by the present authors reviewed studies on publication and related biases. Since then, many new studies on publication and related biases have been published. This report aims to update the 2000 *HTA* monograph on publication bias by synthesising findings from previous studies and newly indentified ones.

Objectives

- To identify and appraise empirical studies on publication and related biases published since 1998.
- To assess the usefulness and limitations of available methods to deal with publication and related biases.
- To examine in a random sample of published systematic reviews, measures taken by the authors to prevent, reduce and detect different types of dissemination bias.

Methods

Part I: Review of evidence and method studies
Study selection

The report included *evidence* studies that provided empirical evidence on the existence, consequences, causes and/or risk factors of dissemination bias; and *method* studies that developed or evaluated methods for preventing, reducing or detecting dissemination bias.

Data sources
The following electronic databases were searched: Cochrane Methodology Register Database (CMRD), MEDLINE, EMBASE, AMED and CINAHL. The main literature search was conducted in August 2008 and a final search of PubMed, PsycINFO and OpenSIGLE was conducted in May 2009 to identify more recently published studies. We also examined reference lists of retrieved studies.

Data extraction and synthesis
The identified studies were classified by one reviewer as *evidence* or *method* studies and checked by a second reviewer. One reviewer extracted data directly into tables (specifically designed according to types of bias or methods), which were checked by a second reviewer. Evidence from empirical studies was summarised narratively. Where appropriate, the results have been quantitatively pooled.

Part II: Survey of published systematic reviews

We searched MEDLINE for systematic reviews published in 2006, and randomly selected 100 reviews of effects of health-care interventions, 50 reviews of diagnostic accuracy, 100 reviews of association between risk factors and health outcomes, and 50 reviews of gene-disease associations. We assessed the methods used to deal with publication and related biases in these systematic reviews.

Results

Empirical evidence on dissemination bias

Updated analyses of data from cohort studies confirmed findings from the previous *HTA* report that studies with significant or positive results are more likely to be published than those with non-significant or negative results. Publication bias occurs mainly before the presentation of findings at conferences and before the submission of manuscripts to journals. Recent high-quality studies have provided convincing evidence that outcome reporting bias exists and has an important impact on the pooled summary in systematic reviews. Studies with significant results tend on average to be published earlier than studies with non-significant results, although the new evidence is less clear than that from the previous review. New empirical evidence suggests that published studies tend to report a greater treatment effect than those

from grey literature. However, for individual cases, the direction of bias is unpredictable, and grey literature studies may be relatively small and of poor quality. The impact of non-English language studies was highly heterogeneous. Exclusion of non-English language studies appears to result in a particularly high risk of bias in some areas of research such as complementary and alternative medicine. The updated review also identified limited evidence on citation bias, duplicate publication bias, place of publication bias, database or index bias, country bias and media attention bias.

Limitations of the available evidence

Empirical studies on publication and related biases have focused mainly on certain areas of research such as clinical trials of health-care interventions. When studies are classified as positive or important, bias may be introduced due to inevitable subjectivity. Much of the empirical evidence comes from case reports, which may be selectively reported because of their striking findings. Cohort studies often included studies that were diverse in terms of design and research questions. It is usually impossible to exclude the impact of confounding factors on the observed association between study results and publication status.

Consequences of research dissemination bias

The most important consequences of publication bias include avoidable suffering of patients and waste of limited resources. This updated review identified only a few new cases that indicate the detrimental impact of publication and related biases. Consequences of publication and related biases are different for different types of research studies. Dissemination bias can jeopardise the integrity of scientific research.

Sources of publication bias

The dissemination profile of a research finding is determined by the interests of research sponsors, investigators, peer-reviewers and editors. The updated review identified further evidence indicating that publication bias is often due to investigators' failure to write up and submit, although it should be recognised that the investigators' decision to write up an article and then submit it may be affected by pressure from research sponsors, preferences of journal editors, and the requirements of the research award system. Newly identified and previous included evidence

suggested that the interests of research sponsors, particularly industry's commercial interests, can restrict the dissemination of the research findings. Studies that can be conducted without the use of large amounts of resource investment, and those that are of great variations in results are more subject to publication bias.

Methods to prevent, reduce or detect publication and related bias

The available methods can be classified according to the stage of a literature review: to prevent publication bias before a literature review (e.g. prospective registration of trials), to detect publication bias during a literature review (e.g. locating unpublished studies, funnel plot and related tests, sensitivity analysis modelling), or to minimise the impact of publication bias after a literature review (e.g. confirmatory large-scale trials, updating the systematic review).

The first step for the prevention of publication bias is a wide public awareness of detrimental consequences of publication bias, and the need for the results of all studies to be made accessible. One important solution to publication bias is the prospective registration of all studies at inception. The compulsory policy of trial registration adopted by the International Committee of Medical Journals in 2004 may be the most influential initiative so far to promote prospective registration of clinical trials. The World Health Organization (WHO) initiated a project in 2005 to set international standards for clinical trial registration. Further action through government regulations (e.g. the FDA Modernisation Act in the USA) may still be required. In spite of the greater risk of publication bias, the prospective registration and disclosure of data from unpublished basic research, observational studies and early stage exploratory trials has faced considerable difficulties.

The development of prospective trial registration itself is not sufficient for the prevention of publication bias. It is important to make sure that results of registered trials are publically accessible. The usefulness of trial registrations relies on systematic reviewers searching them, using the data they provide and spending time contacting trialists where studies have not yet been published.

The recent development of clinical trial registration and electronic publication of results from clinical

trials will facilitate the identification and location of ongoing or unpublished clinical trials. Funnel plot and related statistical tests have been widely used to assess publication bias. Unfortunately, the interpretation of results of funnel plot tests was often too simplistic and likely misleading. Many sophisticated modelling methods have not been widely used in systematic reviews, possibly because of their complexity and lack of user friendly software.

Survey of published systematic reviews

Compared with systematic reviews published in 1996, recent systematic reviews of health-care interventions are making greater efforts to locate and include non-English language studies (47% versus 30%), and grey literature or unpublished studies (53% versus 35%). There was also an increased use of available methods to test for publication bias in recent reviews (22% versus 17%). Grey literature, unpublished studies or non-English language studies were more likely to be searched for in reviews of treatment efficacy or diagnostic accuracy than in reviews of epidemiological studies. However, the risk of publication bias was less likely to be tested in reviews of treatment and diagnosis as compared with reviews of epidemiological studies.

Conclusions

Dissemination of research findings is likely to be a biased process, although the actual impact of such bias is still uncertain, depending on specific circumstances. Therefore, the potential problem of research dissemination bias should be taken into consideration by all who are involved in evidence-based decision making. The recent initiatives in the prospective registration of clinical trials and the endorsement of reporting guidelines may prevent or reduce publication and reporting bias in future systematic reviews of clinical trials, although prospective registration of basic research, early stage clinical studies and observational studies is still underdeveloped. However, trial registers will only be helpful in reducing publication bias if the results of registered trials are accessible. In systematic reviews, measures can be taken to minimise the impact of research dissemination bias by systematically searching for published and unpublished studies. All statistical methods, simple or complex, are by nature indirect and exploratory, and are often based on certain assumptions that can be difficult to justify. The available statistical methods can be useful for the purpose of sensitivity analyses.

Recommendations for future research

- Further empirical research is needed to evaluate the effect of prospective registration of studies, open access policy and improved publication guidelines in the prevention of research dissemination bias.
- The role of developments in computer science and information technology for the prevention of research dissemination bias needs to be investigated by further research.
- The impact of publication bias on health decision-making and the outcomes of patient management need to be investigated by further research.
- Methods that can be used to assess qualitatively the risk of publication bias in systematic reviews need to be developed by further research.
- Further research should focus on the practical application of the available statistical methods.

Chapter 1

Introduction

Synthesis of published research is becoming increasingly important in providing relevant and valid research evidence to clinical and health policy decision-making. However, the validity of research synthesis based on published literature will be threatened if published studies comprise a biased selection of all studies that have been conducted.[1]

A previous *Health Technology Assessment* (*HTA*) monograph published in 2000 comprehensively reviewed studies that provided empirical evidence of publication and related biases, and studies that developed or tested methods for preventing, reducing or detecting publication and related biases.[2] The review found evidence indicating that studies with significant or favourable results were more likely to be published, or were likely to be published earlier than those with non-significant or unimportant results. There was limited and indirect evidence indicating the possibility of full publication bias, outcome reporting bias, duplicate publication bias, and language bias. The review identified little empirical evidence relating to the impact of publication and related biases on health policy, clinical decision-making and the outcome of patient management. Considering that the spectrum of the accessibility of research results (dissemination profile) ranges from completely inaccessible to easily accessible, it was suggested that a single term, 'dissemination bias', could be used to denote all types of publication and related biases.[2]

In the previous *HTA* report published in 2000, the available methods for dealing with dissemination biases were classified according to measures that could be taken before, during or after a literature review: to prevent publication bias before a literature review (e.g. prospective registration of trials), to reduce or detect publication and related biases during a literature review (e.g. locating grey literature or unpublished studies, and funnel plot related methods), and to minimise the impact of publication bias after a literature review (e.g. confirmatory large-scale trials, updating systematic reviews).[2] It was concluded that the ideal solution to publication bias is the prospective, universal registration of all studies at their inception. It was also concluded, although debatable, that available statistical methods for detecting and adjusting publication bias should be mainly used for the purpose of sensitivity analysis.[2]

Since the publication of the 2000 *HTA* report on publication bias, many new empirical studies on publication and related biases have been completed and published. For example, Egger *et al.* (2003) provided further empirical evidence on publication bias, language bias, grey literature bias and MEDLINE index bias,[3] and Moher *et al.* (2003) evaluated language bias in meta-analyses of randomised controlled trials.[4] Recently, more convincing evidence on outcome reporting bias has been published.[5–7] The new empirical evidence may contradict or strengthen the empirical evidence included in the previous *HTA* report. There are also some new published studies that investigated methods for dealing with publication bias (e.g. references 8 and 9). More importantly, perhaps, new initiatives have been introduced to enhance the prospective registration of clinical trials.[10]

This report aims to update the 2000 *HTA* report on publication bias, by incorporating findings from newly identified studies. We first discuss the concepts and definitions about publication and related biases in this chapter. After a description of review objectives and methods in Chapter 2, evidence from empirical studies on the existence and consequences of publication bias is summarised in Chapters 3 to 5. We discuss sources of publication bias in Chapter 6, while methods for dealing with publication bias are examined in Chapters 7 and 8. The results of a survey of systematic reviews published in 2006 are presented in Chapter 9. Finally, the major findings of this updated review are discussed in Chapter 10.

Definition of publication and related biases

The observation that many studies are never published was termed 'the file-drawer problem' by Rosenthal in 1979.[11] The importance of this problem depends on whether or not the published studies are representative of all studies that have

been conducted. If the published studies are the same as, or a random sample of, all studies that have been conducted, there will be no bias and the average estimate based on the published studies will be similar to that based on all studies. If the published studies comprise a biased sample of all studies that have been conducted, the results of a literature review will be misleading.[12] The efficacy of a treatment will be exaggerated if studies with positive results are more likely to be published than those with negative results.

Publication bias is specifically defined as 'the tendency on the parts of investigators, reviewers, and editors to submit or accept manuscripts for publication based on the direction or strength of the study findings'.[13] In the definition of publication bias, there are two basic concepts: study findings and publication status. Study findings are commonly classified as being statistically significant or non-significant. In addition, study results may be classified as being positive or negative, supportive or unsupportive, favoured or disliked, striking or unimportant. It should be noted that the classification of study findings is often dependent on subjective judgement and may be unreliable. For example, people may have a different understanding about what are positive or negative findings.

The formats of publication include full publication in journals, presentation at scientific conferences, reports, book chapters, discussion papers, dissertations or theses. In fact, 'publication is not a dichotomous event: rather it is a continuum'.[14] Although a study that appears in a full report in a journal is generally regarded as published, there may be different opinions about whether it should be classified as published or unpublished when results are presented in other formats.

The accessibility of research results is dependent not only on whether a study is published but also on when, where and in what format this occurs. In the 2000 *HTA* report on publication bias, we used the term 'dissemination profile' to describe the accessibility of research results, or the possibility of research findings being identified by potential users. The spectrum of the dissemination profile ranges from completely inaccessible to easily accessible, according to whether, when, where and how research is presented or stored. Dissemination bias occurs when the dissemination profile of a study is determined by its results. The term dissemination bias could be used to embrace publication bias and other related biases caused by time, type and language of publication, multiple publication, selective citation, database index, and biased media attention (see *Box 1* for definitions).

The advantages of the term 'dissemination bias' are that it avoids the need to define publication status and it is more directly related to accessibility than publication. For example, media attention can have a major impact on dissemination, but it is not normally included within the definition of publication bias. Along with the development of information technology and changes in regulations and policy, data from some 'unpublished' studies may be conveniently accessible to the public, and formal publication in journals is only one of several ways to disseminate research findings. Therefore, dissemination bias is a better expression to replace this broad use of the term publication bias. However, the term 'publication bias' and 'publication and related biases' are already established in research literature and they will also be used for discussion in this report.

BOX 1 *Definitions of dissemination, publication and related biases*

Dissemination bias

Occurs when the dissemination profile of a study's results depends on the direction or strength of its findings. The dissemination profile is defined as the accessibility of research results or the possibility of research findings being identified by potential users. The spectrum of the dissemination profile ranges from completely inaccessible to easily accessible, according to whether, when, where and how research is published.

Publication bias

Occurs when the publication of research results depends on the nature and direction of the results. Because of publication bias, the results of published studies may be systematically different from those of unpublished studies.

Outcome reporting bias

Occurs when a study in which multiple outcomes were measured reports only those that were significant.

Time lag bias

Occurs when the speed of publication depends on the direction and strength of the trial results. For example, studies with significant results may be published earlier than those with non-significant results.

Grey literature bias

Occurs when the results reported in journal articles are systematically different from those presented in reports, working papers, dissertations or conference abstracts.

Full publication bias

Occurs when the full publication of studies that have been initially presented at conferences or in other informal formats is dependent on the direction and/or strength of their findings.

Language bias

Occurs when languages of publication depend on the direction and strength of the study results.

Multiple publication bias (duplicate publication bias)

Occurs when studies with significant or supportive results are more likely to generate multiple publications than studies with non-significant or unsupportive results. Duplicate publication can be classified as 'overt' or 'covert'. Multiple publication bias is particularly difficult to detect if it is covert, when the same data are published in different places or at different times without providing sufficient information about previous or simultaneous publication.

Place of publication bias

In this review, this is defined as occurring when the place of publication is associated with the direction or strength of the study findings. For example, studies with positive results may be more likely to be published in widely circulated journals than studies with negative results. The term was originally used to describe the tendency for a journal to be more enthusiastic towards publishing articles about a given hypothesis than other journals, for reasons of editorial policy or readers' preference.

Citation bias

Occurs when the chance of a study being cited by others is associated with its result. For example, authors of published articles may tend to cite studies that support their position. Thus, retrieving literature by scanning reference lists may produce a biased sample of articles and reference bias may also render the conclusions of an article less reliable.

Database bias (indexing bias)

Occurs when there is biased indexing of published studies in literature databases. A literature database, such as MEDLINE or EMBASE, may not include and index all published studies on a topic. The literature search will be biased when it is based on a database in which the results of indexed studies are systematically different from those of non-indexed studies.

Media attention bias

Occurs when studies with striking results are more likely to be covered by the media (newspapers, radio and television news).

Chapter 2
Review objectives and methods

The current review is an update of a previous *Health Technology Assessment* (*HTA*) report.[2] This review is divided into two parts: a review of empirical and methodological studies on publication and related biases, and a survey of publication bias in a sample of published systematic reviews.

Objectives

1. To identify relevant evidence studies published since 1998. Evidence studies are defined as those that provide empirical evidence on the existence, consequences, causes and risk factors of dissemination bias.
2. To identify relevant method studies published since 1998. Method studies are those that have developed or investigated methods for preventing, reducing or detecting dissemination bias.
3. To categorise evidence and method studies identified according to a conceptual framework of dissemination profile, and to critically appraise studies that provided direct empirical evidence.
4. To synthesise findings from newly identified and previously included studies to enable us to assess whether each type of dissemination bias does exist, and if so the extent of the effect that it may have on results of systematic reviews and hence decision-making.
5. To assess the usefulness and limitations of available methods through synthesis of the methodological studies.
6. To examine measures taken in a representative sample of published systematic reviews to prevent, reduce and detect different types of dissemination bias. We included both narrative and quantitative (meta-analytic) systematic reviews that evaluated effect of health-care interventions, systematic reviews that evaluated the accuracy of diagnostic tests, systematic reviews that evaluated association between genes and disease, and systematic reviews of epidemiological studies that evaluated association of risk factors and health outcomes.
7. To bring together current evidence on the existence and scale of each type of dissemination bias, effects of methods to combat these biases, and current use of these methods to create recommendations for reviewers, policy-makers and health professionals.

Review of empirical and methodological studies

Criteria for inclusion

We included studies that provide empirical evidence on the existence, consequences, causes and/or risk factors of types of dissemination bias; and method studies that develop or evaluate methods for preventing, reducing or detecting dissemination bias in biomedical or health-related research. Evidence studies are those that provided empirical evidence on the existence, consequences, causes and risk factors of types of dissemination bias; and method studies are those whose main objectives involved one of the following: to develop or evaluate methods for preventing, reducing or detecting dissemination bias. In some cases, a study may be considered as both an evidence study and a method study.

Literature search strategy

The following health-related or biomedical bibliographic databases were searched to identify relevant studies pertaining to empirical evidence and methodological issues concerning publication and related biases: MEDLINE, the Cochrane Methodology Register Database (CMRD), EMBASE, AMED and the Cumulative Index to Nursing and Allied Health Literature (CINAHL). The strategies used to search these electronic databases are presented in Appendix 1. The period searched was from 1998 to August 2008. A further search of PubMed (from 2008 to 2009), PsycINFO (from 1998 to 2009), and OpenSIGLE (from 1998 to 2009) was carried out in May 2009 by one reviewer to identify relevantly published or grey literature studies. References (titles with or without abstracts) gathered by searching MEDLINE and CMRD were independently examined by two reviewers. References from other databases were

assessed by one reviewer because they were mostly duplicates of those from MEDLINE.

Literature searches for methodological studies are often difficult because of ill-defined boundaries and inappropriate indexing in commonly used bibliographic databases.[15] In addition, a large number of relevant issues need to be considered in this methodological review. It is hence possible that many relevant studies may be missed by formal searches of electronic databases. Therefore, an iterative approach for literature search was adopted by examining the reference lists of retrieved studies, and examining citations of identified key studies, to identify additional relevant studies. A more focused search of databases was also conducted during the review of specific issues.

Classification of identified relevant studies

According to findings from the previous *HTA* report,[2] the relevant evidence and method studies were numerous in quantity and substantially diverse in quality. To facilitate subsequent assessment

and synthesis, identified studies were classified according to a framework of study classification (*Figure 1*).

The identified studies were initially classified by one reviewer as evidence or method studies. Empirical evidence studies were further subcategorised into various types of dissemination bias according to a framework of dissemination profile: non-publication (never or delayed); incomplete publication (outcome reporting or abstract bias); limited accessibility to publication (grey literature, language or database bias); other biased dissemination (citation, duplicate or media attention bias). Some studies were included in more than one category.

The evidence studies were separated into two groups – direct and indirect evidence studies. Direct evidence referred to findings that could be used directly to indicate dissemination bias, including admissions of bias on the part of those involved in the publication process, comparison of the results of published and unpublished studies, and the prospective and retrospective follow-up
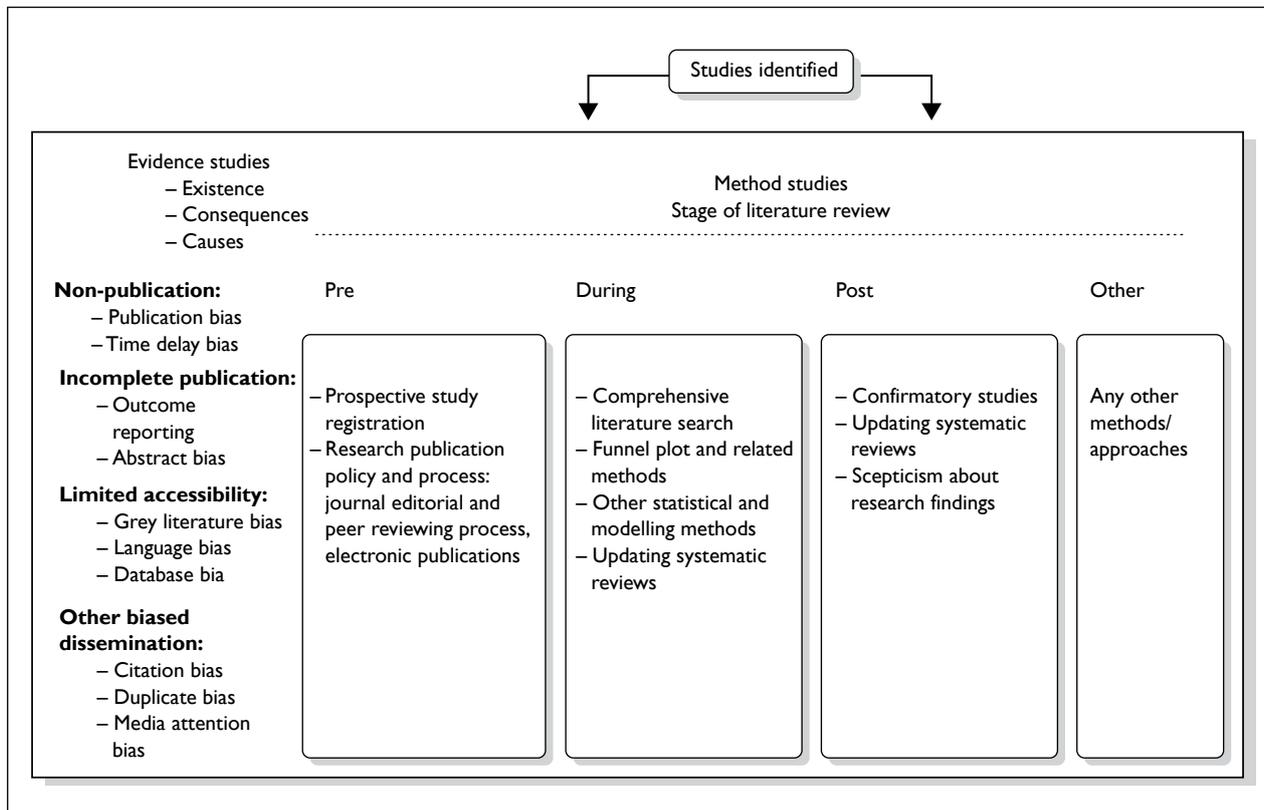

FIGURE 1 *Classification of identified relevant studies on publication and related biases.*

of dissemination profile of cohorts of studies. Indirect evidence referred to findings that could presumably have some relation with dissemination bias but where other alternative explanations could not be completely excluded. The availability of empirical evidence is very different for different types of research dissemination bias. This updated review focused on direct evidence, although indirect evidence was also considered when direct evidence was limited or absent.

The initial search of the electronic databases yielded a total of 1353 records, with much duplication, many studies being indexed in several different databases. These search results were assessed by one reviewer and 705 potentially relevant articles were identified. These studies were then independently assessed by two reviewers based on their abstracts. Finally, 300 studies were included, of which 109 were classified as evidence studies, 52 as method studies and 9 as both evidence and methods. The remaining studies were classified as background or other studies.

Data extraction and synthesis

We planned to apply a checklist of quality assessment critically to appraise studies that provided empirical evidence (Appendix 2). However, we found it was extremely difficult because of poor reliability of the checklist; different reviewers often disagreed about the overall quality of studies. The task of quality assessment was made more difficult because the designs and objectives of relevant studies in this review were highly diverse. Considering the very limited time available, we decided not to apply the checklist for quality assessment. However, we did try to identify and summarise the main limitations in studies that provided empirical evidence on publication bias in the review, although this assessment of study validity was not as systematic as specified in the protocol.

Initially, data from the included studies were independently extracted by two reviewers using separate data extraction forms for empirical and methodological studies (Appendix 2 and Appendix 3) and any disagreements were resolved by consensus. However, we found that data extracted using Appendix 2 and Appendix 3 were often insufficient, and the extraction of data from studies directly into study tables was more flexible and efficient. To save time, one reviewer extracted data directly into tables, which were checked by a second reviewer.

Findings from the newly identified studies and the previously identified studies[2] are summarised to assess each type of publication bias and the impact of these biases on the results of the systematic review and consequently decision-making. Evidence and method studies were narratively synthesised. Where judged appropriate, the results have been quantitatively pooled (e.g. the odds ratio of full publication of studies according to results). Heterogeneity across studies within each subgroup was measured using the I^2 statistic.[16] Meta-analyses were carried out using Review Manager (RevMan Version 5.0. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2008).

Assessment of a sample of published reviews

In the previous *HTA* report, 193 systematic reviews taken from the Database of Abstracts of Reviews of Effectiveness (DARE) were used to identify further evidence of dissemination bias and to illustrate the methods used in systematic reviews for dealing with publication bias. However, there were several shortcomings in our previous assessment. Firstly, systematic reviews included in the DARE database might on average have been of better quality than those from the general bibliographic databases (such as MEDLINE) so the representativeness of systematic reviews assessed in the previous *HTA* report was questionable. Secondly, 91% of systematic reviews evaluated effectiveness of health-care interventions and 9% evaluated the accuracy of diagnostic technologies, and these were not separately assessed. The problem of dissemination bias might be different between the two types of systematic reviews. Thirdly, neither reviews of epidemiological studies of association between risk factors and health outcomes nor reviews of studies of association between genes and diseases were included in the previous *HTA* report.

To overcome these shortcomings, in the current updated review we have obtained a representative sample of systematic reviews from the general bibliographic database MEDLINE and have separately assessed them as (1) systematic reviews of studies on effects of health-care interventions or treatment reviews, (2) systematic reviews of epidemiological studies on association between risk factors and health outcomes or epidemiological reviews, (3) systematic reviews of genetic studies on association between genes and disease or genetic reviews, and (4) systematic reviews of studies on accuracy of diagnostic tests or diagnostic reviews.

We have also assessed a sample of systematic reviews that tested for publication bias.

Identifying and sampling of reviews

A search of MEDLINE using 'systematic review' or 'meta-analysis' (in titles or in abstracts) identified 3503 English-language references published in 2006 and 2007. In this updated review, any published literature reviews of primary studies that reported methods for literature search were considered systematic reviews. Editorials, letters and review of reviews were excluded. These references were assessed by one reviewer to identify systematic reviews. Then the identified systematic reviews ($n = 2481$) were categorised by one reviewer into four categories – treatment effect, diagnostic accuracy, risk factors, and gene-disease association reviews – and checked by another reviewer. The final sample of reviews comprised 1448 treatment reviews, 251 diagnostic reviews, 598 epidemiological reviews and 184 genetic reviews. We then obtained computer-generated random numbers to select a random sample of 100 systematic reviews of treatment effects, 50 reviews of studies of diagnostic accuracy, 100 reviews of epidemiological studies, and 50 reviews of gene-disease associations.

For the assessment of reviews that explicitly considered or tested for publication bias, we used a restrictive search strategy limited to systematic reviews or meta-analyses that tested publication bias and were published from 2000 to 2008 in the English language (Appendix 1). The search was conducted in August 2008 and identified 204 potentially relevant reviews. These were assessed by one reviewer to identify those reviews that tested for publication bias and then computer-generated random numbers were used to select a random sample of 50 such reviews.

Data extraction and analysis

Using a data extraction form (Appendix 4, slightly revised according to types of reviews), two reviewers independently extracted data from included systematic reviews. Any disagreements between the two reviewers were resolved by discussion or by scrutiny from a third reviewer. A pre-derived scoring system was tested to assess the reviewers' judgement on efforts taken to reduce publication bias and risk of publication bias to assess the degree of agreement between the two reviewers. According to measures taken to deal with publication and related biases in a systematic review, efforts to minimise publication bias were judged to be 'sufficient', or 'partial sufficient' or 'insufficient'. Risk of publication bias was correspondingly considered to be 'low', or 'moderate' or 'high' (see Chapter 9 for more details).

Data were separately extracted from systematic reviews of effects of health-care interventions, systematic reviews of epidemiological studies, systematic reviews of genetic studies and systematic reviews of studies on accuracy of diagnostic tests. Each category of systematic reviews was analysed separately and then compared. Within each category of systematic reviews, methods used for identifying and preventing or reducing publication and related biases were examined and compared. Data from the reviews that tested for publication biases were assessed separately to find the most commonly used method to test publication bias and risk of publication bias in such reviews. The findings are synthesised and discussed in detail in Chapter 8. Systematic reviews of effects of health-care interventions and systematic reviews of diagnostic accuracy published in the previous *HTA* report were also compared with the present findings to examine whether the reporting and treatment of dissemination bias have improved over time.

Chapter 3
Evidence from cohort studies of publication bias

Evidence of publication bias can be classified as direct or indirect.[17] Indirect evidence includes observations of a disproportionately high percentage of positive findings in the published literature, and a larger effect size in small studies as compared with large studies. This evidence is indirect because factors other than publication bias may also lead to the observed disparities. The existence of publication bias was first suspected by Sterling in 1959, after observing that 97% of studies published in four major psychology journals were statistically significant.[18] In 1995, the same author concluded that the practices leading to publication bias had not changed over a period of 30 years.[19]

Direct evidence includes the admissions of bias on the part of those involved in the publication process (investigators, referees or editors), comparison of the results of published and unpublished studies, and the follow-up of cohorts of registered studies.[2] The 2000 *HTA* report on publication bias included both direct and indirect evidence. Because of a large amount of new direct evidence, this updated review focuses on direct evidence from empirical studies, but indirect evidence will also be considered when direct evidence is limited. Surveys of authors and investigators that provide evidence on publication bias are included in Chapter 5 (sources of publication bias).

This section includes any empirical studies that tracked a cohort of studies before their formal publication and reported the rate of publication by study results. Cohort studies on time to publication and selective outcome reporting will be reviewed later. Included cohort studies of publication bias were classified into four subgroups according to the starting point of follow-up of cohorts: inception cohort studies, regulatory cohort studies, abstract cohort studies and manuscript cohort studies. A study that followed up a cohort of research from their beginning (even if retrospectively) was termed an inception cohort study. A regulatory cohort study refers to a study that examined formal publication of research submitted to regulatory

authorities. An abstract cohort study refers to a study that investigated the subsequent full publication of abstracts presented at conferences. A manuscript cohort study refers to a study of manuscripts submitted to journals. Primary studies included in cohort studies of publication bias may be clinical trials, observational studies or basic research.

Publication of a study was usually defined as the full publication in journals. However, study results may be categorised differently in the included cohort studies. In this review, study results were classified as 'statistically significant ($p \leq 0.05$)' versus 'non-significant ($p > 0.05$)' or 'positive' versus 'non-positive'. Positive results included those that were considered being 'positive', 'favourable', 'significant', 'important', 'striking', 'showed effect' and 'confirmatory'. Non-positive result refers to other results labelled as being 'negative', 'non-significant', 'less or not important', 'invalidating', 'inconclusive', 'questionable', 'null' and 'neutral'.

Inception cohort studies

Five cohort studies were included in the 2000 *HTA* report.[20–24] Cohorts of research protocols approved by Research Ethics Committees (REC), Institutional Review Boards (IRB) or registered by research sponsors were followed up to investigate factors associated with subsequent publication. The study by Ioannidis included randomised controlled trials (RCTs) conducted by two groups of trialists [sponsored by the National Institutes of Health (NIH) from 1986 to 1996] and was focused mainly on time lag bias.[23] The rate of publication ranged from 60% to 98% for studies with statistically significant results and from 20% to 85% for studies with statistically non-significant results. Dickersin (1997) combined the results from four cohort studies[20–22,24] and found that the pooled adjusted odds ratio (OR) for publication bias (publication of studies with significant or important results versus those with unimportant results) was 2.54 (95% CI: 1.44 to 4.47).[25]

The updated review included seven additional inception cohorts that provided data on the rate of publication according to study results (Appendix 5).[26–32] Seven other inception cohort studies were excluded because they did not examine the association between publication and study results.[33–39]

The cohort study by Bardy (1998) included 188 of the 274 drug trials notified to the Finnish National Agency in 1987.[26] Study results were classified as being positive if the risk-benefit ratio was in favour of the drug under investigation, or if the objective of the study was supported. Results were considered inconclusive if the risk-benefit assessment was inconclusive or if the study was non-comparative, whereas studies were judged as negative if the benefit-risk ratio was not in favour of the drug or no different from placebo. The rate of publication was 47% for positive results, 33% for inconclusive results, and 11% for negative results.[26]

Cronin and Sheldon (2004) sent a questionnaire to project leaders of 101 projects sponsored by the UK NHS R&D (research and development) programme to obtain information on study findings and publication status.[27] The method suggested by Dickersin and Min[40] was used to define study results. Studies were categorised as 'showed (an) effect' or not, depending on whether results were statistically significant ($p < 0.05$) or considered to be of great importance. The rate of publication of studies with statistically significant or important results was not statistically significantly different from those with non-significant or non-important results (76% versus 64%).[27]

Two cohort studies by Decullier and colleagues (2005, 2006) followed up biomedical research protocols approved by French RECs in 1994 and 1997.[28,29] In one of the two French studies, results of 501 completed studies were classified by original investigators as being confirmatory, invalidating or inconclusive (see Appendix 5 for details).[28] The rate of publication of studies with confirmatory results was higher than those with inconclusive results (OR = 4.59; 95% CI: 2.21 to 9.54).[28] In the other French study of 47 completed studies, the importance of results was subjectively rated by investigators from 1 to 10, and important results were those graded as > 5.[29] The rate of publication was 70% for studies with important results and 60% for those with less important results (OR = 1.58; 95% CI: 0.37 to 6.71).[29]

Misakian and Bero identified a cohort of 61 passive smoking research projects that were sponsored by

76 organisations between 1981 and 1995.[30] A semi-structured telephone interview of investigators was carried out to verify study results and publication status. Study results were classified as statistically significant or statistically non-significant. The mixed result refers to a situation in which at least one of multiple primary outcomes was statistically significant. The rate of publication was 85% for statistically significant results, 86% for non-significant results, and only 14% for the mixed results.[30]

The publication status of 68 RCTs processed through the pharmacy department of an eye hospital since 1963 was examined by Wormald *et al.*[31] This study was published only as a conference abstract and additional data were provided in Dwan *et al.*[41] The rate of publication was 93% for statistically significant results and 71% for non-significant results.

Zimpel and Windeler investigated the subsequent publication of 140 medical theses on complementary medical subjects.[32] Results were classified as positive or non-positive (this classification is slightly unclear as the article was published in German). Publication status was tracked by searching MEDLINE and by personal communication with authors or supervisors. The rate of publication was 40% for positive results and 28% for negative results.[32]

Regulatory cohort studies

No regulatory cohort studies of publication bias were included in the previous *HTA* report. In this updated review we identified four regulatory cohort studies that examined formal publication of clinical trials submitted to regulatory authorities (Appendix 6).[42–45] Of the four regulatory cohort studies, two did not specify clinical fields[42,44] and two focused on antidepressants.[43,45] One study was not included because the association of journal publication and study results was not reported.[46]

Melander *et al.* conducted a study of 42 randomised placebo-controlled trials of five antidepressants submitted by industry to the Swedish drug regulatory authority for marketing approval.[43] Studies were classified according to whether they found the test drug was significantly more effective than the placebo with the primary outcome. Publication status (including stand-alone, pooled or multiple publications) of the trials was investigated by searching electronic bibliographic

databases and contacting the companies. All 21 studies with significant results were published (stand-alone or pooled) while only 81% of studies with non-significant results were published.[43]

Turner *et al*. examined 74 clinical trials of 12 antidepressant agents submitted to the US Food and Drug Administration (FDA) between 1987 and 2004.[45] Trial results were classified as positive, questionable or negative according to the FDA's regulatory decisions. Publication status of trials was determined by searching literature databases and contacting trial sponsors. The rate of publication was 97% for studies with positive results, 50% for studies with questionable results, and 33% for studies with negative results.[45]

In a study by Lee *et al*., formal publication of 909 trials supporting 90 new drugs approved by the FDA between 1998 and 2000 was verified by searches of PubMed and other databases.[42] Statistical significance of the primary outcome was defined as being $p < 0.05$ or a CI excluding no difference. For equivalency or non-inferiority trials, the statistically significant result refers to those with '$p > 0.05$ or a CI including no difference or a CI excluding the pre-specified difference described in the trial'.[42] It was reported that the rate of formal publication was higher for trials with significant results than those with non-significant results (66% versus 36%).

Similar to the above study, a more recently published study included all efficacy trials supporting new drug applications approved by the FDA from 2001 to 2002.[44] Favourable results were those being significantly ($p < 0.05$) in favour of the new drug or those confirming equivalence in non-inferiority trials. Trials with favourable results were more likely to be published compared with trials with not favourable or unknown results (82% versus 64%).[44]

Cohorts of meeting abstracts

In the 2000 *HTA* report on publication bias,[2] we identified eight reports that examined the association between study results and the subsequent full publication of research initially presented as abstracts in meetings or journals.[47–54] A Cochrane Methodology Review included 79 studies of the subsequent full publication of biomedical research results initially presented as abstracts or in summary form.[55] Sixteen of the 79

studies reported data on the rate of publication by significance or importance of study results. Our updated search identified 22 additional cohort studies of research abstracts that provided data on publication bias (for details of all these studies, see Appendix 7).[56–77] Almost all of the 30 cohort studies of conference abstracts were restricted to a specific clinical field, such as emergency medicine, anaesthesiology, perinatal medicine, cystic fibrosis or oncology. The rate of full publication of meeting abstracts ranged from 37% to 81% for statistically significant results, and from 22% to 70% for non-significant results.

Manuscript cohort studies

We identified four studies of cohorts of manuscripts submitted to journals (Appendix 8).[78–81] Two studies examined manuscripts submitted to general medical journals [*Journal of the American Medical Association* (*JAMA*), *British Medical Journal* (*BMJ*), *The Lancet* and *Annals of Internal Medicine*][78,81] and two used manuscripts submitted to the *Journal of Bone and Joint Surgery (American Version)*.[79,80] The study results of submitted papers were classified according to the significance of statistical tests ($p < 0.05$ or not) in the two studies of manuscripts submitted to general medical journals.[78,81] In the studies of manuscripts submitted to the *Journal of Bone and Joint Surgery*, results were classified as being positive or negative or neutral, although the definitions of these outcomes may be different between the two studies (Appendix 8).[79,80] Results from these studies suggested that the acceptance of submitted papers for publication by journals was not significantly associated with the direction or strength of their findings.

In the study of Olson *et al*.,[81] 133 accepted manuscripts were further examined and it was found that time to publication was not associated with statistical significance (median 7.8 months for positive and 7.6 months for negative results, $p = 0.44$).[82] However, a subgroup analysis of 156 manuscripts with a high level of evidence (level I or II) in the study by Okike *et al*. found that the acceptance rate was significantly higher for studies with positive or neutral results than for studies with negative results (37%, 36% and 5% respectively; $p = 0.02$).[80]

These manuscript cohort studies are generally well designed and conducted. Although no conflict of interest was declared in these studies, this kind of study will always need support or collaboration

from editors of the journals. In prospective studies, editors' decisions on the acceptance of manuscripts may be influenced by their awareness of the ongoing study.[81]

Pooled analyses of cohort studies

Results from different studies of publication bias have been quantitatively combined in previous reviews,[25,55,83] although it is still controversial because of heterogeneity across individual studies.[41] Pooled estimates may improve statistical power and the generalisability of results. In this review, the association between study results and the possibility of subsequent publication was measured by using odds ratios (OR). Heterogeneity across studies within each subgroup was measured using the I^2 statistic.[16] A random-effects model was used in meta-analyses.

The formal publication of statistically significant results ($p < 0.05$) could be compared with that of non-significant results in four inception cohort studies, one regulatory cohort study, 12 abstract cohort studies and two manuscript cohort studies (*Figure 2*). The rate of publication of studies in the four inception cohorts ranged from 60% to 93% for significant results and from 20% to 86% for non-significant results. The rate of full publication of meeting abstracts ranged from 37% to 81% for statistically significant results, and from 22% to 70% for non-significant results. Heterogeneity across the four cohort studies from the inception subgroup was statistically significant ($I^2 = 61\%$, $p = 0.05$). There was no statistically significant heterogeneity across studies within the cohort studies of abstracts and cohort studies of manuscripts. The pooled odds ratio for publication bias by statistical significance of results was 2.40 (95% CI: 1.18 to 4.88) for the four inception cohort studies, 1.62 (95% CI: 1.34 to 1.96) for the 12 abstract cohort studies, and 1.15 (95% CI: 0.64 to 2.10) for the two manuscript cohort studies (*Figure 2*).

To include data from other cohort studies, a positive result was defined as being important or confirmatory or significant, while a 'non-positive' result included negative, non-important, inconclusive or non-significant results. This more inclusive definition of positive results allowed the inclusion of all 12 inception cohort studies, four regulatory cohort studies, 29 abstract cohort studies, and four manuscript cohort studies (*Figure 3*). There was statistically significant heterogeneity across cohort studies within the

inception ($p = 0.06$), regulatory ($p = 0.04$) and abstract subgroups ($p < 0.001$). Pooled estimates of odds ratios consistently indicated that studies with positive results were more likely to be published than studies with non-positive results, but this was not true after the submission to journals (*Figure 3*).

Types of studies included in the cohort studies varied, and included basic experimental, observational and qualitative research, and clinical trials. When the analyses were restricted to clinical trials, the result was not significantly different from that based on all studies. Although the number of cohort studies that could be included was small, clear evidence of publication bias can still be observed when the analysis was restricted to clinical trials (*Figure 4* and *Figure 5*).

We constructed funnel plots separately for the four subgroups of cohort studies (*Figure 6*). The asymmetry of these funnel plots was tested using the method recommended by Peters *et al*.[9] (This is a method of linear regression analysis, using log odds ratio as the dependent variable and the inverse of the total sample size as the independent variable. Please see Chapter 8 for more details about this method.) There was no statistically significant asymmetry for the funnel plots of inception cohort studies ($p = 0.178$), regulatory cohort studies ($p = 0.262$), abstract cohort studies ($p = 0.233$) or manuscript cohort studies ($p = 0.942$).

Main results of the above meta-analyses of cohort studies are summarised in *Table 1*.

Factors associated with publication bias

Some cohort studies have examined the impacts of certain factors on the publication of research. The factors investigated included study design, type of study, sample size, funding source and investigators' characteristics. Easterbrook *et al*. (1991) conducted subgroup analyses to examine susceptibility to publication bias among various subgroups of studies. They found that observational, laboratory-based experimental studies and non-randomised trials had greater risk of publication bias than RCTs. Factors associated with less bias included a concurrent comparison group, a high investigator rating of study importance and a sample size above 20.[22]

Dickersin *et al*. (1992) investigated the association between the risk of publication bias and type of study (observational, clinical trial), multi- or single
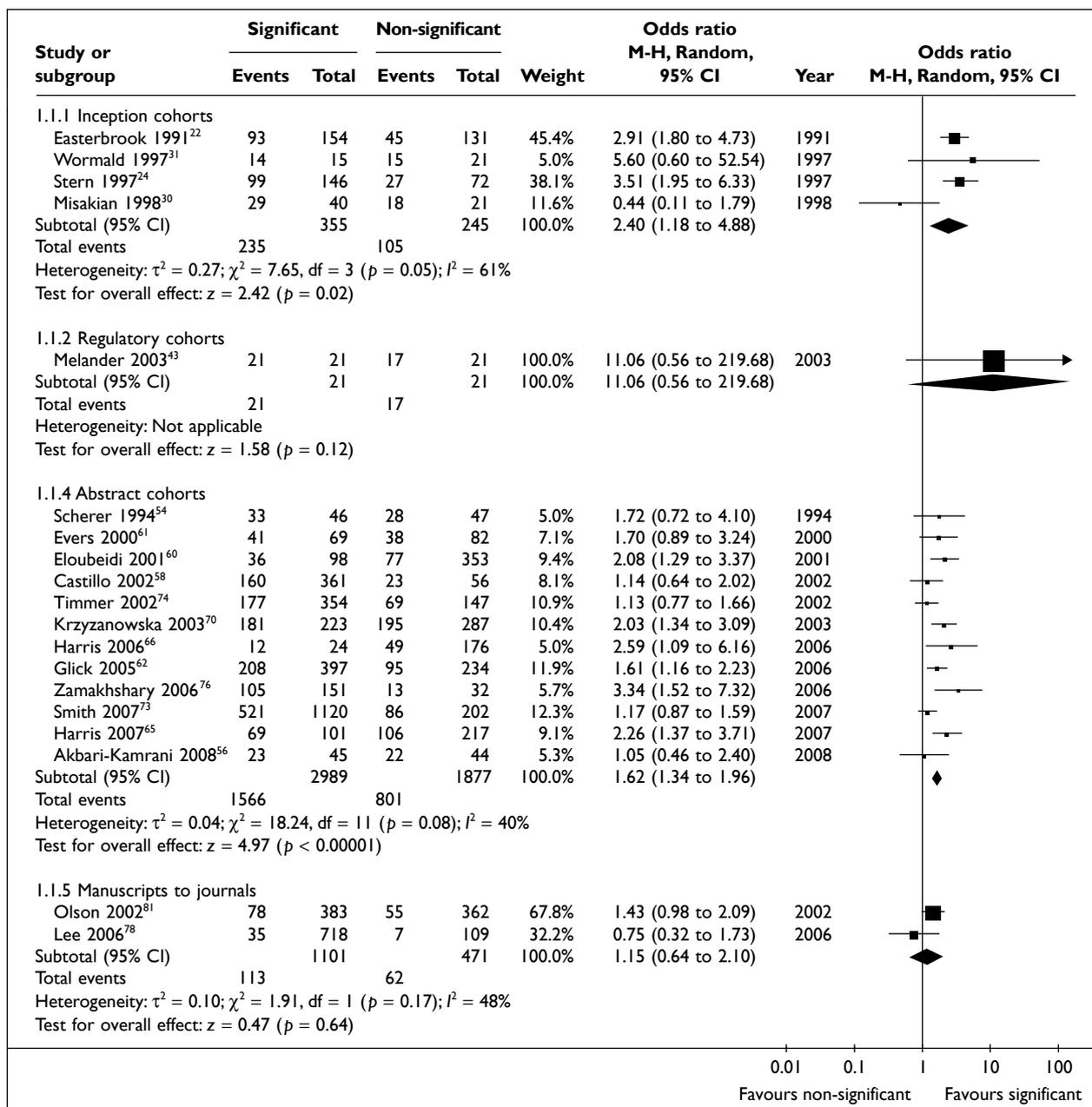
Study or subgroup	Significant		Non-significant		Weight	Odds ratio M-H, Random, 95% CI	Year	Odds ratio M-H, Random, 95% CI
	Events	Total	Events	Total				
1.1.1 Inception cohorts								
Easterbrook 1991[22]	93	154	45	131	45.4%	2.91 (1.80 to 4.73)	1991	
Wormald 1997[31]	14	15	15	21	5.0%	5.60 (0.60 to 52.54)	1997	
Stern 1997[24]	99	146	27	72	38.1%	3.51 (1.95 to 6.33)	1997	
Misakian 1998[30]	29	40	18	21	11.6%	0.44 (0.11 to 1.79)	1998	
Subtotal (95% CI)		355		245	100.0%	2.40 (1.18 to 4.88)		
Total events	235		105					
Heterogeneity: $\tau^2 = 0.27$; $\chi^2 = 7.65$, df = 3 ($p = 0.05$); $I^2 = 61\%$								
Test for overall effect: z = 2.42 ($p = 0.02$)								
1.1.2 Regulatory cohorts								
Melander 2003[43]	21	21	17	21	100.0%	11.06 (0.56 to 219.68)	2003	
Subtotal (95% CI)		21		21	100.0%	11.06 (0.56 to 219.68)		
Total events	21		17					
Heterogeneity: Not applicable								
Test for overall effect: z = 1.58 ($p = 0.12$)								
1.1.4 Abstract cohorts								
Scherer 1994[54]	33	46	28	47	5.0%	1.72 (0.72 to 4.10)	1994	
Evers 2000[61]	41	69	38	82	7.1%	1.70 (0.89 to 3.24)	2000	
Eloubeidi 2001[60]	36	98	77	353	9.4%	2.08 (1.29 to 3.37)	2001	
Castillo 2002[58]	160	361	23	56	8.1%	1.14 (0.64 to 2.02)	2002	
Timmer 2002[74]	177	354	69	147	10.9%	1.13 (0.77 to 1.66)	2002	
Krzyzanowska 2003[70]	181	223	195	287	10.4%	2.03 (1.34 to 3.09)	2003	
Harris 2006[66]	12	24	49	176	5.0%	2.59 (1.09 to 6.16)	2006	
Glick 2005[62]	208	397	95	234	11.9%	1.61 (1.16 to 2.23)	2006	
Zamakhshary 2006[76]	105	151	13	32	5.7%	3.34 (1.52 to 7.32)	2006	
Smith 2007[73]	521	1120	86	202	12.3%	1.17 (0.87 to 1.59)	2007	
Harris 2007[65]	69	101	106	217	9.1%	2.26 (1.37 to 3.71)	2007	
Akbari-Kamrani 2008[56]	23	45	22	44	5.3%	1.05 (0.46 to 2.40)	2008	
Subtotal (95% CI)		2989		1877	100.0%	1.62 (1.34 to 1.96)		
Total events	1566		801					
Heterogeneity: $\tau^2 = 0.04$; $\chi^2 = 18.24$, df = 11 ($p = 0.08$); $I^2 = 40\%$								
Test for overall effect: z = 4.97 ($p < 0.00001$)								
1.1.5 Manuscripts to journals								
Olson 2002[81]	78	383	55	362	67.8%	1.43 (0.98 to 2.09)	2002	
Lee 2006[78]	35	718	7	109	32.2%	0.75 (0.32 to 1.73)	2006	
Subtotal (95% CI)		1101		471	100.0%	1.15 (0.64 to 2.10)		
Total events	113		62					
Heterogeneity: $\tau^2 = 0.10$; $\chi^2 = 1.91$, df = 1 ($p = 0.17$); $I^2 = 48\%$								
Test for overall effect: z = 0.47 ($p = 0.64$)								

0.01   0.1   1   10   100
Favours non-significant   Favours significant

FIGURE 2 *Rate of publication of statistically significant versus non-significant results: all studies. After excluding Misakian and Bero (1998)[30] from inception cohort studies, pooled OR = 3.19 (2.21 to 4.61); heterogeneity: $I^2 = 0\%$, p = 0.79.*

centre, sample size, funding source and principal investigator (PI) characteristics (such as gender, degree and rank). They found that none of the factors examined was associated with publication bias.[20]

Dickersin and Min (1993) reported that the OR for publication bias was 0.84 (95% CI: 0.07 to 9.68) for multicentre studies compared to 21.14 (95% CI: 2.60 to 171.7) for single centre studies.[21] In addition, the risk of publication bias was different between studies with a female PI (OR = 0.47; 95% CI: 0.02 to 11.61) and studies with a male PI (OR = 20.70; 95% CI: 2.61 to 164.2). One interesting explanation for the difference in study publication between female and male PIs posted by Dickersin and Min was that 'women have fewer studies to manage', related to their relatively lower rank (35% of women PIs were professors compared to 65% of male PIs), and are thus less selective in study publication. They did not find an association between publication bias and other study features

such as the use of randomisation or blinding, having a comparison group or a large sample size.[21]

Stern and Simes (1997) found that the risk of publication bias tended to be greater for clinical trials (OR = 3.13; 95% CI: 1.76 to 5.58) than other studies (for all quantitative studies OR = 2.3; 95% CI: 1.47 to 3.66). When analysis was restricted to studies with a sample size ≥ 100, publication bias was still evident (HR = 2.00; 95% CI: 1.09 to 3.66).[24]

Discussions of findings from cohort studies

The updated review identified limited new evidence on publication bias based on a follow-up of research protocols, and a large number of new studies on subsequent publication of meeting abstracts. Updated analyses yielded results similar to those from the 2000 *HTA* report and other

existing reviews: studies with statistically significant or positive results are more likely to be formally published than those with non-significant or non-positive results.[2,25,41,55,83] Dickersin in 1997 combined the results from four inception cohort studies[20–22,24] and found that the pooled adjusted odds ratio for publication bias (publication of studies with significant or important results versus those with unimportant results) was 2.54 (95% CI: 1.44 to 4.47).[25] A recent systematic review of inception cohort studies of clinical trials found the existence of publication bias and outcome reporting bias, although pooled meta-analysis was not conducted due to perceived differences between studies.[41] A Cochrane methodology review of publication bias by Hopewell *et al.*[83] included five inception cohort studies of trials registered before the main results were known,[20,21,23,24,26] in which the pooled odds ratio for publication bias was 3.90 (95% CI: 2.68 to 5.68). In a Cochrane methodology review by Scherer *et al.*, the association between the subsequent full publication and study results
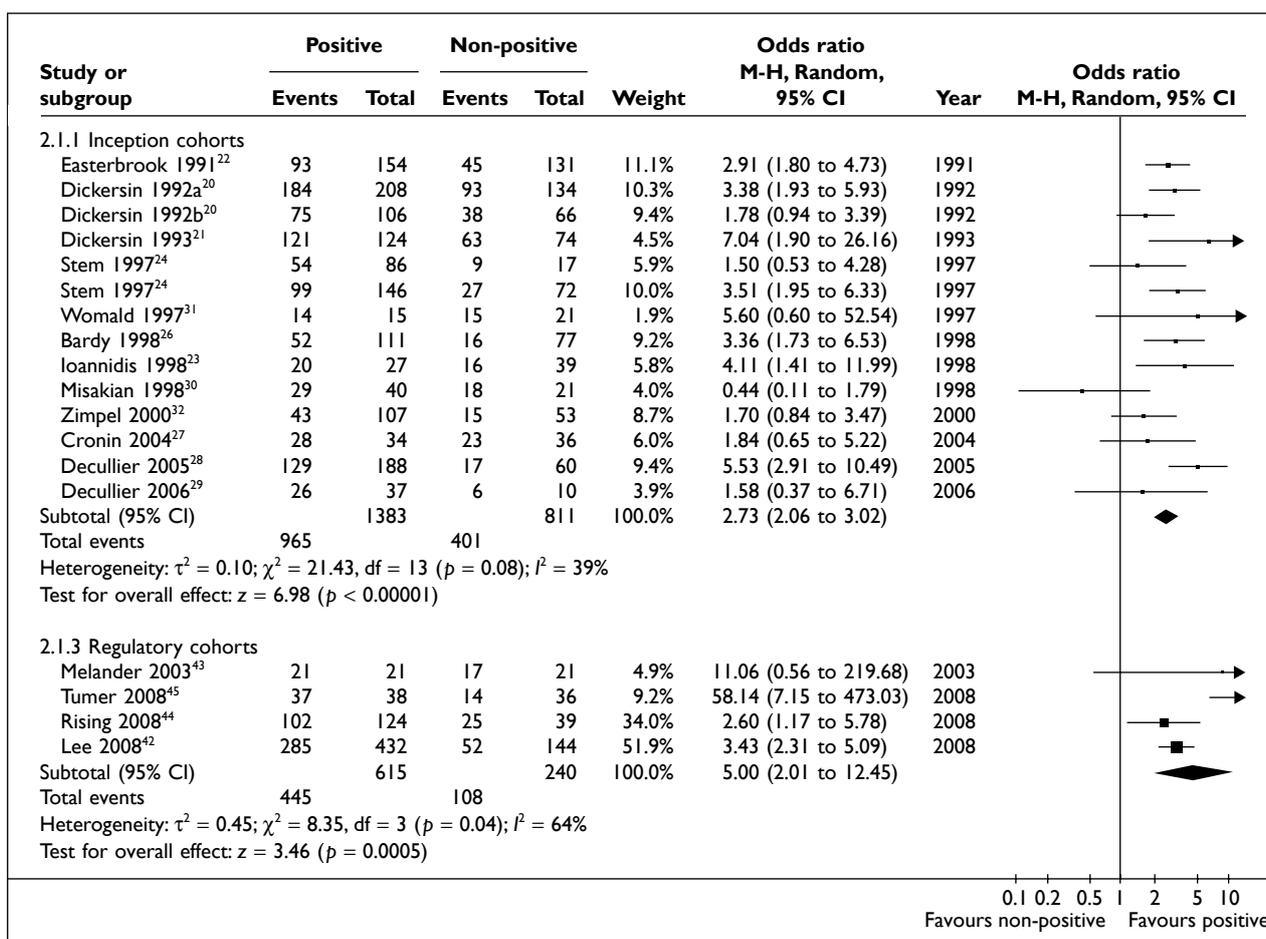
Study or subgroup	Positive		Non-positive		Weight	Odds ratio M-H, Random, 95% CI	Year	Odds ratio M-H, Random, 95% CI
	Events	Total	Events	Total				
2.1.1 Inception cohorts								
Easterbrook 1991[22]	93	154	45	131	11.1%	2.91 (1.80 to 4.73)	1991	
Dickersin 1992a[20]	184	208	93	134	10.3%	3.38 (1.93 to 5.93)	1992	
Dickersin 1992b[20]	75	106	38	66	9.4%	1.78 (0.94 to 3.39)	1992	
Dickersin 1993[21]	121	124	63	74	4.5%	7.04 (1.90 to 26.16)	1993	
Stem 1997[24]	54	86	9	17	5.9%	1.50 (0.53 to 4.28)	1997	
Stem 1997[24]	99	146	27	72	10.0%	3.51 (1.95 to 6.33)	1997	
Womald 1997[31]	14	15	15	21	1.9%	5.60 (0.60 to 52.54)	1997	
Bardy 1998[26]	52	111	16	77	9.2%	3.36 (1.73 to 6.53)	1998	
Ioannidis 1998[23]	20	27	16	39	5.8%	4.11 (1.41 to 11.99)	1998	
Misakian 1998[30]	29	40	18	21	4.0%	0.44 (0.11 to 1.79)	1998	
Zimpel 2000[32]	43	107	15	53	8.7%	1.70 (0.84 to 3.47)	2000	
Cronin 2004[27]	28	34	23	36	6.0%	1.84 (0.65 to 5.22)	2004	
Decullier 2005[28]	129	188	17	60	9.4%	5.53 (2.91 to 10.49)	2005	
Decullier 2006[29]	26	37	6	10	3.9%	1.58 (0.37 to 6.71)	2006	
Subtotal (95% CI)		1383		811	100.0%	2.73 (2.06 to 3.02)		
Total events	965		401					
Heterogeneity: $\tau^2 = 0.10$; $\chi^2 = 21.43$, df = 13 ($p = 0.08$); $I^2 = 39\%$								
Test for overall effect: $z = 6.98$ ($p < 0.00001$)								
2.1.3 Regulatory cohorts								
Melander 2003[43]	21	21	17	21	4.9%	11.06 (0.56 to 219.68)	2003	
Tumer 2008[45]	37	38	14	36	9.2%	58.14 (7.15 to 473.03)	2008	
Rising 2008[44]	102	124	25	39	34.0%	2.60 (1.17 to 5.78)	2008	
Lee 2008[42]	285	432	52	144	51.9%	3.43 (2.31 to 5.09)	2008	
Subtotal (95% CI)		615		240	100.0%	5.00 (2.01 to 12.45)		
Total events	445		108					
Heterogeneity: $\tau^2 = 0.45$; $\chi^2 = 8.35$, df = 3 ($p = 0.04$); $I^2 = 64\%$								
Test for overall effect: $z = 3.46$ ($p = 0.0005$)								

0.1 0.2 0.5   1   2   5   10
Favours non-positive   Favours positive

FIGURE 3 *(continued on next page) Rate of publication of positive versus non-positive results: all studies. After excluding Misakian and Bero (1998)[30] from inception cohort studies, pooled OR = 2.93 (2.31 to 3.71); heterogeneity: $I^2 = 17\%$, p = 0.27.*
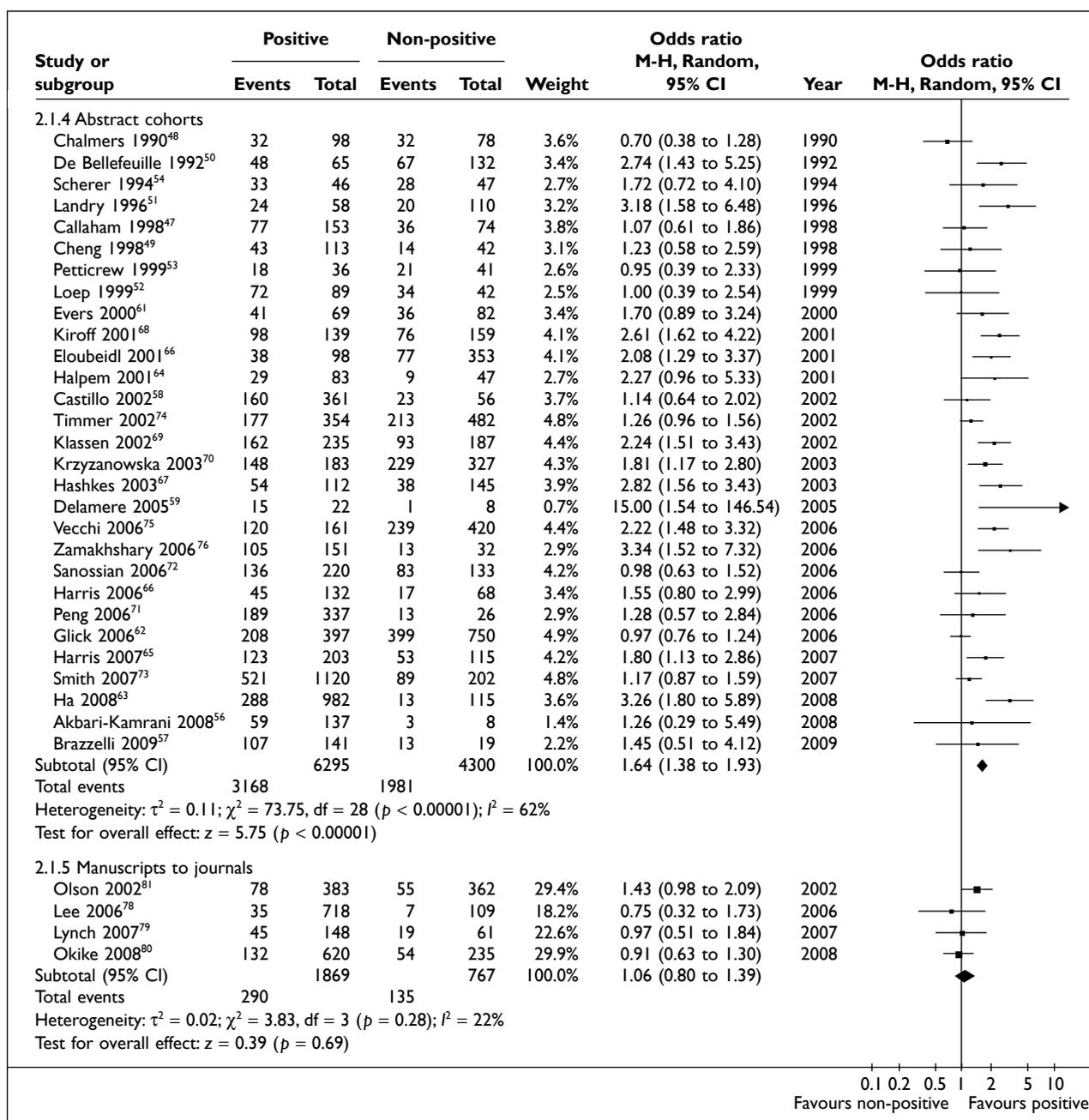
Study or subgroup	Positive		Non-positive		Weight	Odds ratio M-H, Random, 95% CI	Year	Odds ratio M-H, Random, 95% CI
	Events	Total	Events	Total				
2.1.4 Abstract cohorts								
Chalmers 1990[48]	32	98	32	78	3.6%	0.70 (0.38 to 1.28)	1990	
De Bellefeuille 1992[50]	48	65	67	132	3.4%	2.74 (1.43 to 5.25)	1992	
Scherer 1994[54]	33	46	28	47	2.7%	1.72 (0.72 to 4.10)	1994	
Landry 1996[51]	24	58	20	110	3.2%	3.18 (1.58 to 6.48)	1996	
Callaham 1998[47]	77	153	36	74	3.8%	1.07 (0.61 to 1.86)	1998	
Cheng 1998[49]	43	113	14	42	3.1%	1.23 (0.58 to 2.59)	1998	
Petticrew 1999[53]	18	36	21	41	2.6%	0.95 (0.39 to 2.33)	1999	
Loep 1999[52]	72	89	34	42	2.5%	1.00 (0.39 to 2.54)	1999	
Evers 2000[61]	41	69	36	82	3.4%	1.70 (0.89 to 3.24)	2000	
Kiroff 2001[68]	98	139	76	159	4.1%	2.61 (1.62 to 4.22)	2001	
Eloubeidl 2001[66]	38	98	77	353	4.1%	2.08 (1.29 to 3.37)	2001	
Halpem 2001[64]	29	83	9	47	2.7%	2.27 (0.96 to 5.33)	2001	
Castillo 2002[58]	160	361	23	56	3.7%	1.14 (0.64 to 2.02)	2002	
Timmer 2002[74]	177	354	213	482	4.8%	1.26 (0.96 to 1.56)	2002	
Klassen 2002[69]	162	235	93	187	4.4%	2.24 (1.51 to 3.43)	2002	
Krzyzanowska 2003[70]	148	183	229	327	4.3%	1.81 (1.17 to 2.80)	2003	
Hashkes 2003[67]	54	112	38	145	3.9%	2.82 (1.56 to 3.43)	2003	
Delamere 2005[59]	15	22	1	8	0.7%	15.00 (1.54 to 146.54)	2005	
Vecchi 2006[75]	120	161	239	420	4.4%	2.22 (1.48 to 3.32)	2006	
Zamakhshary 2006[76]	105	151	13	32	2.9%	3.34 (1.52 to 7.32)	2006	
Sanossian 2006[72]	136	220	83	133	4.2%	0.98 (0.63 to 1.52)	2006	
Harris 2006[66]	45	132	17	68	3.4%	1.55 (0.80 to 2.99)	2006	
Peng 2006[71]	189	337	13	26	2.9%	1.28 (0.57 to 2.84)	2006	
Glick 2006[62]	208	397	399	750	4.9%	0.97 (0.76 to 1.24)	2006	
Harris 2007[65]	123	203	53	115	4.2%	1.80 (1.13 to 2.86)	2007	
Smith 2007[73]	521	1120	89	202	4.8%	1.17 (0.87 to 1.59)	2007	
Ha 2008[63]	288	982	13	115	3.6%	3.26 (1.80 to 5.89)	2008	
Akbari-Kamrani 2008[56]	59	137	3	8	1.4%	1.26 (0.29 to 5.49)	2008	
Brazzelli 2009[57]	107	141	13	19	2.2%	1.45 (0.51 to 4.12)	2009	
Subtotal (95% CI)		6295		4300	100.0%	1.64 (1.38 to 1.93)		
Total events	3168		1981					

Heterogeneity: $\tau^2 = 0.11$; $\chi^2 = 73.75$, df = 28 ($p < 0.00001$); $I^2 = 62\%$
Test for overall effect: $z = 5.75$ ($p < 0.00001$)

Study or subgroup	Positive		Non-positive		Weight	Odds ratio M-H, Random, 95% CI	Year	Odds ratio M-H, Random, 95% CI
2.1.5 Manuscripts to journals								
Olson 2002[81]	78	383	55	362	29.4%	1.43 (0.98 to 2.09)	2002	
Lee 2006[78]	35	718	7	109	18.2%	0.75 (0.32 to 1.73)	2006	
Lynch 2007[79]	45	148	19	61	22.6%	0.97 (0.51 to 1.84)	2007	
Okike 2008[80]	132	620	54	235	29.9%	0.91 (0.63 to 1.30)	2008	
Subtotal (95% CI)		1869		767	100.0%	1.06 (0.80 to 1.39)		
Total events	290		135					

Heterogeneity: $\tau^2 = 0.02$; $\chi^2 = 3.83$, df = 3 ($p = 0.28$); $I^2 = 22\%$
Test for overall effect: $z = 0.39$ ($p = 0.69$)

0.1 0.2 0.5 1 2 5 10
Favours non-positive　　Favours positive

FIGURE 3 *Rate of publication of positive versus non-positive results: all studies. After excluding Misakian and Bero (1998)[30] from inception cohort studies, pooled OR = 2.93 (2.31 to 3.71); heterogeneity: $I^2 = 17\%$, p = 0.27 (continued).*

was examined in 16 of 79 abstract cohort studies.[55] According to these 16 abstract cohort studies, the subsequent full publication of conference abstracts was statistically significantly associated with positive study results (pooled OR = 1.28; 95% CI: 1.15 to 1.42).

Compared with the previous reviews on cohort studies of publication bias, our review is more inclusive in terms of types of studies and is the first to enable an explicit comparison of results from cohort studies of publication bias with fundamentally different sampling frames. Biased

selection for publication may affect research dissemination over the whole process from before study completion, to presentation of findings at conferences, manuscript submission to journals, and formal publication in journals. It seems that publication bias occurs mainly before the presentation of findings at conferences and before the submission of manuscript to journals (see *Figure 2* and *Figure 3*). The subsequent publication of conference abstracts was still biased but the extent of publication bias tended to be smaller compared with all studies conducted. After submission of

Study or subgroup	Significant		Non-significant		Weight	Odds ratio M-H, Random, 95% CI	Year	Odds ratio M-H, Random, 95% CI
	Events	Total	Events	Total				
3.1.1 Inception cohorts								
Wormald 1997[31]	14	15	15	21	14.0%	5.60 (0.60 to 52.54)	1997	
Stern 1997[24]	55	76	18	54	86.0%	5.24 (2.46 to 11.17)	1997	
Subtotal (95% CI)		91		75	100.0%	5.27 (2.57 to 10.80)		
Total events	69		33					
Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 0.00$, df = 1 ($p = 0.96$); $I^2 = 0\%$								
Test for overall effect: $z = 4.55$ ($p < 0.00001$)								
3.1.3 Regulatory cohorts								
Melander 2003[43]	21	21	17	21	100.0%	11.06 (0.56 to 219.68)	2003	
Subtotal (95% CI)		21		21	100.0%	11.06 (0.56 to 219.68)		
Total events	21		17					
Heterogeneity: Not applicable								
Test for overall effect: $z = 1.58$ ($p = 0.12$)								
3.1.4 Abstract cohorts								
Scherer 1994[54]	33	46	28	47	19.2%	1.72 (0.72 to 4.10)	1994	
Evers 2000[61]	41	69	38	82	13.1%	1.70 (0.89 to 3.24)	2000	
Timmer 2002[74]	84	140	47	99	24.3%	1.66 (0.99 to 2.79)	2002	
Krzyzanowska 2003[70]	181	223	195	287	29.3%	2.03 (1.34 to 3.09)	2003	
Akbari-Kamrani 2008[56]	23	45	22	44	13.9%	1.05 (0.46 to 2.40)	2008	
Subtotal (95% CI)		523		559	100.0%	1.73 (1.33 to 2.25)		
Total events	362		330					
Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 2.01$, df = 4 ($p = 0.73$); $I^2 = 0\%$								
Test for overall effect: $z = 4.10$ ($p < 0.0001$)								
3.1.5 Manuscripts to journals								
Olson 2002[81]	78	383	55	362	100.0%	1.43 (0.98 to 2.09)	2002	
Subtotal (95% CI)		383		362	100.0%	1.43 (0.98 to 2.09)		
Total events	78		55					
Heterogeneity: Not applicable								
Test for overall effect: $z = 1.84$ ($p = 0.07$)								

0.05   0.2   1   5   20
Favours non-significant   Favours significant

FIGURE 4 *Rate of publication of statistically significant versus non-significant results: clinical trials only.*

manuscript to journals, editorial decisions were not clearly associated with study results.

Limitations of the available evidence on publication bias

There are some caveats to the available evidence on publication bias. Study findings have been defined differently among the empirical studies assessing publication bias. The most objective method would be to classify quantitative results as statistically significant ($p < 0.05$) or not. However, this was not always possible or appropriate. When other methods were used to classify study results as important or not, bias may be introduced due to inevitable subjectivity.

The funnel plot asymmetry is not statistically significant for inception, regulatory, abstract cohort studies and manuscript cohort studies (see

Figure 6). However, there are reasons to suspect the existence of publication and reporting bias in studies of meeting abstracts. A large number of reports of full publication of research abstracts were assessed for inclusion into this review but did not mention the association between publication and study results and so were excluded. This association might not have been examined; or not reported when the association was not significant. As an example, Zaretsky and Imrie (2002)[77] reported no significant difference ($p = 0.53$) in the rate of subsequent publication of 57 meeting abstracts between statistically significant and non-significant results; but this study was not included in the analysis as insufficient data were provided.

Large cohort studies on publication bias usually included cases that were highly diverse in terms of research questions, designs and other study characteristics. Many factors (e.g. sample size,

Study or subgroup	Positive		Non-positive		Weight	Odds ratio M-H, Random, 95% CI	Year	Odds ratio M-H, Random, 95% CI
	Events	Total	Events	Total				
4.1.1 Inception cohorts								
Dickersin 1992[20]	84	96	52	72	19.2%	2.69 (1.22 to 5.96)	1992	
Dickersin 1993[21]	121	124	63	74	10.6%	7.04 (1.90 to 26.16)	1993	
Stern 1997a[24]	55	76	18	54	20.0%	5.24 (2.46 to 11.17)	1997	
Wormald 1997[31]	14	15	15	21	4.5%	5.60 (0.60 to 52.54)	1997	
Stern 1997b[24]	14	24	9	13	9.4%	0.62 (0.15 to 2.60)	1997	
Ioannidis 1998[23]	20	27	16	39	13.9%	4.11 (1.41 to 11.99)	1998	
Bardy 1998[26]	52	111	16	77	22.3%	3.36 (1.73 to 6.53)	1998	
Subtotal (95% CI)		473		350	100.0%	3.45 (2.19 to 5.42)		
Total events	360		189					

Heterogeneity: $\tau^2 = 0.10$; $\chi^2 = 8.47$, df = 6 ($p = 0.21$); $I^2 = 29\%$
Test for overall effect: $z = 5.37$ ($p < 0.00001$)

	Events	Total	Events	Total	Weight		Year	
4.1.3 Regulatory cohorts								
Melander 2003[43]	21	21	17	21	4.8%	11.06 (0.56 to 219.68)	2003	
Rising 2008[44]	102	124	25	39	33.8%	2.60 (1.17 to 5.78)	2008	
Tumer 2008[45]	37	38	14	36	9.0%	58.14 (7.15 to 473.03)	2008	
Lee 2008[42]	285	432	52	144	52.4%	3.43 (2.31 to 5.09)	2008	
Subtotal (95% CI)		615		240	100.0%	5.00 (2.01 to 12.45)		
Total events	445		108					

Heterogeneity: $\tau^2 = 0.45$; $\chi^2 = 8.35$, df = 3 ($p = 0.04$); $I^2 = 64\%$
Test for overall effect: $z = 3.46$ ($p = 0.0005$)

	Events	Total	Events	Total	Weight		Year	
4.1.4 Abstract cohorts								
Chalmers 1990[48]	32	98	32	78	10.3%	0.70 (0.38 to 1.29)	1990	
Scherer 1994[54]	33	46	28	47	7.7%	1.72 (0.72 to 4.10)	1994	
Cheng 1998[49]	43	113	14	42	8.9%	1.23 (0.58 to 2.59)	1998	
Loep 1999[52]	72	89	34	42	7.1%	1.00 (0.39 to 2.54)	1999	
Evers 2000[61]	41	69	38	82	9.9%	1.70 (0.89 to 3.24)	2000	
Klassen 2002[69]	162	235	93	187	12.8%	2.24 (1.51 to 3.34)	2002	
Timmer 2002[74]	84	140	86	186	12.3%	1.74 (1.12 to 2.72)	2002	
Krzyzanowska 2003[70]	148	183	229	327	12.4%	1.81 (1.17 to 2.80)	2003	
Delamere 2005[59]	15	22	1	8	1.9%	15.00 (1.54 to 146.54)	2005	
Vecchi 2006[75]	120	161	239	420	12.8%	2.22 (1.48 to 3.32)	2006	
Akbari-Kamrani 2008[56]	59	137	3	8	3.9%	1.26 (0.29 to 5.49)	2008	
Subtotal (95% CI)		1293		1427	100.0%	1.65 (1.28 to 2.12)		
Total events	809		797					

Heterogeneity: $\tau^2 = 0.07$; $\chi^2 = 17.23$; df = 10 ($p = 0.07$); $I^2 = 42\%$
Test for overall effect: $z = 3.91$ ($p < 0.0001$)

	Events	Total	Events	Total	Weight		Year	
4.1.5 Manuscripts to journals								
Olson 2002[81]	78	383	55	362	100.0%	1.43 (0.98 to 2.09)	2002	
Subtotal (95% CI)		383		362	100.0%	1.43 (0.98 to 2.09)		
Total events	78		55					

Heterogeneity: Not applicable
Test for overall effect: $z = 1.84$ ($p = 0.07$)

0.005   0.1   1   10   200
Favours non-positive   Favours positive

FIGURE 5 *Rate of publication of positive versus non-positive results: clinical trials only.*

design, research question and investigators' characteristics) may be associated with both study results and the possibility of publication. Adjusted analyses by some factors may be conducted but it was generally impossible to exclude the impact of confounding factors on the observed association between study results and formal publication.

There is very limited and conflicting evidence on factors (such as study design, sample size, etc.) that may be associated with publication bias. To improve the understanding of factors associated with publication bias, findings from qualitative research on the process of research dissemination may be helpful.[84,85]

17
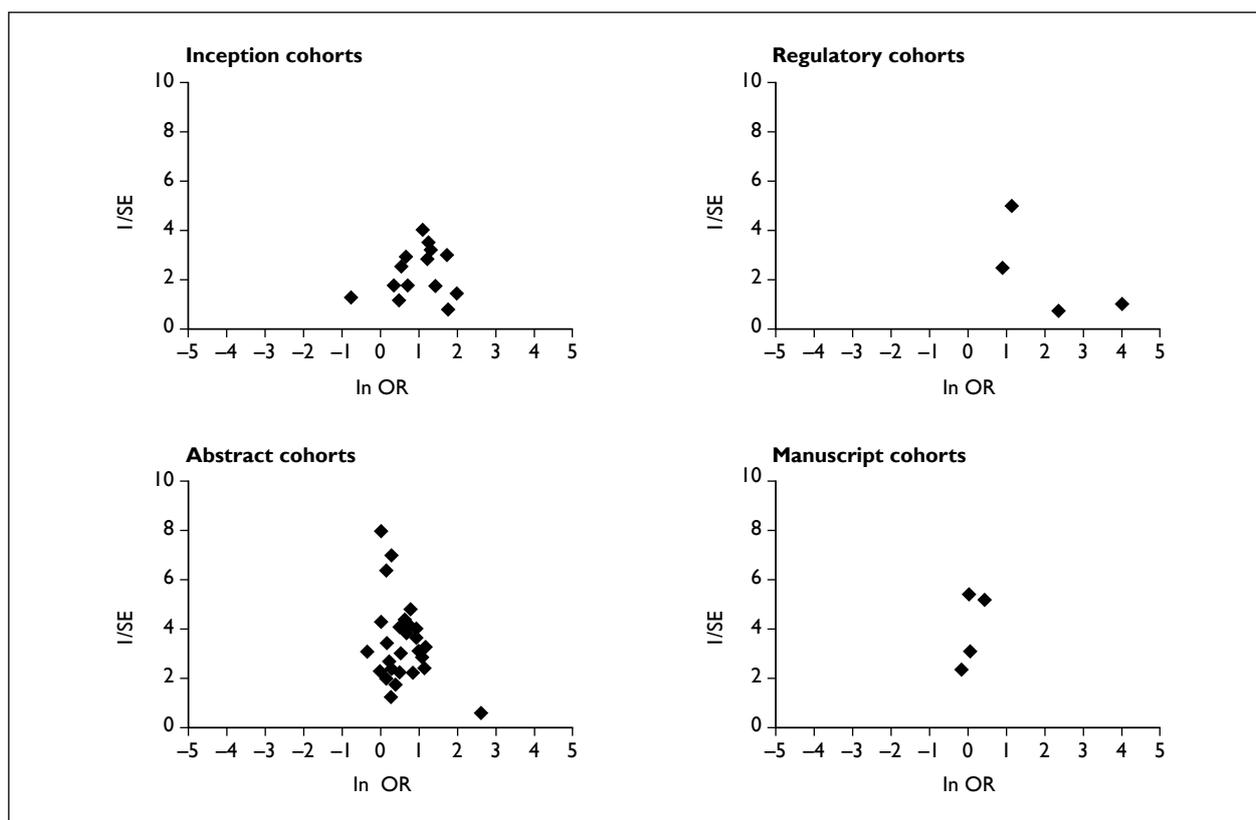
FIGURE 6 *Funnel plots – publication of positive and non-positive results. Funnel plot asymmetry test using the method of Peters et al.:[9] p = 0.178 for inception cohort studies; p = 0.262 for regulatory cohort studies; p = 0.233 for abstract cohort studies; and p = 0.942 for manuscript cohort studies.*

TABLE 1 *Results of meta-analyses of cohort studies of publication bias*

Cohort category	No. of cohort studies	Pooled odds ratio (95% CI)	Heterogeneity test: I^2 (*p* value)
Statistically significant vs non-significant results			
Inception cohorts	4	2.40 (1.18 to 4.88)	61% (0.05)
Regulatory cohorts	1	11.06 (0.56 to 219.68)	
Abstract cohorts	12	1.62 (1.34 to 1.96)	22% (0.24)
Manuscript cohorts	2	1.15 (0.64 to 2.10)	48% (0.17)
Positive vs non-positive results			
Inception cohorts	14	2.73 (2.06 to 3.62)	39% (0.06)
Regulatory cohorts	4	5.00 (2.01 to 12.45)	64% (0.04)
Abstract cohorts	29	1.62 (1.38 to 1.93)	62% (<0.001)
Manuscript cohorts	4	1.06 (0.80 to 1.39)	22% (0.28)

There was statistically significant heterogeneity within subgroups of inception, regulatory and abstract cohort studies (see *Table 1*). The observed heterogeneity may be a result of differences in study designs, research questions, how the cohorts were assembled, definitions of study results, and so on. For example, the statistically significant heterogeneity across inception cohort studies was due to one study by Misakian and Bero (see *Figure 2* and *Figure 3*).[30] After excluding this cohort study, there was no longer significant heterogeneity across inception cohort studies. The cohort study by Misakian and Bero included research on health effects of passive smoking, and the impact of statistical significance of results on publication may be different from studies of other research topics.[30]

The four cohorts of trials submitted to regulatory authorities showed greater extent of publication bias than other subgroups of cohort studies (see *Figure 3*).[42–45] Only 855 primary studies were included in the regulatory cohort studies, and two of the four regulatory cohort studies focused on trials of antidepressants.[43,45] Therefore, the four regulatory cohort studies may be a biased selection of all possible cases.

Conclusions

Despite many caveats about the available empirical evidence on publication bias, there is little doubt that dissemination of research findings is likely to be a biased process. There is consistent empirical evidence that the publication of a study that exhibits statistically significant or 'important' results is more likely to occur than the publication of a study that does not show such results. Indirect evidence indicates that publication bias occurs mainly before the presentation of findings at conferences and the submission of manuscripts to journals.

Chapter 4

Evidence of different types of dissemination bias

This chapter reviews available empirical evidence on different types of research dissemination bias, including outcome reporting bias, time lag publication bias, grey literature bias, language bias, citation bias, duplicate or multiple publication bias, place of publication bias, database bias, country bias and media attention bias.

Outcome reporting bias

Outcome reporting bias occurs when studies with multiple outcomes report only some of the outcomes measured and the selection of an outcome for reporting is associated with the statistical significance or importance of the result. This bias is due to the incomplete reporting within published studies, and is also called 'within-study reporting bias' in order to distinguish it from selective non-reporting of a whole study,[86] or publication bias 'in situ'.[87]

Number of outcomes measured within trials

The existence of a large number of measured or calculated outcomes within a study is the prerequisite of selective reporting bias, which is present in almost all research studies. The selection of outcomes to report can be further classified into three categories:[87] (1) the selection of outcomes investigated, (2) the selection of methods to measure the selected outcome, and (3) the selection of results of multiple subgroup analyses. A large number of results can be generated by the combination of all possible choices. Pocock *et al.* (1987) found that the median number of reported end points was six per trial.[88] They also discussed selective reporting of results and related issues of subgroup analyses, repeated measurements over time, multiple treatment groups, and multiple tests of significance.[88] Tannock (1996) examined 32 RCTs published in 1992 and found that the median number of therapeutic end points per trial was five (range 2–19) and 13 trials did not define their primary end point.[89] Each of the 32 trials, on average, reported six (range 1–31) statistical

comparisons of major outcome parameters; and more than half of the implied statistical comparisons had not been reported.[89]

The number of outcomes estimated from published articles may underestimate the actual number of outcomes measured in trial studies. Based on information from trial protocols, Chan *et al.*[6,7] found that the median number of efficacy outcomes was 20 per trial and the median number of harm outcomes was six or five per trial.

Although outcome reporting bias was highly suspected, there was very limited empirical research included in the 2000 *HTA* report.[90] The updated review has identified many recently published empirical studies that provided direct evidence with which to assess the existence and extent of outcome reporting bias (see Appendix 9 for details of the included studies).

Direct evidence on outcome reporting bias

The most direct evidence on the existence and extent of outcome reporting bias is from studies that compared outcomes specified in research protocols and those reported in subsequent articles. A pilot study by Hahn *et al.* (2002)[36] was the first attempt to compare outcomes specified in trial protocols approved by a local REC and results reported in subsequent publications. They compared outcomes in 15 pairs of protocols and journal articles. Six of the 15 studies stated primary outcome variables in their protocols and four used the same outcomes as primary outcomes in the reports. An analysis plan was mentioned in eight studies, but the plan was followed in only one published report.

Chan and colleagues provided the most direct evidence on outcome reporting bias by investigating a cohort of 102 RCT protocols approved by the Danish REC from 1994 to 1995,[6] and another cohort of 48 RCT protocols approved by the Canadian Institutes of Health Research from 1990 to 1998.[7] Data on unreported and reported

outcomes were collated based on trial protocols, subsequently published journal articles, and a survey of trialists. If a published article provided insufficient data for meta-analysis, the outcome was defined as being incompletely reported. They found that 50% of efficacy outcomes and 65% of harm outcomes were incompletely reported in the Danish cohort; and 31% and 59% respectively in the Canadian cohort. Primary outcomes specified in protocols were different from primary outcomes stated in the corresponding journal articles in 62% (Danish cohort) and 40% (Canadian cohort) of cases. Statistically significant outcomes were more likely to be fully reported than non-significant outcomes. The odds ratio of an efficacy outcome being fully reported if it were statistically significant versus non-significant was 2.4 (95% CI: 1.4 to 4.0) in the Danish cohort and 2.7 (95% CI: 1.5 to 5.0) in the Canadian cohort. The biased reporting of significant outcomes appears more severe for harm data, the odds ratios were 4.7 (95% CI: 1.8 to 12.0) and 7.7 (95% CI: 0.5 to 111) respectively.[6,7]

Further work by Chan and Altman (2005) identified 519 RCTs indexed in PubMed in December 2000, and they conducted a survey of trialists to obtain information on unreported outcomes.[5] The median proportion of incompletely reported outcomes per trial was 42% for efficacy outcomes and 50% for harm outcomes. The pooled odds ratio for outcome reporting bias was 2.0 (95% CI: 1.6 to 2.7) for efficacy outcomes and 1.9 (95% CI: 1.1 to 3.5) for harm outcomes. Reasons given by authors for not reporting efficacy and harm outcomes included space constraints (47% and 25%), lack of clinical importance (37% and 75%), and being statistically non-significant (24% and 50%).[5]

Ghersi *et al*. (2006) compared 103 published RCTs and their protocols approved by Central Sydney Area Health Service REC from 1992 to 1996.[91] They found that 17% of primary outcomes specified in the protocols and 15% reported in articles differed between the protocols and publications. Trials for which all comparisons were statistically significant were more likely to report all outcomes fully ($p = 0.06$). As the study by Ghersi *et al*. was presented only as an abstract there was a lack of information on its study design and other results.[91]

Other evidence

One consequence of outcome reporting bias is that many trials cannot be included in meta-analyses because of incompletely reported outcomes in published papers. Although unreported data may be available from trialists, such communication can be time consuming and often does not result in additional data becoming available.

McCormack *et al*.[92] compared the results of a meta-analysis based on published data with an updated IPD (individual patient data, where the complete original datasets of the included studies are used to pool study data, rather than simply using summary measures from published reports) meta-analysis of trials of hernia surgery. For the outcome of hernia recurrence, the number of contributing RCTs was similar and the results were not significantly different between the two analyses. For the outcome of persisting pain, IPD meta-analysis included many more RCTs (3 versus 20) and provided qualitatively divergent results, as compared with the meta-analysis of published data. This case study indicates that some outcomes (e.g. persisting pain) may be more vulnerable to selective reporting than other outcomes (e.g. hernia recurrence).[92]

In a recent study, Bekkering *et al*. examined 767 observational studies (with 3284 results) and found that only 61% of the reported results could be used in meta-analyses investigating dose–response associations between diet and prostate or bladder cancer.[93] Usable results were more likely to indicate the existence of the association than those that were not usable.[93]

Furukawa *et al*.[94] investigated the association between the proportion of contributing RCTs and the pooled estimates of 156 Cochrane systematic reviews. A median of 46% [interquartile range (IQR) 20% to 75%] of identified RCTs in each meta-analysis contributed to the pooled estimates. The results of regression analysis revealed a general trend that the greater the proportion of contributing RCTs the smaller the treatment effect. It was concluded that outcome reporting may be biased.[94]

A methodological study by Williamson and Gamble[95] provided a motivating example and four cases in which results of Cochrane reviews were compared with results of sensitivity analysis (by imputation) when within-study outcome reporting bias was suspected. The example was a meta-analysis of beta-lactam versus a combination of beta-lactam and aminoglycoside in the treatment of cancer patients with neutropenia, in which only five of the nine eligible RCTs could be included. They found that the pooled treatment effect was considerably decreased in sensitivity analysis where

the missing results were imputed. For the other four selected cases, within-study selection was suspected in several trials but the impact on the conclusions of the meta-analyses was minimal.[95]

Scharf and Colevas compared adverse events reported in 22 published articles and corresponding protocols or data from Clinical Data Update System (CDUS, the National Cancer Institute's electronic database of clinical trial information).[96] The study found considerable mismatch in high-grade adverse events between the articles and the CDUS data, but it was not clear whether the mismatch was due to bias. It is important to note that published articles under-reported low-grade adverse effects: only 58% of low-grade adverse effects recorded in the CDUS database were reported in articles.[96]

Selective reporting of multiple alternative analyses

Bias may be introduced by selective reporting of multiple results generated by different analyses for a given outcome. Melander *et al.*[43] compared 42 trials of five selective serotonin reuptake inhibitor (SSRI) antidepressants submitted to the Swedish Drug Regulatory Authority for marketing approval and published articles. The study considered only one outcome, response rate. Both intention-to-treat (ITT) analysis and per protocol analysis were presented in 41 of the 42 submitted reports, but in only two stand-alone publications. The stand-alone publications tended to report the more favourable result by per protocol analysis.[43]

Compared with clinical trials, epidemiological studies may be more susceptible to selective reporting of results because of their exploratory nature. Kyzas *et al.* (2005)[97] conducted a meta-analysis to assess the association between a prognostic factor, the tumour suppressor protein 53 (TP53), and mortality outcome of patients with head and neck squamous cell cancer. They compared the results using (1) data from 18 studies that were indexed with 'survival' or 'mortality' in MEDLINE or EMBASE, (2) data from 13 published studies that were not indexed with 'survival' or 'mortality' in MEDLINE or EMBASE, and (3) data retrieved from authors for 11 studies in which data on mortality were collected but no usable data were reported. The pooled relative risk for mortality was 1.27 (95% CI: 1.06 to 1.53) using 18 published and indexed studies, 1.13 (95% CI: 0.81 to 1.59) using 13 published but not indexed studies, and 0.97 (95% CI: 0.72 to 1.29) using retrieved data.

According to the study by Kyzas *et al.*,[97] TP53 status can be measured by different methods. When available, Kyzas *et al.* used immunohistochemistry data, and a TP53-positive status was defined as nuclear staining in at least 10% of tumour cells or at least moderate staining in qualitative scales. They also standardised all-cause mortality to 24 months of follow-up. It was found that the association was stronger by using the definitions preferred by each publication (RR 1.38; 1.13–1.67) than when definitions were standardised (RR 1.27; 1.06–1.53).

Kavvoura *et al.* investigated the discrepancy between abstracts and full papers of epidemiological studies using 389 abstracts and 50 randomly selected full papers.[98] In the abstracts, 88% reported one or more statistically significant relative risks and only 43% reported one or more non-significant relative risks. The prevalence of significant results was less prominent in full texts of the articles. A median of nine (IQR 5–16) significant and six (IQR 3–16) non-significant relative risks were presented in the full text of the 50 articles. They also found that 'investigators selectively present contrasts between more extreme groups when relative risks are inherently lower'.[98]

Summary of evidence on outcome reporting bias

The most direct evidence is from the two studies that compared outcomes specified in trial protocols and outcomes reported in subsequent publications.[6,7] The results of unreported outcomes were obtained by a survey of original investigators. Due to low response rates and insufficient data for 2×2 tables, many included cases were excluded from the calculation of odds ratios. In the Danish cohort of 102 trials,[6] the odds ratio for reporting bias was based on 50 trials for efficacy outcomes and 18 trials for harm outcomes. Thirty trials for efficacy and only four trials for harm outcomes were used in the Canadian cohort of 48 trials.[7] This low response rate is likely to lead to an underestimation of outcome reporting bias.

Trials included in these two studies were mostly published before the CONSORT statement appeared in 2001, so it would be interesting to investigate whether outcome reporting bias has been reduced in trials published since 2001.

Findings from ongoing studies may provide further empirical evidence. For example, one study by Ghersi *et al.* compared 103 published RCTs and

their protocols, and was presented in 2006 as a conference abstract. It is now available only in abstract form with insufficient study detail.[91] We also identified an ongoing MRC-funded study[99] that aims to investigate the proportion and impact of within-study outcome reporting in an unselected cohort of 300 Cochrane systematic reviews.

Although case studies often yield evidence of limited usefulness, they may provide evidence indicating what further research is required. For example, the study by McCormack *et al.* of IPD meta-analysis indicated that some subjectively assessed outcomes may be more vulnerable to reporting bias than objectively assessed outcomes.[92]

Time lag bias

When the speed of publication depends on the direction and strength of the study results, this is referred to as time lag bias.[100] Empirical evidence on time lag bias could be separated into two categories: (1) the relationship between the study results and time to publication, and (2) changes in reported effect size over time.

Time to publication

The process of research is usually complex and involves several important milestones. These include development of the research proposal, approval by a research ethics committee, obtaining research funding, recruitment of participants, completion of follow-up, submission of manuscripts to a journal, and final publication in peer-reviewed journals. Measurement of elapsed 'time to publication' could be considered to start from any of these milestones, for example, from the date of REC approval, funding received, initiation or completion of enrolment, completion of follow-up, or manuscript submission.

Four cohort studies were analysed in the 2000 *HTA* report (see Appendix 10). Simes (1987) examined the time from trial closure to publication, using 38 published or unpublished trials on advanced ovarian cancer or multiple myeloma.[101] All six trials that showed a statistically significant survival difference were published within 5 years of study closure, while 5 of the 32 trials that showed no significant difference were published more than 5 years after study closure, and 7 of the 32 trials with non-significant results were not yet published.[101]

In a survey of 218 quantitative studies approved by a hospital Ethics Committee in Australia, Stern

and Simes observed that the median time from granting of ethical approval to the first publication in a peer-reviewed journal was 4.8 years for studies with significant results as compared with 8.0 years for studies with null results (HR 2.32; 95% CI: 1.47 to 3.66).[24] Adjustment for other factors that affect publication (e.g. research design and funding source) did not change this result materially. When only the large quantitative studies (sample size > 100) were analysed, the time lag bias remained evident (HR 2.00; 95% CI: 1.09 to 3.66).[24] Studies with non-significant trend ($0.05 < p < 0.10$) were published later compared with studies with null results ($p > 0.10$) (HR 0.43; 95% CI: 0.15 to 1.24). For qualitative studies ($n = 103$), there was no clear evidence of time lag bias involving studies with unimportant or negative results.[24]

Further empirical evidence on time lag bias came from a cohort of 66 completed phase 2 or phase 3 trials, conducted between 1986 and 1996 by a clinical trials group on AIDS.[23] The results were classified as 'positive' if an experimental therapy for AIDS was significantly ($p < 0.05$) better than the control therapy. 'Negative results' included those with no statistically significant difference and those in favour of the control therapy. Definition of publication was that the trial findings had to be published in a peer-reviewed journal. The median time from start of enrolment to publication was 4.3 years for positive trials as compared to 6.4 years for negative trials ($p < 0.001$). Positive trials were submitted for publication more rapidly after completion (median 1.0 year versus 1.6 years; $p = 0.001$) and were published more rapidly after submission (median 0.8 years versus 1.1 years; $p = 0.04$), compared with negative trials.[23]

Misakian and Bero identified 61 completed studies through a survey of 89 organisations that supported research on the health impact of passive smoking.[30] Time to publication was assessed from the start date of funding because it was difficult to decide the time of study completion. 'Published studies' were those that appeared in a peer-reviewed, non-peer-reviewed, or in-press publication, but not if published only as abstracts. The median time from funding start to publication was 5 years (95% CI: 4 to 7) for statistically non-significant studies, and 3 years (95% CI: 3 to 5) for statistically significant studies ($p = 0.004$). Multivariate analysis revealed that time to publication was associated with the statistical significance of the results ($p = 0.004$), experimental study design ($p = 0.01$), study size ($p = 0.01$) and animals as subjects ($p = 0.03$).[30]

The updated review identified five new empirical studies of the interval before manuscript submission to publication (Appendix 10),[27,37,102–104] and one study of time from manuscript submission to publication.[81]

Min and Dickersin found that statistically significant or important results were associated with time from completion of enrolment to full publication (HR = 1.75; 95% CI: 1.14 to 2.93), according to a study of 242 observational studies initiated at Johns Hopkins University.[104] However, Cronin and Sheldon examined a cohort of 70 studies sponsored by the UK NHS R&D programme and did not observe a significant difference in time from study completion to publication by whether a study result was significant ($p < 0.05$ or important) or not (HR = 0.53; 95% CI: 0.25 to 1.1).[27] Similarly, study results were not found to be significantly associated with time to publication, according to findings from the remaining three empirical studies (Appendix 10); however, in all of the studies where data were provided the trend suggested a shorter time for statistically significant or 'positive' studies than for non-significant or negative ones, and the cohorts of studies were often small.[37,102,103]

Dickersin *et al.* tracked 133 manuscripts of comparative studies accepted for publication by the *Journal of the American Medical Association* (*JAMA*) and investigated time from manuscript submission to publication.[82] Results were classified as positive if a statistically significant ($p < 0.05$) difference was reported for the primary outcomes. Seventy-eight (59%) manuscripts reported positive results, 51 (38%) reported negative results and the results of four (3%) articles were unclear. The median time interval between submission and publication was 7.8 months for positive studies versus 7.6 months for negative studies ($p = 0.44$).[82] Findings of this study indicated that time lag bias (if any) may likely occur before, not after, the manuscript submission to journals.

We also identified six studies that investigated the time from abstract presentation at meetings to subsequent full publication (Appendix 10).[49,60,61,69,70,105] The study by Krzyzanowska *et al.*[70] included 510 abstracts of large ($n > 200$) phase 3 trials presented at an oncology meeting between 1989 and 1998. They found that trials with statistically significant results were published earlier than those with non-significant results (median time to publication 2.2 versus 3.0 years; HR = 1.4; 95% CI: 1.1 to 1.7).[70] Findings from

two other studies[60,69] also suggested that time from abstract presentation to full publication was associated with significant results. However, the observed association between study results and time from abstract presentation to publication was not statistically significant in the remaining three studies (two of which were small in terms of the number of abstracts assessed).[49,61,105]

Change in reported effect size over time

The 2000 *HTA* report on publication bias[2] discussed only two brief reports on the temporal trend of reported effect size.[106,107] Rothwell and Robertson found that the treatment effect was overestimated by early trials as compared with subsequent trials in 20 of the 26 meta-analyses of clinical trials.[106] In another report, a significant correlation ($p < 0.10$) between the year of publication and the treatment effect was observed in 4 of the 30 meta-analyses published in *BMJ* or *JAMA* during 1992–6.[107]

The updated review included two new case studies of pharmaceutical interventions[108,109] and two studies in research on ecology or evolution[110,111] (Appendix 10). Gehr *et al.* found that reported effect size significantly decreased over time in three of the four meta-analyses of studies on pharmaceutical interventions.[108] In a case study of *N*-acetylcysteine in the prevention of contrast-induced nephropathy, Vaitkus and Brar found that trials published earlier reported more favourable results than trials published later.[109]

In a study of 44 meta-analyses in ecology, Jennions and Moller found a significant relationship between year of publication and estimated effect size, and the association remained significant even after controlling for sample size ($p < 0.01$).[110] However, in another case study in the area of ecology, Leimu and Koricheva did not find a significant association between the effect size and year of publication in two meta-analyses of studies testing plant defence theories.[111]

Ioannidis and Trikalinos hypothesised that highly contradictory results are more likely to be rapidly published than other results.[112] They investigated changes in between-study variance over time in 44 meta-analyses of epidemiological studies on genetic associations and in 37 meta-analyses of clinical trials of health-care interventions. It was found that early published studies tended to be more heterogeneous than later published studies

in meta-analyses of genetic associations. There was no significant change in the between-study variance over time in meta-analyses of clinical trials.[112]

Summary of evidence on time lag bias

Empirical evidence on time lag bias came mainly from studies that investigated time to publication using cohorts of studies. Because of the strict inclusion criteria, the Cochrane Methodology Review on time lag bias[113] included only two cohort studies.[23,24] We have adopted a more comprehensive approach and included more relevant studies (Appendix 10).

Four of the five newly identified cohort studies on time lag bias (before submission) did not find a significant association between time to publication and study results.[27,37,101,103] These four studies were relatively small, and the sources of the sample were diverse. It is unclear whether poor quality studies may tend to overestimate or underestimate the association between study results and time to publication. Of the six newly included studies of time from abstract presentation at meetings to full publication, three large studies reported significant time lag bias[60,69,70] while the other three studies (one large, two small) did not.[49,61,105]

Considering all the available evidence from cohort studies on time to publication, we conclude that on average studies with significant or important results still tend to be published earlier than studies with non-significant results. However, this conclusion may not be generalisable to many individual cases. Studies included in the identified cohort studies were usually divergent in terms of research questions and design, so that the observed association between study results and time to publication may be influenced by other confounding factors.[109] Limited evidence suggests that study findings that were difficult to interpret, for example, when the *p* value was greater than 0.05 but smaller than 0.10 or when results were mixed negative or positive, may take even longer to be published than studies that had clear negative results.

If it does exist, time lag bias is likely to occur before manuscript submission for journal publication.[81]

One consequence of time lag bias (earlier publication of significant results) may be a diminishing effect size reported by studies over time. Therefore, temporal trends of reported effect size in meta-analysis may indicate the existence of time lag bias (the 'fading of reported effectiveness' coined by Gehr *et al.*).[108] Compared with studies included in cohort studies of time to publication, studies in meta-analyses for the investigation of temporal trends of reported effect size were much more similar in term of participants, interventions and outcomes. However, time lag bias is only one of several possible explanations for changes in reported treatment effect over time.[108] It is surprising that only very limited research has been conducted to investigate temporal trends of reported effect size in meta-analysis.

Grey literature bias

The distinction between grey literature and unpublished or published studies may sometimes be ambiguous. Studies presented in the form of grey literature may be considered as published or as unpublished, according to different definitions.[2] The Third International Conference on Grey Literature defined grey literature as 'that which is produced on all levels of governmental, academic, business and industry in print and electronic formats, but which is not controlled by commercial publishers'.[114] Grey literature consists of an immense range, which includes brochures, pamphlets, internal reports, memoranda, databases on ongoing research, newsletters, conference proceedings and abstracts, technical reports, assignments and dissertations[115] as well as personal correspondence, web pages, data archives, policy documents and book chapters.

A survey published in 1993 found that 31% of published meta-analyses included unpublished data.[116] The proportion of people who supported the inclusion of unpublished data in meta-analysis at that time was 78% for meta-analysts or methodologists, while it was only 47% for journal editors.[116] Taus *et al.* reported in 1999 that 11% of the 814 references of included studies in 75 neurological reviews from the Cochrane Library were from books, theses or other unpublished sources.[117] Tetzlaff *et al.* presented a survey in 2006 that found that while both editors and review methodologists had become more in favour of including grey literature, editors were still less inclined towards the inclusion of grey literature (69%) than systematic reviewers or methodologists (85%).[118]

Empirical evidence on grey literature bias may be separated into two categories: (1) the subsequent

full publication of a cohort of grey literature (such as meeting abstracts) according to study results, and (2) a comparison of results of fully published studies and grey literature studies that aimed to answer the same research question. Evidence from the first category has been summarised in the cohort studies section. This section focuses on empirical studies that compared results from published and corresponding grey literature. Unpublished and grey literature studies were not separately considered, since it is usually impossible to distinguish the two.

Studies of multiple meta-analyses

The previous *HTA* report[2] included one study using multiple meta-analyses.[119] The updated review identified five new studies of multiple meta-analyses in which the results of published studies were compared with those estimated by using grey literature (see Appendix 11 for six included studies of at least 10 meta-analyses, and 16 individual case studies).[3,45,120–122]

McAuley *et al*. (2000)[119] investigated the impact of exclusion of grey literature from meta-analyses on the estimate of intervention effectiveness. In a sample of 135 meta-analyses, of which 41 meta-analyses (30%) included some form of grey literature (between 4.5% and 75% of included studies) the removal of grey literature resulted in an increase in the estimate of treatment effect of at least 10% in nine meta-analyses and a reduction of treatment effect of at least 10% in five. On average, published literature yielded significantly larger estimates of treatment effect, by 15% compared with grey literature [ratio of odds ratio (ROR) = 1.15; 95% CI: 1.04 to 1.28]. This study concluded that the exclusion of grey literature can lead to overestimation of treatment effects.[119]

The empirical study by Egger *et al*. (2003) included 60 meta-analyses in which results of published studies ($n = 630$) could be compared with those of grey literature trials ($n = 153$).[3] Estimated treatment effect based on grey literature ranged from 97% more to 209% less beneficial than those based on corresponding published trials. Pooled effect estimates from the grey literature were on average 7% greater than those from published trials (ROR 1.07; 95% CI: 0.98 to 1.15).[3] However, published trials tended to have a larger sample size and be of better quality than unpublished trials. Therefore, there is a possibility that bias could be introduced by including poor quality grey literature.[3]

A recent study by Turner *et al*. (2008) compared published and unpublished clinical trials of 12 antidepressant agents.[45] From the FDA database they identified 74 clinical trials of 12 antidepressants approved by the FDA between 1987 and 2004. Of the 74 trials, 23 were unpublished. It was found that the standardised effect size (Hedges' *g*) using data from the journal articles (0.41; 95% CI: 0.36 to 0.45) was on average 32% (ranged from 11% to 69%) greater than the effect size using data from the FDA reviews (0.31; 95% CI: 0.27 to 0.35) (sign test $p < 0.001$). In addition, negative or questionable findings (according to the FDA's decision) were less likely to be published, and if published, the result was often conveyed as a positive outcome.[45]

The remaining three studies of multiple meta-analyses in which the impact of inclusion of grey literature could be investigated[120–122] consistently found that published studies on average yielded a greater estimate of treatment effect compared with grey literature studies, although the difference was not statistically significant individually and the direction of bias was unpredictable for individual reviews.

Five of the six studies of multiple meta-analyses were included in the Cochrane methodology review on grey literature bias.[123] Using data from three studies[3,119,122] suitable for combining in meta-analysis, Hopewell *et al*. estimated that published trials on average suggested a 9% greater treatment effect than grey literature trials (pooled ROR for grey literature versus published trials = 1.09; 95% CI: 1.03 to 1.16).[123]

Case studies of grey literature bias

The previous *HTA* report[2] reviewed several case studies that investigated the impact of inclusion of grey literature. In the field of psychological and educational research, several case studies reported a tendency for the average effects reported in journal articles to be greater than the effects reported in the corresponding dissertations.[124–126] In the medical and health field, the previous *HTA* report included four case studies.[127–130] However, the validity of empirical evidence from case studies may be questionable because of selective reporting.[123]

According to findings from the 16 case studies included in this review (see Appendix 11 for details of the individual studies), published studies

tended to report a greater estimate of effect sizes than grey literature, although the difference was statistically significant in only some studies.[109,131–134] For example, in an IPD meta-analysis of paternal cell immunisation for recurrent miscarriage, Jeng *et al.* (1995) found that the estimated relative risk was 1.29 (95% CI: 1.03 to 1.60) using data from four published trials, and it was 1.01 (95% CI: 0.74 to 1.28) by using data from four unpublished studies.[131] A meta-analysis of animal experimental studies of nicotinamide for stroke found that abstracts reported a statistically significantly lower estimate of effect size than fully published studies ($p < 0.001$).[132]

One case study by MacLean *et al.* was at first (in 1999) presented as a meeting abstract[130] and then fully published in 2003.[135] The study compared data from published studies with data from FDA New Drug Application Reviews for assessing non-steroidal anti-inflammatory drug (NSAID)-associated dyspepsia. The quality of unpublished data from FDA reviews was comparable with that of published data. The pooled relative risk for NSAID-induced dyspepsia was 1.07 (95% CI: 0.70 to 1.63) using the FDA data, and 1.21 (95% CI: 0.81 to 1.81) using data from published trials. Meta-regression analyses found that estimates varied significantly by NSAID dose ($p = 0.037$) but were not related to whether the study was published or not ($p = 0.73$).[135] The reported difference between the published and grey literature appeared greater in the abstract[130] than that in the full publication.[135]

It is interesting to compare two case studies on the same topic.[136,137] Whittington *et al.* (2004) compared the results of published and unpublished data from clinical trials of SSRIs in childhood depression.[136] They found that the results of published trials indicated a favourable risk-benefit profile for some SSRIs, while unpublished data tended to be unfavourable. They concluded that 'non-publication of trials, for whatever reason, or the omission of important data from published trials, can lead to erroneous recommendations for treatment'.[136] In another case study by Wallace *et al.* (2006), a cumulative meta-analytic approach was used to synthesise evidence from trials of SSRIs in paediatric depression.[137] Although the unpublished data tended to suggest that the SSRIs were less efficacious and more harmful, the overall interpretation of evidence on efficacy and safety would not change on inclusion of unpublished trials.[137]

Batt *et al.* (2004) compared the quality, quantity and nature of grey and published evidence on costs and cost-effectiveness of strategies to increase coverage of routine immunisations in low and middle income countries.[138] Of 34 included studies on effectiveness from the grey literature, 63% met the quality criteria set for inclusion, while 57% of published literature met these criteria, suggesting that in this area grey literature is of higher quality. Inclusion of grey literature almost doubled the number of included studies, covered different geographical areas, covered operational research and finance (rather than the economics and policy-making covered in published literature) and was more up to date. There were no statistically significant differences between published and grey literature in terms of effectiveness (final coverage or changes in coverage).[138]

Summary of evidence on grey literature bias

There is good evidence that published literature tends to be more positive about the effectiveness of interventions than corresponding grey literature, although this can vary in individual reviews. The quality of grey literature studies can be higher, lower or the same as the corresponding published studies.

Studies of cohorts of meeting abstracts found that a large number of abstracts presented at conference meetings will not be published in full, and the subsequent publication of abstracts is associated with study results (see Cohorts of meeting abstracts and Pooled analyses of cohort studies). Such studies included abstracts of studies on diverse research questions, and it is difficult to exclude the influence of many confounding factors on the association between study results and subsequent full publication. Therefore, findings from such cohort studies provided only indirect evidence on grey literature bias.

More direct evidence on grey literature bias came from studies that compared the result of published studies and grey literature within a meta-analysis. Many case studies were identified, but the interpretation of findings from these case studies was complicated due to concern over possible selective reporting. Therefore, empirical studies that used an unbiased sample of multiple meta-analyses provided the most valid evidence on grey literature bias. The updated review identified several recent studies of multiple meta-analyses

in which the results of published studies could be compared with that of grey literature.

Available evidence from good quality studies suggested that published studies tend to report a greater treatment effect compared with a more complete set of data from published and unpublished studies, but that for individual reviews the effects may not always be in this direction. Grey literature studies may be relatively small and of relatively poor quality, although again this is not always the case. The impact of grey literature in meta-analysis is usually small, although occasionally data from grey literature may have important clinical implications. A case-by-case approach is required to decide whether grey literature should be comprehensively searched and included in systematic reviews. The inclusion of grey literature may sometimes introduce bias, as will exclusion of grey literature in other cases.[3,139]

The most commonly included unpublished data used in reviews are conference abstracts, but there are difficulties in using data from abstracts as they provide limited information, may be on partial datasets and may be misleading when compared with later full publications. Evidence on the importance and utility of other types of unpublished material is less clear.

Language bias

Many prestigious international scientific journals are published in English, and journals published in English are more likely to have greater journal impact factors (JIF).[140] However, writing for journals published in English can be more difficult for researchers who are non-native English speakers.[141–143]

Quality of studies published in English and non-English languages

When fictitious manuscripts with identical methodological flaws were sent to referees, Nylenna *et al*. (1994) found that Scandinavian referees awarded higher quality scores to English-language manuscripts than to the manuscripts in a referee's own national language.[144]

Moher *et al*. (1996) compared completeness of reporting of 133 trials published in English and 96 trials published in French, German, Italian or Spanish.[145] They found no significant difference

between English and non-English trials in the completeness of reporting or overall quality score (51.0% versus 46.2%). It was therefore concluded that all trial reports should be included in systematic reviews irrespective of the language in which they are published.[145]

Junker (1998)[146] identified deficiencies in the quality of reporting of 32 German and 89 English-language reports of placebo-controlled trials published by the same group of authors. The mean quality score was 8.4 (on a scale of 0 to 18), with a non-significant difference in the mean quality score between English and German-language reports (0.27; 95% CI: −0.97 to 1.52). However, Junker's assessment is somewhat limited because the investigators looked only at published papers involving German-speaking authors from a single research group.[146]

More recently, Moher *et al*. (2003) found that there were only minor differences in the quality of reports between RCTs published in English and in non-English languages in a study of 42 meta-analyses.[4] However, Egger *et al*. (2003) observed that on average, 115 non-English-language trials tended to include fewer participants, were more likely to show statistically significant results, and were of lower methodological quality, than 485 other trials published in English.[3]

Therefore, studies published in languages other than English cannot be generally excluded for the reason of study quality.

The previous *HTA* report on publication bias[2] included a study of multiple meta-analyses[147] and a study that compared 40 pairs of RCTs published in German and in English.[148] Since then, two major *HTA*-supported studies have been completed and these provide more evidence on the differences in estimated treatment effects between English and non-English language trials in meta-analysis (see Appendix 12 for details of the six studies included in this section).[3,4]

Comparison of studies published in different languages

Egger *et al*. (1997)[148] identified 40 pairs of RCTs, each pair comprising an RCT published in German, and a matched RCT by the same author published in English during the same period. The investigators found that design characteristics and quality features were similar between RCTs published in German, and RCTs conducted in

German-speaking Europe that were published in English. Statistically significant results ($p < 0.05$) were reported in 35% of German language articles and 62% of English language articles (OR 3.75; 95% CI: 1.25 to 11.3). Logistic regression analysis found that a statistically significant finding was the only variable that was associated with a trial's publication in English-language journals. It was therefore concluded that 'authors are more likely to publish RCTs in an English language journal if the results were statistically significant'.[149]

A similar study by Heres *et al.* (2004) reported similar findings in a comparison of 21 pairs of trials in the field of neuroscience matched by the key authors.[150] In this instance, significant results were reported in 33% of German-language articles as compared with 57% of the English-language articles (Wilcoxon's test p = 0.14).[150]

Studies of multiple meta-analyses

Direct evidence on the impact of language bias comes from evaluations of multiple meta-analyses where the results of studies on the same research question but published in different languages could be compared (Appendix 12).

Gregoire *et al.* (1995) studied meta-analyses published in eight medical journals between January 1991 and April 1993.[147] They found that 28 of the 36 meta-analyses had language restrictions. By repeating the same searches without language restrictions in these 28 meta-analyses, they identified 19 individual studies that had not been included for language reasons. The inclusion of eight of these 19 studies to the five corresponding meta-analyses did not change the findings. However, inclusion of the other 11 studies to the remaining seven corresponding meta-analyses had the potential to modify the results. The most important difference was the change in the 95% CI of the overall OR estimated in a meta-analysis of selective decontamination of the digestive tract in intensive care units. The pooled OR was 0.70 (95% CI: 0.45 to 1.09) in the original meta-analysis, and this became 0.67 (95% CI: 0.47 to 0.95) after including a study published in a Swiss journal.[147]

Moher *et al.* (2000) examined a set of 19 meta-analyses to investigate whether different estimates of treatment effect were obtained in meta-analyses restricted to English-language studies compared with those without this restriction. Language-restricted meta-analyses, compared with meta-analyses involving non-English language studies, did not differ with respect to the overall estimate of effectiveness (ROR 0.96; 95% CI: 0.78 to 1.18).[151] Meta-analyses without language restrictions had narrower confidence intervals (average width 0.79; 95% CI: 0.51 to 1.07) compared with language-restricted meta-analyses (average width 0.92; 95% CI: 0.53 to 1.32), which represents a statistically significant relative difference in precision of 16%. These findings were limited by small sample size, small sampling frame, limited clinical topics and limited interventions. The meta-analyses that did include non-English-language trials had a very low number of such studies. Moreover, the majority (13/19) of the meta-analyses included only one trial published in languages other than English.[151]

A further study using 42 meta-analyses (including 529 English- and 133 non-English-language trials) was conducted by Moher and his colleagues (2003).[4,152] The 42 meta-analyses included 34 meta-analyses of conventional interventions, and eight meta-analyses of complementary and alternative medicine. The exclusion of trials in languages other than English, compared with their inclusion, did not yield a significantly different estimate of treatment effect overall (ROR 1.11; 95% CI: 0.92 to 1.34), or when the meta-analyses looked only at conventional interventions (ROR 1.02; 95% CI: 0.83 to 1.26). However, in meta-analyses of complementary medicine, exclusion of non-English trials resulted in a 63% smaller protective effect (ROR 1.63; 95% CI: 1.03 to 2.60). The authors concluded that language bias is unlikely to be a problem for many meta-analyses in the field of conventional medicine, but it may substantially alter the results of meta-analyses of complementary medicine.[4]

Egger *et al.* (2003) provided further empirical evidence by independently examining the influence of non-English-language trials in a sample of meta-analyses.[3,153] They identified 50 meta-analyses that included a total of 485 English-language trials and 115 non-English-language trials. Within these meta-analyses, treatment effect estimates were on average 16% more beneficial in non-English-language trials (ROR 0.84; 95% CI: 0.74 to 0.97) but with considerable heterogeneity. Excluding non-English-language studies led to a variety of changes, from a reduction in benefit of 42% to an increase of 23%. The exclusion of non-English-language studies resulted in greater benefit of the intervention in five meta-analyses, reduction in benefit in 16 meta-analyses, with little or no effect (<5%) in 29 meta-analyses. The average precision of treatment effect estimates decreased

from 8.34 to 7.68 after exclusion of non-English language trials.[3]

Summary of evidence on language bias

The impact of excluding non-English-language studies in systematic reviews appears to be highly heterogeneous. Different types of non-English-language studies (involving different areas of health care, and from different countries) may either be more or less likely to show statistically significant effects than comparable English-language studies, and they may be of lower or similar methodological quality. However, a common finding was that exclusion of non-English-language studies reduced the precision of the estimate of effect.

While there are specific areas where omitting non-English-language studies appears to result in a very high risk of bias (studies in the area of complementary medicine, for example) their exclusion in other areas may, or may not, result in bias. If exclusion does result in bias it is impossible to assess beforehand which direction this bias may take, as it may inflate or deflate the apparent effect size. This will be difficult to assess unless non-English-language studies are first included and later excluded. The best way to ensure that a review does not contain language bias is to search for and include relevant non-English language studies. The cost-effectiveness of this strategy (given the additional searching and translation time and costs) is unclear.

Citation bias

In published articles, references to other studies are cited for various reasons, for example, to show the importance of a research question, to borrow methods and techniques, or to give positive credit to the material referenced.[154] The chance of a study being cited by others may be associated with many factors like the journal impact factor, nationality of authors, working partnerships, etc. Citation bias occurs when the probability that a study will be cited is associated with the study result.

The previous *HTA* report[2] in 2000 included several studies that provided empirical evidence on citation bias.[155–160] Five recently published empirical studies on citation bias were identified in this updated review (see Appendix 13 for the included studies).[161–165] The previously cited studies will be discussed first, followed by the newer studies.

Shadish *et al.* (1995) randomly selected one citation from each of 283 articles published in three psychological journals and asked each author about the most important reason for citing the selected references.[155] It was found that citation was most commonly used to support the author's argument, while study quality was not considered in most cases.[155]

In one study examining the judgement and decision literature, it was found that poor-performance results were significantly more likely to be cited than positive good-performance results.[156] This could not be explained by the journal's popularity or the year of publication. This did suggest citation bias but the results were questionable since the poor-performance and good-performance articles were published in different journals and reported different evidence.[166]

Gotzsche (1987) examined the existence of citation bias by using 111 comparative trials on non-steroidal anti-inflammatory drugs.[157] The trial result was defined as positive if the benefit:harm ratio was in favour of the experimental drug. The pattern of citation was then classified as positive, neutral or negative selection of references by comparing the proportion of references reporting positive and negative results. For example, selection was classified as positive when the proportion of trials with a positive outcome in the reference list was higher than that in all available trials. Among the 76 trials in which citation bias was probable, the selection of references was classified as neutral in 10, negative in 22, and positive in 44. In conclusion, positive selection of references is more likely to happen than neutral and negative selection, suggesting citation bias.[157]

Ravnskov (1995) examined citations in three authoritative reviews on diet–heart issues and found that only one of six relevant RCTs with a negative outcome was cited and by only one of the three reviews. However, two, four and six non-randomised trials with a positive outcome were cited in each review respectively, suggesting that 'fundamental parts of the diet-heart idea are based on biased quotations'.[158]

Hutchison *et al.* (1995) assessed citation bias by comparing the proportion of relevant supportive and non-supportive trials used in 17 reviews on the clinical effectiveness of pneumococcal vaccine. Supportive trials were defined as those that reported significantly fewer failures in vaccinated

subjects than among the controls. It was found that unsupportive trials were more likely to be cited than supportive trials (11.9% versus 5.8%). The tendency to cite recent trials may be one reason for this disproportionate citation of unsupportive studies because six of the seven trials published after 1980 were unsupportive and all seven trials published before 1980 were supportive.[159]

In an assessment by Song *et al.* (1997) of published narrative reviews on the prophylactic removal of impacted third molars, it was found that reviews with similar aims included very different evidence on which to draw conclusions.[160] Of 69 studies that were discussed in nine general reviews about the association between pathology and impacted third molars, one was quoted in five reviews while 43 were cited only once. This discrepancy in the use of relevant studies cannot be reasonably explained by the year of publication or quality criteria. This selective citation of studies corresponded with conflicting conclusions from these narrative reviews.[160]

Five more recent studies have been added to this update. Chapman *et al.* (2009) examined association between citation frequency and reported prevalence in studies of smoking among schizophrenia patients.[162] They found that a 10% increase in reported prevalence of smoking was associated with a 61% (95% CI: 30% to 98%) increase in citation rate.[162] Another study by Callaham *et al.* (2002) evaluated how 204 emergency medicine studies presented to a meeting in 1991 were cited, and the factors associated with citation.[161] Predictors for citation frequency were the impact factor of the journal, the presence of a control group, newsworthiness score and sample size, while biased citation of positive outcomes was not observed.[161]

Kjaergard and Gluud (2002) reviewed 530 hepato-biliary disease trials to assess whether trials with statistically significant outcomes were cited more often than those with non-significant results.[163] They found a significant positive association between a statistically significant study outcome and citation frequency. The citation frequency was also associated with disease area and adequate generation of allocation sequence.[163]

In another study of 368 research papers published in four psychiatric journals, Nieminen *et al.* (2007) found that citation rate was related to *p*-value.[164] Median number of citations for papers reporting 'significant' and 'non-significant' results was 33 versus 16 respectively. Compared with studies with non-significant results, the ratio of citation rate for studies with significant results was 1.63 (95% CI: 1.32 to 2.02).[164]

Schmidt and Gotzsche (2005) investigated reference bias in 42 narrative reviews of physical interventions on house dust mite antigens.[165] Reference selection in each review was classified as positive, neutral or negative according to whether the proportion of trials with a statistically significant outcome in the review was higher than that among all trials available. For example, positive selection of references meant that the proportion of studies with positive results cited in a review was higher than the proportion of positive trials in all relevant trials available. Of the 38 reviews in which physical interventions were recommended, 10 reviews were neutral in terms of reference selection, 27 reviews had a positive selection of references and one a negative selection. The four reviews that did not recommend physical interventions all had a negative selection of references.[165]

Summary of evidence on citation bias

Empirical evidence indicates that studies with positive or significant results are on average associated with a higher frequency of citation, although this may not always be the case in specific areas of the literature. Non-systematic narrative reviews are a specific area where biased citation of research findings can result in misleading conclusions.

Duplicate (multiple) publication

Duplicate, redundant, repetitive or multiple publications are defined as submission of similar manuscripts to more than one journal or the republication of the same data in two or more journals.[167] It has been estimated that 10–25% of the published literature in biomedical sciences represents duplicate or redundant publications.[168] The publications may overlap partially or completely, representing a similar portion or major component of a study, and may share the same hypotheses, methods, results and/or discussion.

Multiple publications of the same data in different journals has been condemned mainly for wasting journal space and editors' and referees' time, as well as readers' time.[167,169–171] However, publication of the same data in different ways may help to

disseminate important research results, providing any previous or parallel publications have been explicitly referenced. However, researchers and journal editors may have different understanding about duplicate publication and it is sometimes difficult to distinguish the unacceptable redundant publication from the acceptable 'parallel' publication.[172] Recently, a database of duplicate publication and potential plagiarism (Déjà vu) has been developed by using a text similarity algorithm to identify extremely similar references from MEDLINE.[173]

Duplicate publication can be classified as 'overt' or 'covert'.[174] Overt duplicate publication is defined as reanalysis of data from a study with appropriate cross-referencing of original reports. Covert duplicate publication is when the same data are published in different places or at different times without adequate reference to a previous or parallel publication. Bias may be introduced in systematic reviews by including data from the same study more than once because of covert duplicate publication.

Empirical evidence on duplicate publication bias

The previous *HTA* report on publication bias included several case studies that provided empirical evidence on duplicate publication bias.[22,174–176] No new published studies and only one conference abstract were identified in this updated review.

Gotzsche (1989) examined 44 multiple publications of 31 controlled trials of NSAIDs in rheumatoid arthritis and found important reported differences in design, exclusion of protocol violators, number of effect variables, number of side effects, and the significance levels between duplicated publications of the same studies.[175] The conclusion became more positive for the new drugs in the late publications of three trials. He also suggested that multiple publications were difficult to detect because the first author and the number of authors cited often differ.[175]

Tramer *et al.* (1997) assessed the impact in a meta-analysis of duplicate data on efficacy estimates of ondansetron on postoperative emesis. It was found that, for three trials that were published in six reports, there was no cross-referencing.[174] The estimated number-needed-to-treat (NNT) to prevent one vomit within 24 hours was 9.5 (95% CI: 6.9 to 15) in 16 non-duplicated reports and

3.9 (95% CI: 3.3 to 4.8) in the three reports that were duplicated. The efficacy was overestimated by including duplicated data (NNT = 4.9; 95% CI: 4.4 to 5.6) compared with the report without duplicated data (NNT = 6.4; 95% CI: 5.3 to 7.9). Tramer *et al.* also discussed difficulties in identifying duplicated publications of the same trial data. For example, the same trial might report a different number of patients or different patient characteristics, or use completely different authors in separate publications.[174]

Huston and Moher (1996) found that identifying the data from single centres of multicentre trials of risperidone for schizophrenia was far from simple because of the chronology of publications, changing authorship, lack of transparency in reporting, and frequent citation of abstracts and unpublished reports.[176] For example, a North American trial had been reported in part, transparently, and not so transparently, in six different publications by using different author names. It had also been cited in several unpublished forms.[176]

Easterbrook *et al.* (1991) conducted a survey of studies approved by an REC and found that studies with significant results were more likely to generate multiple publications and more likely to be published in journals with a high citation impact factor when compared with those with non-significant results.[22] Vandekerckhove *et al.* (1993) identified a review of RCTs of infertility treatment and found that 'six studies with a significant result (but none with a non-significant result) were reported in four publications from the same institution'.[177]

The updating identified only an abstract by Martin *et al.* (2004) in which they examined the impact of including duplicate publications in a meta-analysis of off-pump versus on-pump coronary artery bypass surgery.[178] Trials were classified as covert duplicates when there was no citation of the original publication and non-covert duplicates when the publication declared the duplication or cited the original publication. The authors found that a total of 15 (34%) of the 44 trials were duplicate publications. Of the 15 duplicate published trials, 10 were covert and five were non-covert publications. However, there was no significant difference in the estimate of mortality when duplicates were included (OR 0.85; 95% CI: 0.46 to 1.57) or not included (OR 0.86; 95% CI: 0.48 to 1.54).[178]

Summary of evidence on duplicate bias

We identified only very limited empirical evidence from case studies about the existence of duplicate publication bias. However, it is clear that covert duplicate publication of data from the same study may introduce bias in systematic reviews as the weights carried by particular studies are magnified.

Place of publication bias

Ben-Shlomo and Davey-Smith (1994) found that the *BMJ* published more research articles supporting the 'early life hypothesis' (about the impact of early life development on the risk of adult disease) than the *The Lancet*.[179] They suggested that there may be 'place of publication' bias because, for reasons of editorial policy or readers' preference, one journal is more enthusiastic towards publishing articles about a given hypothesis than other journals.[179]

In a study that compared published and registered trials in advanced ovarian cancer, Simes (1986) found that trials with significant results ($p < 0.05$) in favour of the treatment tended to be published in prominent journals (such as the *New England Journal of Medicine* and *Cancer*), while trials with non-significant results tended to be published in less widely circulated journals.[127] Bero *et al.* (1994) compared 297 symposium articles in journal supplements and a sample of 100 journal articles on environmental tobacco smoking published between 1995 and 1993, and found that 'symposium articles were more likely to agree with the tobacco industry's position (46% vs. 20%)'.[180]

This updated review includes a study by Penel and Adenis[181] that examined the association between the results of 74 phase II trials investigating anticancer targeted therapies and the impact factors of journals publishing these trials. Positive trials were defined as those with an objective response rate equal or superior to the prespecified efficacy threshold, and negative trials were those with a response rate lower than expected. It was found that positive results were more likely to be published in journals with high impact factors, compared with negative results ($p = 0.004$; median 6.14 versus 2.71).[181]

We also identified a new study on location bias that examined the results of clinical trials of complementary and alternative medicine (CAM) therapies published in mainstream medical journals or in complementary medicine journals.

Pittler *et al.* (2000) identified 19 systematic reviews that included 351 controlled trials of complementary medicine.[182] Mainstream medical journals with a high impact factor tended to publish a relatively low proportion of trials with significant results compared with complementary medicine journals (50% versus 63%). They suspected that 'this may reflect the reluctance of authors to submit positive trial reports to these "flagship" orthodox journals, perceiving them to be hostile to CAM'.[182]

Country bias

The causes of variable results from studies on the same topic between different countries are complex, and selective publication is only one possible explanation. The variable results between different countries were studied by Ottenbacher and DiFabio (1985).[183] They observed that the estimated efficacy of spinal manipulation therapy was greater in studies reported in English-language journals published outside the USA than for similar studies in journals published in the USA (average effect size 0.45 versus 0.29). It was suggested that this finding might be explained by the existence of publication bias and/or other intervention characteristics.[183]

A study by Vickers *et al.* (1998) examined 666 abstracts from MEDLINE of clinical trials published up to 1995.[184] The proportion of positive results (when the test treatment was superior to control) in trials comparing acupuncture with controls was 100% for 50 trials originating from China, Taiwan, Japan and Hong Kong.[184] Conversely the results were 56.7% for 180 trials originating from 14 western countries such as the USA, UK, Sweden, Denmark, Germany and Canada. The study also identified that the percentage of positive results in trials of interventions other than acupuncture was 99% for trials originating from China, 97% from the USSR/Russia, 95% from Taiwan, 89% from Japan and 75% from England. It was concluded that publication bias was a possible explanation for the unusually high proportions of positive results reported from some countries.[184] Tang *et al.* (1999) confirmed the existence of publication bias in Chinese journals of traditional medicine by presenting an asymmetric funnel plot of 49 trials of acupuncture in the treatment of stroke.[185]

Continuing the theme of more positive results appearing in published work from specific countries, Pan *et al.* (2005) explored country

bias in the area of genetic epidemiology.[186] They worked with 13 gene-disease associations with existing meta-analyses of at least 15 non-Chinese studies and searched for relevant Chinese studies. Of the 161 studies found (augmenting the 301 non-Chinese studies already included in the meta-analyses) only 20 were included in MEDLINE. Despite having smaller sample sizes than the non-Chinese studies, significantly more Chinese studies showed statistically significant associations (48% versus 18%) and the largest effects were seen in the small sample of MEDLINE indexed Chinese studies. This reinforces the finding that there are large bodies of literature that are commonly missed from meta-analyses using only MEDLINE, but that such bodies may display high levels of publication bias so that caution is needed in interpreting the results of such groups.

Lack of publication by authors from developing countries may lead to 'country bias', both under-representing the research questions of such areas and causing an important gap in our ability to locate and synthesise the results of the whole body of conducted research. For example, King (2004) found that 31 countries accounted for 98% of the world's highly cited papers, the remaining 192 countries accounting for less than 2%.[187] If the results of such studies are different from the results of similar studies by researchers from developed nations then we will observe publication bias.

Database indexing bias

Database indexing bias occurs when there is biased indexing of published studies in literature databases.[188] A literature database, such as MEDLINE or EMBASE, may not include and index all published studies on a topic.[189–191] The literature search will be biased when it is based on a database in which the results of indexed studies are systematically different from those of non-indexed studies. This bias is likely because the result of a study may determine whether and where the study is published.

This updated review identified no new studies on database indexing bias. The following two studies were included in the previous *HTA* report. A study by Zielinski in 1995 estimated that about 98% of journals indexed in the major literature databases were from western developed countries.[192] Nieminen and Isohanni (1999) suggested that there was a bias against European journals in medical literature databases because 27% of psychiatric research papers by Finnish

authors published in English were not indexed in MEDLINE.[193]

Media attention bias

The general population gets most of its information about the latest developments in science and medicine from the popular media. How the press presents the findings of these developments has a very powerful influence on public perception. Media attention bias occurs when studies with striking results are more likely to be covered by newspapers, radio and television news. The overly optimistic portrayal of the scientific findings to the public affects the public participation in policy discussions and creates unrealistic expectation of the potential benefits of a new scientific development.[194] It was not clear whether media coverage was influenced by people's opinions about what is important, or whether people's judgements were influenced by the media coverage, although both directions of influence are possible.[195]

The 2000 *HTA* report on publication bias[2] included limited evidence on media attention bias. Combs and Slovic in 1979 found that the coverage by two newspapers in the USA about causes of death was not related to the statistical frequency of their occurrence.[195] The newspaper overemphasised homicides, accidents and disasters, and under-reported diseases as causes of death. Violent accidents and homicides make more interesting and exciting stories than diseases.[195]

Houn *et al.* (1995) examined the popular press coverage of research in the USA in 1985 and in 1992 on the association between alcohol and breast cancer.[196] They identified 58 scientific articles and 89 newspaper or magazine stories. Only 11 of these 58 scientific articles were cited in the newspaper or magazine stories. Press stories cited all scientific articles that were published in *JAMA* and the *NEJM* but articles published in other journals were often ignored by the newspaper and magazine reports. There was no significant difference between the scientific articles and press stories in the frequency of reporting positive, negative or neutral results. It was concluded that 'the vast majority of scientific studies on alcohol and breast cancer were ignored in press reports'.[196]

Koren and Klein (1991)[197] compared newspaper coverage in the USA of one positive study that reported a significant association between radiation exposure and cancer risk[198] and one negative study

that did not,[199] published in the same issue of *JAMA* in 1991. Nine of the 19 newspaper reports covered only the positive study. In the other 10 reports that covered both the positive and the negative studies, the average number of words was 354 for the positive result and 192 for the negative result. It was suggested that the number, length and quality of newspaper reports on the positive study were greater than news reports on the negative study, which suggests a bias against news reports of studies that show no effects or no adverse effects.[197]

This updated review identified two studies that examined the media coverage of abstracts presented at scientific meetings. Schwartz *et al.* (2002) examined 252 news stories about 147 research articles presented at scientific meetings in 1998, and found that the 43 abstracts that received prominent news coverage were no more likely to be formally published.[200] Woloshin and Schwartz (2006) found that the media coverage of scientific meetings in major international outlets in 2003 often failed to report basic study facts, so that the public would be likely to be misled about the validity and relevance of the science presented, especially as there were no published findings to refer back to for confirmation.[201]

Whiteman *et al.* (2001) examined the scientific publications that do and do not support an association between hormone replacement therapy (HRT) and breast cancer to assess whether they were cited in the popular media in similar proportions.[202] A total of 32 scientific publications were identified, 20 (63%) of which had positive conclusions in which the results supported the HRT–breast cancer association, and 12 (38%) did not. Of the 203 citations in the media reports, 82% were of positive studies and 18% were of negative studies, representing a significant excess of citations of positive publications ($p < 0.01$).[202]

The reporting of clinical trials of herbal remedies by the popular media may be influenced by the disclosure of funding information and competing interest in the scientific and medical literature. Koper *et al.* (2006)[203] used a coding frame analysis technique to systematically compare newspaper articles with the reporting of the same trials in the medical literature. The analysis of 389 newspaper articles from the UK, USA and Canada indicated that media coverage of conflicts of interest had an effect on the overall tone of the article.[203]

Limitations of the available evidence

Empirical studies on publication and related biases have focused mainly on certain areas of research such as clinical trials of health-care interventions. There is only very limited evidence on publication bias in many other research fields including basic research and observational studies.

Studies of publication and related biases themselves may be as vulnerable as other studies to the selective publication and reporting of significant or striking findings.[1] Much of the empirical evidence comes from case reports that may be selectively reported because of their striking findings.

Studies that are less selective are able to provide more convincing evidence on publication and related biases, including cohorts of research protocols, submitted or registered studies. However, many empirical studies were based on cohorts of studies that were diverse in terms of design and research questions. It is usually impossible to exclude the impact of confounding factors on the observed association between study results and publication status. There is very limited and conflicting evidence on factors that may be associated with the direction and extent of publication and related biases.

Findings from individual empirical studies were often heterogeneous, and pooled estimates of publication and related biases can indicate some average trends but may not be generalisable to many individual cases. A case-by-case approach is required to gauge the possible impact of publication and related biases and to decide appropriate measures to deal with these biases.

Conclusions

The 2000 *HTA* report included very limited evidence on outcome reporting bias. Recently published studies have provided convincing evidence that outcome reporting bias exists and is likely to have important effects on pooled summary data within systematic reviews. Limited evidence indicates that harm and subjectively assessed outcomes may be more vulnerable to biased selective reporting than efficacy and objectively assessed outcomes.

Studies with significant or positive results tend, on average, to be published earlier than studies with non-significant or negative results. However, new evidence is less clear about time lag bias than was suggested in the previous review. One consequence of time lag bias would be a diminishing effect size reported by studies over time, although very limited research has been conducted to investigate temporal trends of reported effect size in meta-analysis.

The updated review identified substantially new evidence on grey literature bias. Evidence suggests that published studies tend to report a greater treatment effect than those of grey literature or unpublished studies. However, for individual cases, the direction of bias is unpredictable, and grey literature studies may be relatively small and of poor quality, although this is not always the case. In some reviews inclusion of data from grey literature or unpublished studies have important clinical implications.

Substantially new evidence on language bias has been identified. The impact of excluding non-English-language studies from systematic reviews was highly heterogeneous. Exclusion of non-English-language studies from systematic reviews may be associated with greater, similar or smaller estimates of treatment effects. However, exclusion of non-English-language studies appears to result in a particularly high risk of bias in some areas of research such as complementary and alternative medicine.

Empirical evidence indicates that studies with significant or positive results are on average associated with a higher frequency of citation. Non-systematic narrative reviews are a specific area where biased citation of research findings can result in misleading conclusions.

The updated review identified very limited new evidence on duplicate publication bias, although it is clear that covert duplicate publication of data may introduce bias in systematic reviews. Available evidence on the existence of place of publication bias, database or index bias, country bias and media attention bias is still very limited. The impact of these biases could be prevented in well-conducted systematic reviews.[204]

There is limited evidence on place of publication bias, database bias, country bias and media attention bias. It is helpful to be aware of the potential existence of these biases.

Chapter 5
Consequences of dissemination bias

Evidence from empirical studies reviewed in Chapter 3 suggests that the dissemination profile of research findings may be associated with the strength or direction of study results. As a direct consequence of publication and related biases, published studies may provide misleading estimates of treatment effects or associations between variables. The previous 2000 *HTA* report identified very little direct evidence on the impact of publication and related biases on health policy, clinical decision-making and the outcome of patient management.[2] In this updated review, we considered consequences of publication bias according to types of studies, classifying them into three categories: basic research, observational studies and clinical trials.

Basic research studies

Many new treatments are initially investigated in basic laboratory and animal research. Based on findings from basic research, clinical trials may be conducted to test an intervention in humans. However, subsequent clinical trials often fail to provide confirmatory positive results, inconsistent with findings from basic animal research.[205] One of several possible explanations for the observed discrepancies in results between basic research and clinical trials is biased publication of positive results of basic studies.[206] If positive results from basic research are more likely to be published than negative results, results of published studies of basic research will represent an overestimation of potential treatment effects. It is unlikely that clinical trials that are designed based on false-positive findings from basic research will provide a positive result.

Empirical evidence on the existence and impact of publication bias is very limited in the field of basic laboratory and animal research. This updated review included a case study of a neuroprotective drug, nicotinamide, for focal cerebral ischaemia in animal experimental studies.[132] The animal experimental studies suggested potential efficacy of neuroprotective drugs, but clinical trials failed to confirm these drugs' efficacy. Macleod *et al.* (2004) conducted a systematic review of animal experimental studies of nicotinamide. They found that animal studies that were fully published showed a greater effect (effect size 0.306; 95% CI: 0.241 to 0.371) than studies that were presented in abstract form (0.162; 95% CI: 0.066 to 0.258). It was suspected that some studies with negative results may not be available even in abstract form.[132]

Observational studies

A large number of epidemiological studies have been conducted to investigate various risk factors associated with diseases.[207] However, there are contradictory findings from epidemiological studies regarding many risk factors.[208] For example, the results of epidemiological studies were contradictory regarding the risk of hair dyes, coffee, oat bran, oral contraceptives, environmental exposure to residential radon, and the presence of DDT metabolites in the bloodstream.[209]

Ioannidis and Trikalinos found that early published studies of genetic associations tended to be extremely contradictory, and hypothesised that 'highly contradictory results are most tantalizing and attractive to investigators and editors'.[112] Two further studies (all by Ioannidis and his colleagues) found considerable outcome reporting bias in studies of cancer prognostic factors,[97] and in studies of epidemiological risks.[98] Therefore, publication and related biases may be an important reason for many of the controversies surrounding the results of epidemiological studies.

Clinical trials

The impact of publication bias in clinical trials will depend on the extent of bias, and the underlying effects evaluated. The worst scenario would be where a harmful intervention is falsely reported as effective because of publication bias, which may result in patients receiving a harmful treatment. If an ineffective intervention is falsely considered as effective, patients may receive an ineffective treatment and be denied effective treatments. For an effective intervention, its effects may be

overestimated because of publication bias. New interventions are generally more expensive than conventional interventions, so overestimation of the efficacy of new interventions is likely to result in increased cost without a corresponding improvement in outcome.

A perinatal trial observed that routine hospitalisation was associated with more unwanted outcomes in women with uncomplicated twin pregnancies, but this finding remained unpublished for 7 years.[210] Chalmers pointed out that 'at the very least, this delay led to continued inappropriate deployment of limited resources; at worst, it may have resulted in the continued use of a harmful policy'.[210]

The non-publication of research findings may indirectly harm patients who are involved in future research. For example, a clinical study may find that an intervention is harmful but this finding is not published. Other investigators may subsequently repeat the same research, testing the harmful treatment on different patients. In 1980, a trial tested lorcainide in patients with acute and recovering myocardial infarction. More deaths were observed in the lorcainide group than in the placebo group (9/48 versus 1/47).[211] The trial results were not published because the development of lorcainide was stopped for 'commercial reasons'. About a decade later, an increased mortality was observed among patients treated with the related agents, encainide and flecainide, in two trials.[212,213] Encainide, flecainide and lorcainide all belong to a class of I_C antiarrhythmic agents. If the results of the trial in 1980 had been published, the increased mortality of patients included in the two later trials might have been avoided.

Recently there were several high-profile cases of alleged publication or reporting bias in drug trials. Rofecoxib was withdrawn from the market by Merck on 30 September 2004, because of increased risk of myocardial infarction and stroke according to unpublished data from a clinical trial.[214] Editors of *NEJM* expressed their concern about a trial of rofecoxib, published in the journal in 2000, in which three cases of myocardial infarction were not disclosed in the article.[215] Although authors of the trial denied any wrongdoing,[216] the *NEJM* editors restated their concern.[217] Further, in the year before the withdrawal, the authors of a systematic review had written directly to the primary author of every published trial of rofecoxib asking about

cardiovascular events and major bleeds; however, they received only a single reply and it did not provide data on cardiovascular events.[218]

In a more recent case study, Psaty and Kronmal (2008) found biased reporting of findings from clinical trials of rofecoxib for Alzheimer's disease or cognitive impairment.[219] Before its withdrawal, rofecoxib had been used in more than 80 million patients.[214] Biased reporting of findings from trials may have encouraged more patients to use rofecoxib and delayed the detection of harmful effects of rofecoxib.

In 2003, medicine regulatory authorities in several countries advised that a new antidepressant, paroxetine, should not be used in children with depression (with several other SSRIs added to the list later on), based mainly on findings from unpublished trials from industry.[220,221] This case actually suggests that formal publication may not be the most effective and timely approach to disseminating important research findings. Findings from clinical trials indicated that these SSRI antidepressants were ineffective and associated with increased suicidality and aggression in children with depression. Kondro and Sibbald (2004) revealed a drug company's internal document in which staff were advised to withhold data about SSRI use in children.[222] and GlaxoSmithKline was threatened with legal action over concealment of trial results.[223] A meta-analysis by Turner *et al.* examined 74 FDA-registered clinical trials of antidepressants and found that trials with positive results were more likely to be published than those with negative results.[45] Ioannidis suspected that some relevant trials conducted after market approval may not be included even in the FDA database.[224]

Summary

The most important consequences of publication bias include avoidable suffering of patients and waste of limited resources. This updated review identified only a couple of new cases that indicate the detrimental impact of publication and related biases. Consequences of publication and related biases are different for different types of research studies. Because of the possibility of such bias, the integrity of scientific research could have been jeopardised.

Chapter 6
Sources of publication bias

Biased selection for publication may occur to a varying degree during all stages of the publication process, from author submission and peer review to editorial decision, due to a variety of reasons.[20,225,226] Since bias is a natural human phenomenon,[227] publication bias may be introduced intentionally or unintentionally, consciously or unconsciously. This chapter provides an updated review of evidence about the responsibility of investigators, journal editors or peer-reviewers, and research sponsors for the existence of publication bias. Other study-level factors (including sample size, underlying true effect, study design and quality) that may exacerbate the risk of publication of a biased selection of studies are then discussed.

Investigators and authors

There are various reasons for not writing up an article or not submitting it, such as pressure from research sponsors and instructions from journal editors. The previous *HTA* report[2] included nine studies of reasons given by investigators for not publishing studies.[20–22,30,50,54,228–230] We identified an additional 12 studies in this updated review (Appendix 14).[28,35,67,70,72,231–237] It should be noted that studies often used different ways to categorise reasons for non-publication and reasons given in a study may not be independent of each other. For example, citing 'result not important enough' as the reason may be the cause of other given reasons such as 'not worth the trouble' and 'not enough time'.

Studies included in the previous *HTA* report and those newly identified reported similar reasons for not publishing (Appendix 14). Of the 21 studies included, there are five studies of investigators of protocol cohorts, 11 studies of authors of meeting abstracts, and five studies of other or miscellaneous authors. Percentages of specific reasons from individual studies were transformed to log odds and pooled using random-effects model, although there was significant heterogeneity across studies (*Figure* 7). The main reasons for non-publication were lack of time or low priority (34.5%; 95% CI: 27.4% to 42.3%), results not important enough

(19.6%; 95% CI: 12.0% to 30.4%) and journal rejection (10.2%; 95% CI: 5.5% to 18.2%) (see *Figure* 7). Pooled percentages of specific reasons were similar across different types of empirical studies, except that the lack of time or low interest were significantly higher in studies of meeting abstracts (43.1%; 95% CI: 35.9% to 50.6%) than in studies of protocol cohorts (23.8%; 95% CI: 15.9% to 34.0%) or studies of other authors (20.7%; 95% CI: 7.7% to 44.9%) (see *Figure* 7). In the five studies of meeting abstracts, fear of journal rejection was given as a reason for 23.7% (95% CI: 8.9% to 49.6%) of unpublished studies.

It should be noted that 'lack of time' may be more likely used as the excuse for not publishing unimportant results. The same researcher may have several different simultaneous studies that need attention, and may be reluctant to spend already limited time on the preparation of manuscripts for studies with unimportant or non-significant results that are less likely to be accepted by high-profile journals. In a qualitative study of causes of publication bias in genetic epidemiology, an experienced researcher in genetic epidemiology admitted that because of time constraints and 'piles' of results available, efforts will inevitably focus on the publication of 'wonderful results', not 'negative results'.[84] These findings indicated that investigators may be the main source of publication bias, for not writing up or submitting studies with 'unimportant' results.

Blumenthal *et al.* conducted a postal survey of 3394 life sciences faculty members at 50 universities that received the most funding from the NIH in 1993.[231] Delay to publication by more than 6 months in the last 3 years was reported at least once by 19% of the 2167 respondents. Principal reasons given by respondents for delay to publication included patent application submission (46%), protection of scientific lead (31%), patent negotiation (26%), time for resolution of intellectual property ownership (17%) and slow dissemination of undesired results (28%).[231]

According to a recent survey of 119 authors of papers published in six general medical journals, authors still considered that good study quality,
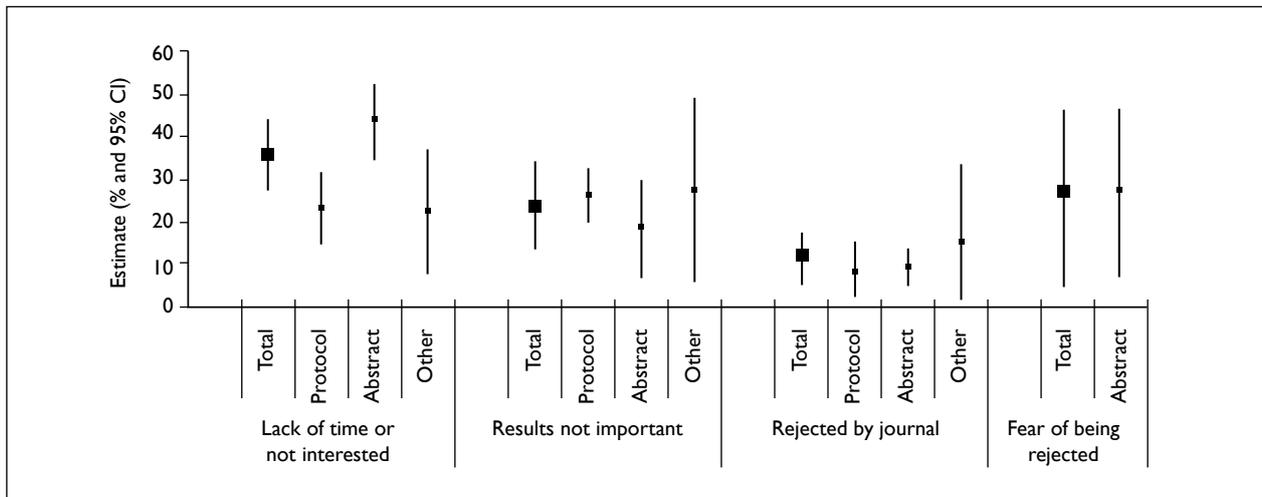
FIGURE 7 *Reasons given by investigators for not publishing: scatter plot.*

manuscript writing and statistical significance of results were important factors associated with the possibility of a study being published.[238]

Findings from surveys of investigators are supported by evidence from other studies. Stern and Simes found that quantitative studies with significant results were more likely to be submitted than studies with null results (78% versus 54%, $p < 0.001$).[24] Ioannidis found that studies with positive results were often submitted for publication more rapidly after completion than were negative studies.[23]

Cain and Detsky believed that even physicians can be biased.[227] A study investigating enthusiasm for radiotherapy after radical mastectomy when stage was not distinguished showed 21 out of 29 radiotherapists were enthusiastic compared with 5 out of 34 authors in other specialties.[239] A systematic review of risk of strokes and death following endarterectomy for symptomatic carotid stenosis showed a higher risk in studies where neurologists assessed the patients and lowest where the single author was affiliated to a department of surgery (7.7% versus 2.3%).[240]

Authors' criteria for selecting journals was investigated in a study of all active clinical and research faculty at Stanford University School of Medicine.[241] A response rate of 63.7% with factors ranked from unimportant (1) to very important (6) showed journal prestige (5.2), makeup of journal's readership (4.8), whether the journal publishes articles on the topic (4.8) and likelihood of manuscript acceptance (4.4) to be the important

factors at primary submission. For subsequent submission, manuscript acceptance (5.0) and whether the manuscript usually publishes articles on the topic (4.7) were the most important factors determining submission.[241]

McCambridge (2007) discussed a case of publication bias in reviews of drug education.[242] A series of systematic reviews of drug education in schools were conducted by Tobler *et al.*, and formally published in 1986, 1997 and 2000.[243–245] Findings from these reviews indicated that interventions delivered by mental health clinicians were more effective than those by others; and interactive programmes were effective and non-interactive programmes were not. These findings have had considerable impact on research, policy and practice. Through personal communication, McCambridge obtained some unpublished results of updated meta-analysis conducted by the same team of the previous three systematic reviews.[242] According to the unpublished results of updated meta-analysis, differences between different intervention programmes are no longer statistically significant, but these findings have not been formally published in peer-reviewed journals at 4 years after the end of the review project. The non-positive finding is likely to be one of the reasons for non-publication.[242]

Editorial review process
Editorial policies

Little is known about the actual editorial process itself. A semistructured interview of editors of

three leading biomedical journals found a great diversity in editorial policies and procedures between the journals.[246] A retrospective review of studies published in 2006 found that medical journals were more likely to publish reports from their own editorial board than from other journals.[247] Although editorial rejection was not a frequent reason given by investigators for studies remaining unpublished (see *Figure 7*), authors may not submit articles with 'unimportant' results because of anticipated rejection according to journals' instructions to authors and their own (or colleagues') experience.

The 2000 *HTA* report on publication bias included several studies that surveyed authors or investigators about manuscript submission for publication.[23,230,248] A study by Weber *et al.* reported that anticipated rejection by journals was cited as a reason for failure to submit a manuscript by 20% of 179 authors.[230] In another study, 17 of 45 submitted trials were rejected by at least one journal, and at least four negative trials with over 300 patients each were rejected two or three times, while no positive trial was multiply rejected.[23] In a survey of 80 authors of articles published in psychology or educational journals in 1988, 61% of the 68 respondents agreed that, if the research result is not statistically significant, there is little chance of the manuscript being published.[248] Several new studies identified in the updated review provide results similar to those reported in the previous studies.[67,233,236,237] Anticipated rejection by journals was the reason for not submitting a study given by 10% of investigators in the article by Vuckovic Dekic *et al.*,[237] by 13% in Hashkes and Uziel,[67] by 13% in Sprague *et al.*,[236] and by up to 26% in Hartling *et al.*[233]

The 2000 *HTA* report also included several studies that surveyed journal editors. A survey of 429 editors or members of advisory boards of 19 leading journals in management and the related social sciences in 1974 found that non-significant results, replications, lack of new data, similarity to recently published articles, or having previously been presented at meetings were factors associated with reduced chance of acceptance.[249] A survey in 1996 of 36 editors of English-language journals found that editors primarily valued the significance and importance of the research above validity of the experimental and statistical methods.[250] Originality and clinical significance of results are also important criteria for manuscript acceptance. Negative results may have less of an effect on clinical practice, supporting their

publication in pay-to-publish journals or open access electronic journals unless they show a widely used intervention is ineffective.[19,251] Qualitative criteria for assessing study importance include originality of results, predictability, triviality, narrow interest, highly specialised and few/no clinical implications.[252] Unoriginality accounted for 14% of all reasons given for rejection of manuscripts in 1989 by the *American Journal of Surgery*.[253] Confirmatory studies, either positive or negative, have a low chance of being accepted.[254,255]

The updated review included only one new relevant study that surveyed journal editors. A survey of the editors of 33 medical journals owned by not-for-profit organisations showed that 70% reported having complete editorial freedom and the remainder reported having a high level of freedom.[256] However, 42% reported being pressurised by the association's leadership, 30% by senior staff and 39% by rank-and-file members. Ultimately 48% of the journal's board of directors had authority to hire and 55% to fire the editor indicating that editorial independence from journal owners needs protection.[256]

Several cases of inappropriate instructions to authors by journal editors that may lead to publication bias were reported in the 2000 *HTA* report on publication bias. For example, a journal on diabetes clearly stated that 'mere confirmation of known facts will be accepted only in exceptional cases; the same applies to reports of experiments and observations having no positive outcome'.[255] More journal editors may have realised the detrimental impact of selective publication of positive results, and we are not currently aware of explicit journal instructions to authors that may be a cause of publication bias. However, further efforts will be required to translate this change in journal editorial policies to the submission behaviour of authors and investigators.

Journal peer review

Journal peer review has been defined as 'the assessment by experts (peers) of material submitted for publication in scientific and technical periodicals'.[257] Unacceptable biases in the peer review process include biases related to certain types of author (prestige, gender, nationality), or certain types of manuscript (language, innovation, positive/negative results).[225,258] The process of journal peer reviewing is a complex process and there are many studies on different types of biases in peer reviewing. This report considers only

biased peer reviewing process as a possible cause of selective publication according to study results. The 2000 *HTA* report on publication bias included several studies that used sham papers to investigate publication bias in the peer review process. No new relevant studies have been identified in this update review, and studies included in the previous *HTA* report are discussed below.

Mahoney sent a sham paper with identical experimental procedures but different results to 75 journal referees and found poor agreement between reviewers and bias against the manuscript that reported results conflicting with referees' own perspectives (confirmatory bias).[259] A further study gave a similar result, in which a sham paper about transcutaneous electrical nerve stimulation (TENS) was sent to 33 referees identified as pro- or contra-TENS.[260] Referees' judgement was associated with preconception and experience, and inter-rater reliability was again found to be poor.[260] Abbot and Ernst sent four versions of a sham study in complementary medicine to 200 authors, and found that the poor quality manuscript was rejected significantly more often than the good quality manuscript (55% versus 16%; $p < 0.05$), and no evidence of peer-reviewer bias against a positive or negative outcome.[261]

Abstract reviewing may be less predictable than reviewing a full article. Ector *et al.* compared the agreement between reviewers in grading abstracts submitted to a conference (the sixth European Symposium on Cardiac Pacing).[262] Each abstract was graded on a scale of 1 to 10 by two peer-reviewers. There was no statistically significant correlation between reviewers in 13 of the 28 pairs. It was suggested that reviewing abstracts is less predictable and more likely to be biased than reviewing a full article.[262] A study of 1983 posters submitted for three annual conferences for bias was conducted by Blackburn *et al.*[263] Posters having authorship that included at least one reviewer received higher ratings than those having only non-reviewing authors.[263]

Wager *et al.* compared reviews from reviewers selected by authors and those selected by editors, and found that reviewer source had no impact on review quality or tone but that author-nominated reviewers were significantly more likely to recommend acceptance and less likely to recommend rejection than editor-chosen reviewers after initial review.[264] Another similar study also found that editor-selected reviewers were less likely to recommend acceptance than author-

chosen reviewers, although there was no significant difference in review quality or speed between them.[265]

Geographical bias can also influence peer review. In a Scandinavian study, two versions of a sham paper with methodological flaws, one in Scandinavian and one in English, were sent to 180 Scandinavian reviewers.[144] The 156 referees who returned 312 reviews considered the English-language version significantly better than the Scandinavian version ($p < 0.05$).[144] A retrospective analysis of original submissions received by the journal *Gastroenterology* in 1995 and 1996 also showed geographical bias.[266] There were 2355 US and 1297 non-US reviewers ($p = 0.31$), with US reviewers recommending acceptance of papers submitted by US authors more often than non-US reviewers ($p = 0.001$). Non-US reviewers ranked US papers slightly more favourably than non-US papers ($p = 0.09$), with US reviewers ranking US papers much more favourably ($p = 0.001$).[266] However, a study of 3444 papers submitted to the journal *Cardiovascular Research* between 1997 and 2002 showed that US reviewers assigned significantly higher priority to manuscripts regardless of where the manuscript was from ($p < 0.0005$).[267] The same study also found that manuscripts received significantly higher priority ratings when reviewers and authors originated from the same country ($p < 0.05$).[267]

Gender bias during peer review of manuscripts and grant proposals has also been demonstrated in several studies. A study of manuscripts received by *JAMA* in 1991 comprising 1698 male and 462 female authors with eight male editors, five female editors, 2452 male and 930 female reviewers showed significant gender differences.[268] Female editors were assigned manuscripts from females more often than males ($p < 0.001$). Female editors also used more reviewers per manuscript if sent for other review and rejected more manuscripts ($p < 0.001$).[268] However, articles submitted were not accepted at significantly different rates based on gender ($p < 0.4$). A Scandinavian study used a sham paper with either a female or male author to assess gender bias in 1637 randomly selected Swedish physicians.[269] Female authors were ranked higher than male authors, with female assessors upgrading female authors more than male authors and male assessors reflecting no gender differences.[269] A study conducted by Caelleigh *et al.* of 50 female and 50 male reviewers showed no gender bias when assessing an empirical study with two versions, one attributing lower forecast income of women to

intrinsic gender factors and the other attributing the difference to extrinsic social learning factors.[270]

In a retrospective study, the effect of institutional prestige on referees' recommendation and editorial decision was assessed.[271] Institutional prestige was determined according to the monetary value of research and training grants and contracts funded by the National Institutes of Health. The association between the recommendation for acceptance and institutional prestige was observed for the 147 brief reports (i.e. case reports and similar short papers) but not for 258 major papers (such as case series, research reports and epidemiological studies).[271]

Study results and journal editorial decisions

Rejection by journals was given by investigators as a reason for 5% to 33% of non-publication of studies (see *Figure 7*). If the decision to accept or reject studies for publication is not based on study findings, the rejection of studies by journals will not result in publication bias.

The 2000 *HTA* report on publication bias included several studies that provided limited evidence on the acceptance of manuscript by study results. Epstein sent two versions of a fictitious paper with either positive or negative results to 146 social work journals.[272] The positive manuscript was accepted in 35% and the negative manuscript was accepted in 25% ($p > 0.05$).[272] A study found that 17 of 45 submitted trials were rejected by at least one journal, four negative trials were rejected two or three times and no positive trial was multiply rejected.[23] However, another study found no difference in the rate of publication of submitted manuscripts between studies with significant results and studies with null results (87% versus 82%, $p = 0.54$).[24] In a case–control study of 100 accepted and 100 rejected papers in two Spanish medical journals, it was found that publication status was associated with high study quality, not positive findings.[273]

The updated review identified four studies that followed cohorts of manuscripts submitted to journals (Appendix 15).[78–81] Results of these four studies of manuscript cohorts have been discussed in Chapter 3. Two studies examined manuscripts submitted to general medical journals (*JAMA, BMJ, The Lancet,* and *Annals of Internal Medicine*)[78,81] and two used manuscripts submitted to the *Journal of Bone and Joint Surgery (American Version).*[79,80] The study results of submitted papers were classified according to the significance of statistical testing ($p < 0.05$ or not) in the two studies of manuscripts submitted to general medical journals.[78,81] In the studies of manuscripts submitted to the *Journal of Bone and Joint Surgery*, results were classified as being positive, negative or neutral, although the definitions of these outcomes may be different between the two studies (Appendix 15).[79,80]

Figure 8 shows the results from the four studies and the pooled odds ratio of acceptance rate (OR 1.06; 95% CI: 0.80 to 1.39), which suggested that the acceptance of submitted papers for publication by journals was not significantly associated with the direction or strength of their findings. In addition, Olson *et al.*[81] further examined 133 accepted manuscripts and found that time to publication was not associated with statistical significance (median 7.8 months for positive and 7.6 months for negative results, $p = 0.44$).[82]

Because the acceptance of manuscripts for publication by journal editors was not determined by the direction or strength of study results, the existence of publication bias may be largely due to biased selection of studies to submit by investigators. This may also be supported by the fact that a large proportion of submitted papers showed statistically significant results (51% to 87%) or positive results (71% to 72%) in the four cohort studies. Since any author will inevitably consider the possibility of their manuscripts being accepted before submission, submitted studies with negative results may be a biased selection of all studies with negative results.

In Olson *et al.*'s cohort study of manuscripts submitted to *JAMA*, there was a tendency that studies with significant results had a higher rate of acceptance than studies with non-significant or unclear results (20.4% versus 15.2%, $p = 0.07$).[81] In the cohort study by Okike *et al.*, a subgroup analysis of 156 manuscripts with a high level of evidence (level I or II) found that the acceptance rate was significantly higher for studies with positive or neutral results than for studies with negative results (37%, 36% and 5% respectively; $p = 0.02$).[80]

The studies included in Appendix 15 are generally well designed and conducted. Although no conflict of interest was declared in the four cohort studies of submitted manuscripts, this kind of study will always need support or collaboration from editors of the journal. In prospective studies, editors' decisions on the acceptance of manuscripts may

Study or subgroup	Positive		Non-positive			Odds ratio M-H, Random, 95% CI	Odds ratio M-H, Random, 95% CI
	Events	Total	Events	Total	Weight		
Lee 2006[78]	35	718	7	109	9.9%	0.75 (0.32 to 1.73)	
Lynch 2007[79]	45	148	19	61	15.7%	0.97 (0.51 to 1.84)	
Okike 2008[80]	132	620	54	235	38.5%	0.91 (0.63 to 1.30)	
Olson 2002[81]	78	383	55	362	35.8%	1.43 (0.98 to 2.09)	
Total (95% CI)		1869		767	100.0%	1.06 (0.80 to 1.39)	
Total events	290		135				
Heterogeneity: $\tau^2 = 0.02$; $\chi^2 = 3.83$, df = 3 ($p = 0.28$); $I^2 = 22\%$							
Test for overall effect: $z = 0.39$ ($p = 0.69$)							

0.2 0.5 1 2 5
Favours non-significant   Favours significant

FIGURE 8 *Acceptance rate and results of studies submitted to journals for publication. Unadjusted odds ratio.*

be influenced by their awareness of the ongoing study.[81] Therefore, biased selection for publication by journals cannot be completely ruled out.

Readers and users of research findings

Journal editors' policy may reflect readers' preferences, and it has been suggested that editors should find ways to incorporate the reader's perspective into the peer review process and study the effects of their efforts. It is likely that readers' preferences for certain findings may be an important reason for the biased publication of studies in journals. We have identified no new relevant studies in the updated review, although two studies were discussed in the previous *HTA* report on publication bias. A survey of 452 readers showed readers were generally satisfied with the quality of manuscripts but dissatisfied with the lack of manuscripts relevant to medical practice.[274] The difference of opinion between readers and peer-reviewers may be attributable to clinicians avoiding unestablished treatments but journals being more likely to accept for publication manuscripts with novel treatments.[274]

Research funding bodies and commercial interests

Research commissioning bias may contribute to publication bias since industry sponsors research and often own the data, making them susceptible to manipulation and suppression. Rosenberg noted the conflict between dissemination of research findings with the protection of investors who have

supported the research that pervades modern science.[275] An editorial in *JAMA* noted that 35% of signed agreements in a sample of university-industry research centres allowed the sponsor to delete information from publication, 53% allowed publication to be delayed and 30% allowed both.[276]

The 2000 *HTA* report on publication bias included several studies that investigated association between study results and industry sponsorship in biomedical research. A study of clinical trials published in 1984 in five general medical journals showed 89% of drug company-funded trials supported a new therapy compared with 61% of generally funded trials ($p = 0.002$).[277] Another study of 56 RCTs published between 1987 and 1990 and concerning NSAIDs in the treatment of arthritis showed that in all trials manufacturer-associated drugs were reported to be comparable with (71%) or superior to (29%) the control drugs.[278] Of the 22 trials that reported a drug with less toxicity, the manufacturer-associated drug's safety was reported to be superior in 86% of cases with justification provided in only 55%, suggesting selective publication or biased interpretation of results in manufacturer-associated trials.[278]

Stelfox *et al.* examined the published safety profiles of calcium-channel antagonists and the financial association of authors with the pharmaceutical industry.[279] They identified 77 articles and a questionnaire was sent to 86 authors of 70 articles, of whom 69 authors completed the survey. Of the authors that supported the safety of calcium-channel antagonists, 96% had financial relationships with the manufacturers compared to 60% of neutral authors and 37% of critical authors ($p < 0.001$).[279] Therefore, the

importance of full disclosure of relationships with pharmaceutical manufacturers in journal articles was highlighted.[279–284]

The updated review identified several recently published reviews or primary studies about industry sponsorship in biomedical research.[285–287] A systematic review published in 2003 by Bekelman *et al.* found that industry research funding was received by 23% to 28% of academic researchers, and was associated with restrictions on open collaboration, data access or publication of results.[285] Pooling of results from eight studies found that industry sponsorship was statistically significantly associated with pro-industry conclusions (pooled OR 3.60; 95% CI: 2.63 to 4.91).[285] A similar systematic review by Lexchin *et al.*, also published in 2003, found that 'research funded by drug companies was less likely to be published than research funded by other sources', and industry-sponsored studies were more likely to report outcomes favouring the sponsor compared with studies supported by others (pooled OR 4.05; 95% CI: 2.98 to 5.51).[286] Findings of the above two systematic reviews in 2003[285,286] are confirmed by results of recently published studies.[288–298] Jorgensen *et al.* compared Cochrane reviews with industry-supported meta-analyses of the same drugs. They found that industry-supported meta-analyses of drugs 'were less transparent, had few reservations about methodological limitations of the included trials, and had more favourable conclusions than the corresponding Cochrane reviews'.[299]

Sawka and Thabane pointed out significant heterogeneity across studies in the meta-analysis by Bekelman *et al.* of results of industry-sponsored studies, and suggested that the pooled odds ratio is 'unconventional'.[300] In many studies included in the two systematic reviews,[285,286] industry-sponsored studies may not be comparable with non-industry-sponsored studies from many perspectives,[301] although they had similar methodological quality. Studies that included homogeneous research in terms of patients and interventions seemed less likely to find significant differences in results between industry-sponsored and non-industry-sponsored research. For example, in a meta-analysis of trials of antimuscarinic medications for overactive bladder, Tulikangas *et al.* found 'no difference in outcomes when comparing studies funded by industry or not for tolterodine and oxybutynin'.[302] Barden *et al.* investigated industry bias using comparable trials on acute pain and

migraine and found no evidence indicating that industry-sponsored trials on acute pain and migraine were biased.[301]

Therefore, publication bias (including outcome reporting bias) is only one of several possible explanations for observed association between favourable results and the industry sponsorship. However, direct evidence showing industry's commercial interests as a source of publication bias does exist.[303] Identified case studies on biased reporting of research due to commercial interests are summarised in Appendix 16.[136,214,215,217,220,276,304–332]

Some pharmaceutical companies attempted to suppress the publication of 'negative' results by taking legal action, and all cases occurred before 2000.[276,304–309] One company took legal action against an investigator in order to stop the publication of negative results from a study on deferiprone in patients with thalassaemia.[305,306] A study by Dong *et al.* showing bioequivalence of generic and brand name levothyroxine was suppressed by the pharmaceutical company due to the deleterious effect of the results on the price of the company's product.[333] A pharmaceutical company also tried to suppress a systematic review that would have had a negative economic impact on statins.[334] Publication of a meta-analysis with unsupportive results of bovine somatotrophin was blocked by a pharmaceutical company using its legal rights over the raw data.[304]

It seems that industry is no longer able to suppress the publication of results of entire sponsored research. However, the updated review identified several new cases in which results of industry-sponsored research were selectively reported or misrepresented in publication (Appendix 16).[136,217,219,220,222,323,324,331,332,335] Non-publication of 'negative' results was common.[317–320,322,336] For example, Psaty and Kronmal compared published and unpublished mortality findings in two trials of rofecoxib for Alzheimer's disease.[219] The two published articles only mentioned on-treatment mortality in the text without any statistical analyses, and concluded that rofecoxib is well tolerated.[330,337] However, the company's unpublished intention-to-treat analyses and the independent analyses based on data provided by the sponsor in the New Jersey Vioxx litigation found a statistically significant increase in total mortality (HR 2.99; 95% CI 1.55 to 5.56; and HR 2.13; 95% CI: 1.55 to 5.77, respectively).[219]

47

The direct evidence included in Appendix 16 was mainly restricted to some high-profile cases where investigators determined to challenge industry's suppression of publication, or where the open access policy facilitated the identification of discrepancies between published and unpublished results. There may be many hidden cases where research results were not disclosed because investigators gave in to the pressure from research sponsors.

A more recently published systematic review included seven studies that compared the reporting of adverse effects according to funding sources.[287] There was no clear evidence that the reporting of the raw adverse effects data was biased. However, a drug was more likely to be interpreted as safe by industry-funded authors compared with authors without pharmaceutical funding.[287]

Variation in study results

If the results from all possible studies were the same or similar, selected publication of results would not be biased. Greater variation in the results may be associated with an increased risk of publication bias. Factors that influence variation in study results include small sample size, small or moderate effect size, subjective nature of outcome measurement, and complex interventions. However, we have not been able to identify any studies that provide direct empirical evidence on results variation and publication bias. In this section, the updated review

Size *n*	True OR = 1.0	True OR = 0.8	True OR = 0.6
50	−0.017	0.223	0.270
100	0.019	0.206	0.164
200	0.038	0.139	0.118
300	0.016	0.125	0.042
400	−0.009	0.083	0.022
500	−0.026	0.084	0.014
600	0.013	0.064	0.001
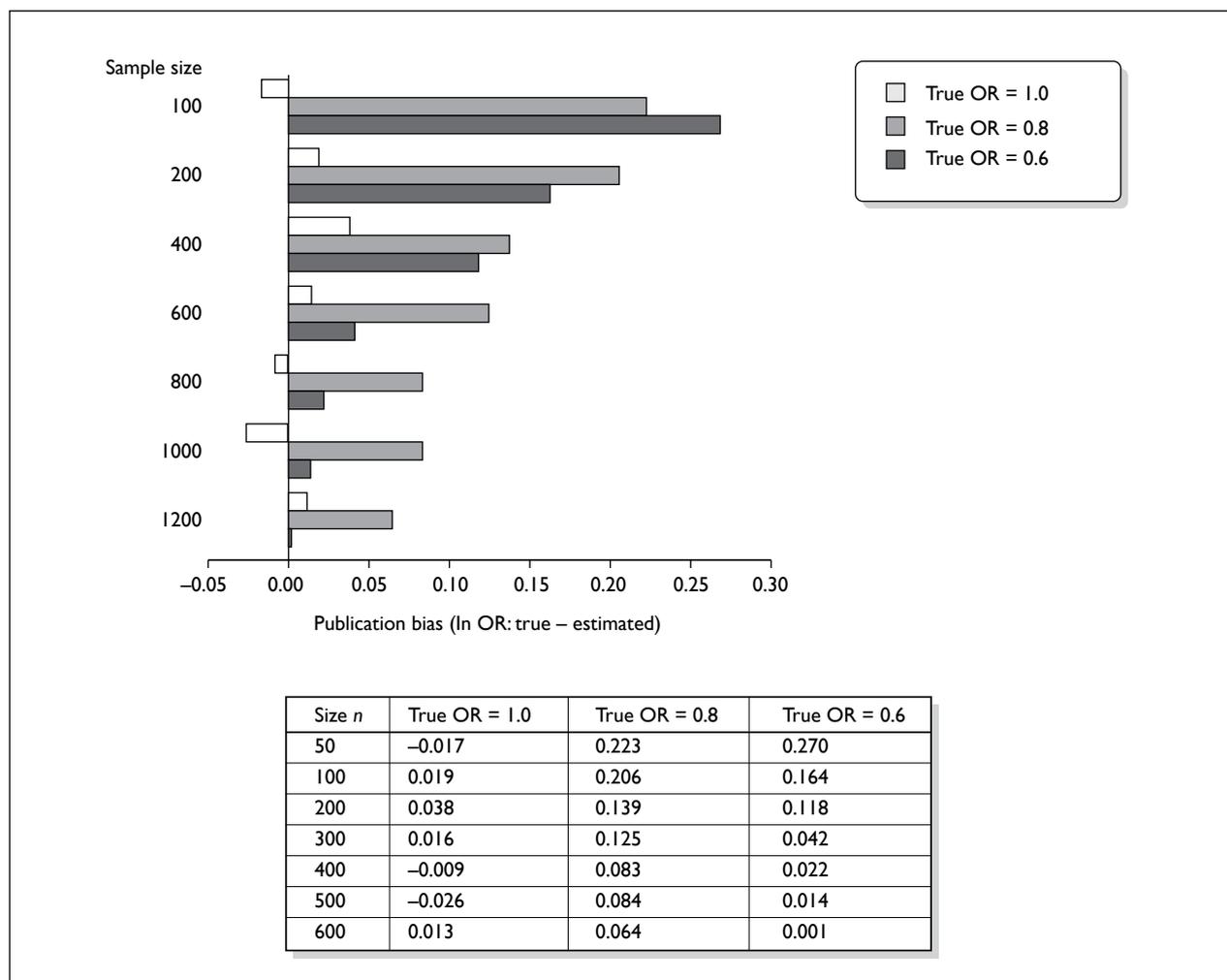
FIGURE 9 *Extent of publication bias and range of sample sizes: results of 1000 times simulation (selection based on p values; control rate = 0.20).*

True OR (ln OR)

True OR	n = 100	n = 200	n = 400
1.0	0.044	0.030	−0.023
0.9	0.073	0.066	0.134
0.8	0.148	0.177	0.164
0.7	0.285	0.199	0.146
0.6	0.269	0.177	0.097
0.5	0.284	0.172	0.053
0.4	0.203	0.060	0.003
0.3	0.131	0.034	−0.013
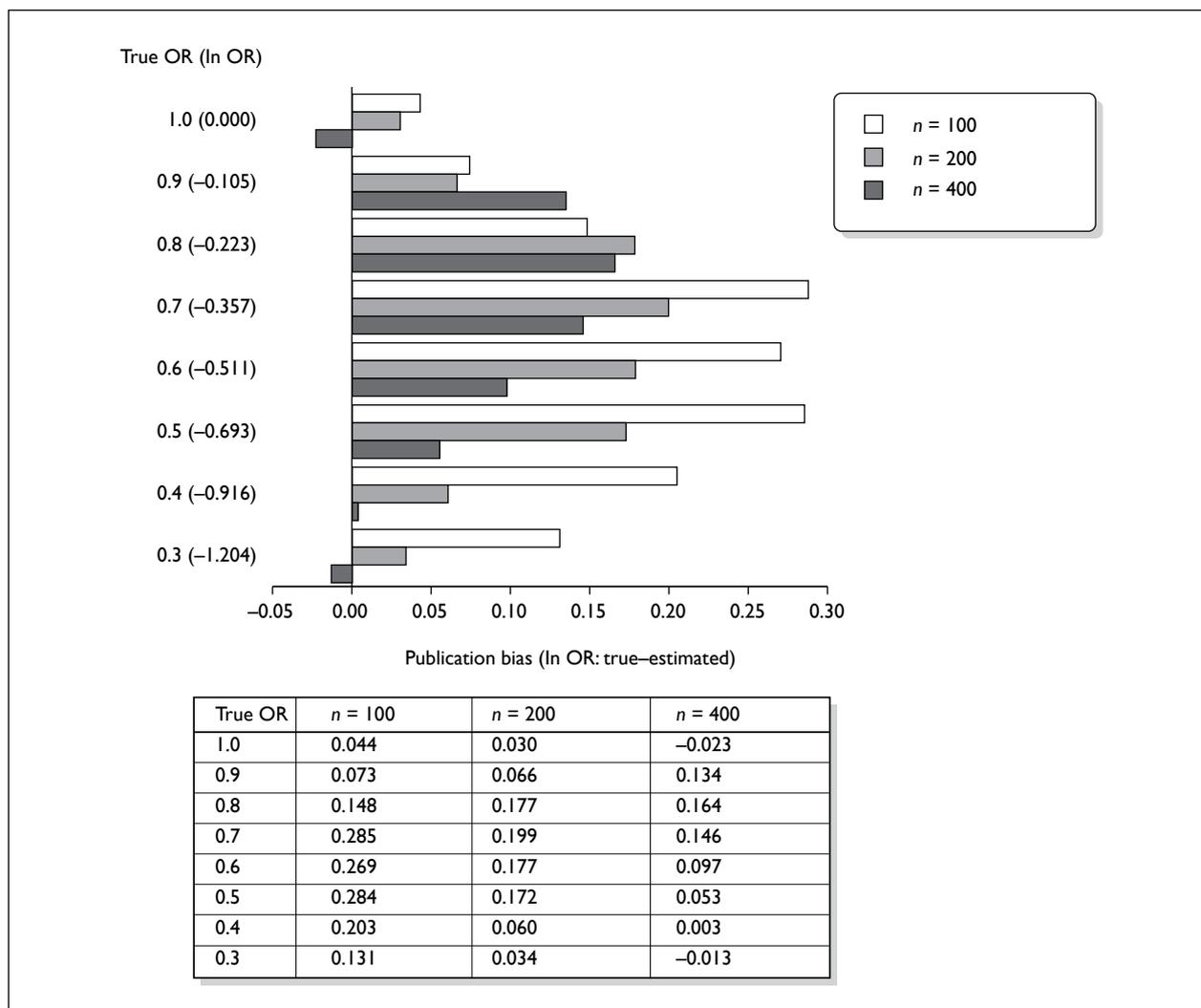
FIGURE 10 *Extent of publication bias and the true effect size (ln OR): results of 1000 times simulation (selection based on* p *values; control rate = 0.20).*

Small sample size

Studies with small sample sizes tend to produce variable results and present a range of results to select for publication. Simulations have demonstrated that small sample size is associated with considerable publication bias when only studies with significant results are published.[338,339]

In practice, a small study with a non-significant result may be readily abandoned without trying to publish because it is easy and cheap to carry out in terms of time, staff and other resources invested. In addition, small trials may often be poorly designed and conducted. Therefore, the risk of

publication bias will be great if many small trials have been conducted.[340] However, small trials may still be helpful in many aspects, and publication bias should not be considered 'as a good reason to discourage trials with low power'.[341]

Figure 9 shows the results of a stochastic simulation investigating the relationship between publication bias and the range of possible sample sizes. Given a true odds ratio of 0.73 and other conditions assumed in the simulation, the estimated odds ratio is 0.23 when the sample sizes range from 20 to 100, 0.44 when the sample sizes range from 20 to 500, and 0.70 when the sample sizes range from 20 to 5000. When the possible sample sizes range from 20 to 10,000, the estimated odds ratio is 0.72, nearly identical to the true value of 0.73. Thus, the extent of bias due to selective publication of

significant results is reduced when there are many large-scale trials.

Small effect size

The simulation results also indicated that the extent of bias, by selecting the significant results to publish, is greater when the true effect is small or moderate than when the true effect is zero or large.[339] *Figure 10* shows the results of a computer simulation about the relation between the true effect (log odds ratio) and the extent of bias. The difference between the true and the biased effect was large when the true effect is small compared with that when the treatment effect is zero or larger. Therefore, a small or moderate effect (or weak association) can be considered as a risk factor for publication bias. This risk factor may exist in most cases because clinical trials are mainly designed to assess health-care interventions with small or moderate (but clinically important) effects.

Study design and other quality characteristics

The design quality of studies may be associated with the risk of publication bias. Non-randomised studies, single-centre studies, and phase I and II trials might be more susceptible to publication bias than randomised studies, multicentre studies and phase III trials.[12,342] Risk factors for publication bias were assessed but not consistently identified across several cohort studies of publication bias.[20–22,24] Irwig and colleagues[343] suggested that publication bias is more of a problem for diagnostic tests than for randomised trials because 'many studies of test accuracy may use data collected primarily as part of clinical care, there may be no clear record of attempted evaluations'. It is therefore useful to estimate how easy it would be for investigators to abandon a completed study with unimportant results without publication, according to some study characteristics.

Summary

Investigators, peer-reviewers, editors and funding bodies may all be responsible for the existence of publication bias. The dissemination profile of a research finding is determined by the interests of research sponsors, investigators, peer-reviewers and editors. Evidence from newly identified studies confirmed findings from the previous *HTA* report that publication bias is often due to investigators' failure to write up and submit. However, it should be recognised that the investigators' decision to write up an article and then submit it may be affected by pressure from research sponsors, instruction from journal editors, and requirements of the research award system. Newly identified as well as previous included evidence indicates that the interests of research sponsors, particularly industry's commercial interests, can restrict the dissemination of the research findings. Large differences in likely study results across similar studies that can be easily conducted and abandoned will further exacerbate the biased selection of findings for publication.

Chapter 7

Prevention of publication bias

Measures that may prevent publication bias should be logically designed according to the likely sources of such bias. Although investigators, peer-reviewers, editors and funding bodies may all be responsible for the existence of publication bias, the importance of these responsibilities in terms of preventing publication bias may be different. As discussed in Chapter 6, dissemination biases are related to many complicated factors, and these factors are inter-related. People's tendency to notice only a portion of relevant research results has complicated social, cultural, political, economic and psychological bases. In spite of these difficulties, it is possible that biased publication of research and the impact of publication bias may be prevented to a certain extent and its impact minimised. In this chapter, we review measures that may help to reduce the existence and impact of publication bias, including changes to the publication of research, electronic publishing, an open access policy, the prospective registration of studies at inception and confirmatory large-scale studies.

Changes in publication process

Because of the huge number of published studies and specialist information needs, health-care practitioners, policy-makers and researchers must selectively receive information that is perceived relevant. At the same time, curiosity about new and atypical events by general readers means that to maintain a journal's circulation, editors may have to accept studies for publication according to readers' preference and type of information that readers required.

Investigators might not write up and submit studies with negative results because of anticipated rejection according to journals' instructions to authors and their own experience. The 2000 *HTA* report on publication bias listed some measures that could reduce publication bias by journals, including accepting manuscripts for publication mainly based on research protocols,[344] making prospective registration of trials a precondition for their publication,[345] disclosing conflict of interest

or competing interests,[280] and electronically publishing and archiving research.[346] Early initiatives included that by *The Lancet*, a general medical journal, which in 1997 began to assess and register selected protocols of randomised trials and systematic reviews, and to provide a commitment to publish the main findings of the study.[347] In the same year, over 100 medical journals around the world invited readers to send in information on unpublished trials in a so called 'trial amnesty'.[348]

Recently, biomedical journals have launched several initiatives and made important progress in the prevention of publication bias. The international guidelines for writing and editing publications[349] may help to prevent incomplete and biased reporting by the endorsement of sound reporting guidelines for specific study designs, including CONSORT (Consolidated Standards of Reporting Trials) for randomised controlled trials, STARD (Statement for Reporting Studies of Diagnostic Accuracy) for studies of diagnostic accuracy, QUOROM (Quality of Reporting of Meta-analyses) for systematic reviews, and STROBE (Standards for the Reporting of Observational Studies in Epidemiology) for observational studies in epidemiology.[349] To help editors, peer-reviewers and authors to ensure transparent and complete reporting of health research, the EQUATOR (Enhancing the Quality and Transparency of Health Research) network has been developed.[350–352] Further empirical evidence is required to indicate the impact of reporting guidelines on reporting bias.

Below, we discuss the prevention of publication bias by improved peer review, disclosure of competing interests, and electronic publication. The role of medical journals on the prospective registration of trials will be discussed later under 'Prospective registration of trials'.

Peer review process

Journal peer review has been defined as 'the assessment by experts (peers) of material submitted for publication in scientific and technical periodicals'.[257] The aim of peer review is to improve the general quality of published studies

and screen out articles with flawed methodology or conclusions.[253,353] A large number of studies on the peer reviewing process are available but few are directly relevant to the prevention of publication bias. The previous *HTA* report on publication bias included some studies showing that the general quality of peer review had not been improved by blinding peer-reviewers to authors' identities.[354–357] A recently published systematic review of studies on editorial peer review by Jefferson *et al.*[358] identified 19 comparative studies on editorial peer review. They included nine RCTs that investigated the effect of blinding on peer review. Of the nine RCTs, five found no effect of blinding on review quality and four found that blinding affected review quality although these studies highlighted the difficulty of ensuring robust blinding procedures. The effect of a submission checklist was investigated in two studies, with one showing no benefit and the other showing some benefit (though this study had a small sample size). The limitations of this systematic review (as noted by the authors) included atypical settings, involvement of few major journals, small numbers of reviews and reviewers, and methodological weaknesses making validity of the studies reviewed difficult to assess.[358]

Disclosure of commercial interest

In order to reduce bias due to research funding many journals require authors to disclose their 'conflict of interests' or 'competing interests'.[349] The updated review identified several studies of disclosure of commercial interest of authors. Cooper *et al.* performed a study of the characteristics of conflict of interest policies of biomedical journals with regards to authors, peer-reviewers and editors using a survey.[359] The response rate for the survey was 67%, with 93% of journals reporting having an author conflict of interest policy and 11% reporting that they restricted author submissions based on the conflict of interest policy. However, whilst 77% of journals reported collecting conflict of interest information, only 57% published author disclosures. Of interest, only 3% of respondents published conflict of interest disclosures of peer-reviewers and 12% published editor conflict of interest disclosures.[359]

Conversely, a study by Cain *et al.* showed that disclosure may exacerbate bias rather than prevent it with the possibility of a conflict of interest statement subconsciously absolving the author of responsibility.[360] Whilst a conflict of interest statement may appear transparent in acknowledging bias, it is insufficient to prevent it.[361]

Electronic publication

The volume of electronic publishing has greatly increased due to its advantages of rapid publication, no limit in length of articles, no limit in numbers of studies, interconnected articles, cost-effective dissemination, and cost-effective archiving.[362–365] Because of reduced or no space limitations, electronic publishing may reduce publication and related biases by publishing research protocols and by allowing studies to be judged on their design and methodology rather than the immediate relevance of findings to current practice, novelty or exciting results.[366]

Publication of research protocols prior to study completion has been recommended as a measure to prevent poor medical research.[346] The development of electronic publishing has provided great potential for the publication of research protocols. Any discrepancies between the research protocols and published studies will become transparent and outcome reporting bias may be prevented by the publication of protocols.[346]

Electronic journals with no space limitations may encourage publication of studies with negative or no significant results as well as those that replicate previous studies.[366] Peer review in the context of electronic publishing is still important to ensure quality, but this could also be published in conjunction with the article. For example, online BioMed journals publish any accepted papers with their initially submitted versions and comments from peer-reviewers plus responses from authors to these comments.

There are two forms of electronic publishing: printed journals with electronic supplementary materials, and electronic only journals. Sim and Rennels have suggested using the Trial Bank Model to publish traditional prose form studies and a concurrent electronic database of additional data.[367] This model has been adopted by the *BMJ* using the 'electronic long–paper short' (ELPS) mode of publishing.[368] Medical journals have two basic functions: medical recorder and medical newspaper.[369] The ELPS model, and the web-based supplementary material model, seem logical choices for these two different but related basic functions.

As a specific remedy for publication bias, the *Journal of Negative Results in Biomedicine*, an open access online journal, was introduced in 2002 with a remit to publish studies with negative results.[370] This journal is indexed by MEDLINE, EMBASE, Scopus and Google Scholar, making retrieval of the negative results it publishes more likely during systematic review. From inception to August 2008 it had published 60 articles.

The recent development of electronic publishing has provided great opportunities for preventing publication and related biases, but we have found little direct evidence of how effective they are in practice. Electronic publishing itself won't be able to resolve biased publication and reporting of research results. Many online open access journals [including *BioMed* and *Public Library of Science* (*PLoS*)] charge authors a fee to cover the publishing costs. It is still unclear what impact this pay-to-publish model has on publication and related biases and the general quality of studies published in these open access electronic journals.

Prospective registration of trials

In 1997, over 100 medical journals around the world invited readers to send in information on unpublished trials in a so called 'trial amnesty'.[348] This was a request to retrospectively register conducted trials, even where outcome data were not provided, to enable other researchers, specifically systematic reviewers, to know of the existence of the study and write to the trialists for further details. One year after its launching, only 165 trials were registered[371] and it was considered a failure by 2004.[372] This failure of retrospectively registering unpublished results led to further support for the development of prospective registration of studies.

Accepting studies for publication based mainly on their pre-submitted research protocol could help to reduce publication bias by ensuring that the publication maintains its pre-stated primary outcomes, and is published regardless of whether that primary outcome shows a statistically significant effect. In 1997, a general medical journal, *The Lancet*, began assessing and registering selected protocols of randomised trials and systematic reviews, and providing a commitment to send for peer review the main clinical findings of the study (there is currently no commitment to publish a final paper)[347] (see also http://www.

thelancet.com/journals/lancet/misc/protocol). Only 75 protocols had been accepted and registered in *The Lancet* by June 2007,[373] and up to the time of writing (August 2008) fewer than 100 protocols had been registered in this way; other journals are yet to follow suit. Registration of study protocols by paper journals, although a good idea, is clearly not going to help to prevent publication bias in the bulk of research, and it is still feasible for a study to be registered with *The Lancet*, but not published by them if the results are not deemed appropriate for whatever reason.

Prepublication of protocols has been recommended as an important measure to prevent poor medical research.[346] Electronic publication of systematic review protocols prior to study completion has become part of the online-only *Cochrane Database of Systematic Reviews*, where all accepted reviews have their protocols peer reviewed and published before the completed review is published. This allows for feedback by any reader on the methodology or question addressed and prevents duplication of effort in defining research that is about to be undertaken. The development of electronic publishing has provided great potential for the publication of research protocols. Theoretically this ensures that any discrepancies between the research protocols and published studies become transparent and outcome reporting bias may be prevented.[346]

Boissel *et al.* defined a clinical trial registry 'as a database of planned, ongoing or completed clinical trials, published as well as unpublished, in which details concerning the trial's objectives, main design features, sample size, and tested treatment are stored'.[374] It has been generally accepted that prospective registration of trials at their inception may prevent publication bias.[103] Even if not all trials are registered, a prospective registration of some trials may provide an unbiased sample of all studies that have been conducted.[1] For example, the International Cancer Research Data Bank was used to assess alkylating agent therapy in advanced ovarian cancer.[127]

The Clinical Trials Registry of the International Committee on Thrombosis and Haemostasis, established in 1974, may be the first registry of clinical trials.[375] In 1988, Easterbrook identified 24 registries of clinical trials.[376] Clinical trials included in these registries were prospectively or retrospectively identified by surveying selected individuals, organisations, pharmaceutical companies or other industries; from conferences

and selected journals; searching other related registries of trials; and by funding bodies or research ethic committees.[376] Currently, the two most important registries of clinical trials are the International Standard Randomised Controlled Trial Number (ISRCTN) register and ClinicalTrials.gov. The ISRCTN (available at http://www.Controlled-trials.com) was launched in 2000 by the publisher Current Science Group and its ownership was transferred to a not-for-profit entity in 2005.[377] ClinicalTrials.gov (at www.ClinicalTrials.gov) was developed in 2004 by the National Library of Medicine. As of August 2008 ISRCTN included over 7000 trial registrations, while ClinicalTrials.gov had registered almost 60,000. Other important trial registries include the Australian Clinical Trials Registry, the Netherlands Trial Registry, and UMIN (University Hospital Medical Information Network) Clinical Trials Registry.[378]

Voluntary registration of clinical trials is often incomplete and a mandatory system by government regulation may be necessary.[379,380] Spain's Royal Decree of 1978 and a Ministerial Order of 1982 established a register of clinical trials.[381] The FDA Modernisation Act of 1997 in the USA also requests the establishment of a federally funded database containing information on both government-funded and privately funded trials of drugs designed to treat serious or life-threatening conditions.[382]

The World Health Organization (WHO) has recognised the importance of trial registration and initiated a project in 2005 to set international standards for clinical trial registration.[383] The WHO Registry Network provides prospective trial registries with a forum to exchange information and work together to establish best practice for clinical trial registration. By establishing international standards, the WHO Trial Registry Portal will help to prevent selective dissemination of trial information by specific trial registers. The Clinical Trials Search Portal (CTSP) could be searched online by users to identify registered trials included in WHO's trial registration database.[384]

The most influential initiative is the introduction of a trials registration policy by the International Committee of Medical Journal Editors (ICMJE) in 2004.[10] Since 2005, as a condition of consideration for publication in ICMJE member journals, trials must register at or before the onset of patient enrolment in a registry that is accessible to the public at no charge, open to all prospective

registrations, and managed by a not-for-profit organisation.[10] A similar compulsory registration of clinical trials is also required by *BMJ*, with modified criteria for suitable registries.[372] The policy of compulsory trial registration required by journals has greatly increased the number of trials registered.[385] The number of trials registered in the ClinicalTrials.gov database increased by more than 70% between May and October 2005, after the implementation of ICMJE's policy.[386]

Even prospective registration of trials may not be sufficient to prevent all biases in the publication of trial results, including outcome reporting bias. The registration of all trial results (published and unpublished) has been advocated.[387,388] Although there is still disagreement about the registration of trial results,[378] publicly available full summary results from trials will help to reduce publication bias and outcome reporting bias. In addition, the prospective registration of research has currently focused mainly on phase III and phase IV trials. Biased selection for publication of early stage trials and non-trial studies (such as basic research, observational studies and studies of diagnostic accuracy) is still far from being resolved. For example, biased publication of early stage phase I studies may increase the failure rate of phase II studies.[389] Choi *et al.* suggested a global registry of anticipated public health studies. However, the establishment and maintenance of a comprehensive registration of non-trial studies will need to overcome more difficulties.[390]

Trials registries will only be helpful in reducing publication bias if systematic reviewers include the registries in their search strategies and results of trials are accessible. Ramsey and Scoggins identified 2028 completed or terminated cancer trials from NIH's ClinicalTrials.gov registry in September 2007.[391] They then searched ClinicalTrials.gov and PubMed for peer-reviewed publications of these trials. It was found that only 19.5% of the 1791 completed cancer trials and 3.4% of the 237 terminated trials had been published in peer-reviewed journals.[391]

Open access policy

Since 1997, the practice of incomplete release of information about licensed drugs in Europe for reasons of commercial interests and intellectual property has been challenged.[392–394] Abraham and Lewis[392] suggested that 'the secrecy and confidentiality of EU medicines regulation is not

essential for a viable pharmaceutical industry', considering that European pharmaceutical companies often obtain data on competitors' products by using the US Freedom of Information Act. There were already 'encouraging signs' in 1998 within the pharmaceutical industry to improve public access to the findings of clinical studies that the industry sponsored.[394] However, there were setbacks in transparent reporting of clinical trials sponsored by the pharmaceutical industry until several recent high-profile cases of incomplete reporting of industry sponsored trials.[380]

Recently, policies of mandatory open access to the results of studies they sponsored have been adopted by many important research funding bodies, including the UK Medical Research Council (MRC), the Wellcome Foundation, the European Research Council, the Canadian Institutes of Health Research, and the National Institutes of Health (NIH) in the USA.[395,396] However, the current policies of open access focus on the openness of results of non-industry-sponsored studies, and it cannot prevent biased publication and reporting of results from industry-sponsored research.

There is no convincing evidence indicating that online open access to published studies increases the number of their citations.[397] A study found that the possibility of an article being found on a non-publisher website was higher for articles published in journals with higher impact factors.[398] Therefore, self-archiving for open access is likely to be selective and may be biased.[399]

Right to publication

Some high-profile cases of publication bias reviewed in Chapter 6 were due to suppression by the industry of publication of negative results from industry-sponsored research. In 1997, a pharmaceutical company changed its policy about dissemination of research it sponsored, allowing investigators 'to publish studies conducted under generally accepted standards of scientific rigour without company prior approval', subject to the 'right to review prepublication drafts to address intellectual property issues'.[400] To prevent publication bias due to industry suppression, Rennie recommended that investigators 'should never allow sponsors veto power',[276] and Rosenberg suggested that scientists should refuse 'to keep information confidential and refuse to sign any agreements for the transfer of information

or reagent that included a requirement of confidentiality.[275] A recently published study found that standards for clinical trial agreements with industry varied considerably among 107 academic medical centres, and subsequent disputes on intellectual property (30%) and control of or access to data (17%) were common.[401]

Research sponsors' guidelines

Funders of clinical research often require investigators of sponsored studies to follow research guidelines. Influential research guidelines include International Conference on Harmonisation (ICH Topic E6), the EU Clinical Trial Directive, the Declaration of Helsinki, and the CONSORT Statement.[41] Dwan *et al.* surveyed research guidelines issued by 17 organisations and 56 charities that funded clinical trials.[41] They found that 11 of these organisations or charities emphasise the publication of both positive and negative results, and three request the adherence to trial protocols in data analysis and that any changes need to be explicitly reported. It was concluded that research funders' guidelines should be improved to prevent selective reporting of outcomes.[41]

Confirmatory large-scale trials

For the purpose of avoiding moderate biases and moderate random errors in assessing or refuting moderate benefits, a large number of patients in randomised controlled trials are required.[402] Large-scale trials are generally believed to be less vulnerable to publication bias; this is the fundamental assumption of many methods for detecting publication bias. When the existence of publication bias is likely and the consequences of such bias are clinically important, a confirmatory, multicentre, large-scale trial may be conducted to provide more convincing evidence. The updated review included no new studies for this section.

Disagreements in results between large-scale trials and corresponding meta-analyses of small trials were observed in empirical studies.[403–405] Although small trials tend to lack statistical power and be more vulnerable to publication bias, the systematic review of small studies may provide useful information about whether a confirmatory large study is required and how to design such a study.[406]

Large-scale confirmatory trials become necessary after a systematic review has reported a clinically significant finding but publication bias cannot be safely excluded as an alternative explanation. Confirmatory large trials are still important even when prospective registries of trials are available. This is because publication bias is only one of many potential threats to trial validity. Compared with a universal register of all trials, confirmatory large trials are more selective about the research areas and objectives, but more flexible at the same time to minimise the impact of other biases, for example, biases related to study design, selection of control, participants and setting.

Summary

The first step for the prevention of publication bias is a wide public awareness of detrimental consequences of publication bias, and the need for the results of all studies to be made accessible. Important actions by government, journals and research sponsors have been taken after several high-profile cases of incomplete disclosure of trial results by pharmaceutical companies were recently brought to light.

Changes in editorial policy, the peer review process, disclosure of commercial interest, electronic publication, trial registration, and open access policy may all help to prevent publication and related biases, although there is as yet little direct evidence as to how well they work in practice. The recent development of electronic publication provides great technical potential to overcome some limitations of conventional printed journals. Publication and related biases may be reduced by electronic publication because of unlimited space, linkage between references, timely publication, and cost-effective dissemination and archiving.

One important solution to publication bias may be the prospective registration of all studies at inception. Voluntary registration of clinical trials is usually incomplete. The policy of compulsory trial registration adopted by the International Committee of Medical Journals in 2004 may be the most influential initiative to promote prospective registration of clinical trials. Further mandatory government regulations may still be required.

The development of prospective trial registration itself is not sufficient for the prevention of publication bias. It is important to make sure that results of registered trials are publically accessible. The usefulness of trial registrations relies on systematic reviewers searching them, using the data they provide and spending time contacting trialists where studies have not yet been published.

Successful efforts so far have focused on biased reporting of phase III and phase IV trials, because of their immediate health consequences. Prospective registration of basic laboratory research, early stage clinical studies and observational studies is still underdeveloped. Open access policy is often mandatory only to public- or charity-supported research. Therefore, although publication bias might be reduced it could still be a problem in many fields of biomedical and health research.

Chapter 8
Reducing or detecting publication and related biases in systematic reviews

Methods that could be useful to reduce or detect publication bias in systematic reviews are discussed in this chapter. Literature searching, locating unpublished trials and assessment of the risk of publication bias will be discussed first. Then methods designed to be used in meta-analysis are discussed, including funnel plot and related statistical methods, fail-safe N, and more sophisticated modelling methods. Finally, the importance of updating systematic reviews is discussed.

Literature searching

If time and resources were unlimited, it is possible that a literature search could identify all published studies relevant to a particular review question. In the real world, a balance must instead be struck between *sensitivity* (the number of relevant studies identified as a proportion of the total number in existence) and *precision* (the number of relevant studies identified as a proportion of the total number retrieved).[2] These two parameters tend to be inversely correlated, such that effort expended on increasing sensitivity is costly in terms of retrieving non-eligible studies, which must subsequently be excluded on an individual basis. Bennett has developed methodology to assess the number of studies potentially missed in a systematic review.[407]

Despite the need to work within a realistic framework, two main issues at this stage of a review can create significant bias in any subsequent analysis: limited exploitation of searching modalities, and low sensitivity within electronic search strategies.

Limited exploitation of searching modalities

Research literature exists in a number of different formats, such as peer-reviewed material published in academic journals, conference abstracts (which may or may not be peer reviewed), other forms of grey literature and personal records.

To carry out an effective systematic review, it is necessary to devise a literature search that can take account of this diversity.[204] Systematic reviewers have a range of searching modalities at their disposal (e.g. electronic bibliographic databases, citation tracking, hand-searching of journals and bibliographies, contacting experts, etc.), but sometimes only one or two are exploited. In particular, search strategies sometimes concentrate almost exclusively on electronic bibliographic databases. This approach may lead to biased searching. For example, studies reporting non-significant results may be overlooked if they tend to be consigned to lower profile journals or other sources that are poorly indexed (or not indexed at all).[408–410] Hand-searching initiatives are one of the few means of addressing the issue of poor indexing.[408,411] Such work may also tend to appear more often as grey literature or other unpublished material (due to publication bias), in which case it may again be overlooked if the search is restricted to standard bibliographic databases.

In addition, reliance on a standard protocol-driven strategy (e.g. prespecified bibliographic database searches supplemented by limited hand-searching) may be associated with the false impression that a search is necessarily exhaustive. Greenhalgh and Peacock recently showed that only 30% of the studies eligible for their review could be obtained from a purely protocol-driven search, compared with the additional inclusions identified through 'snow-balling' (reference and citation tracking) and drawing on personal knowledge (both within and beyond the research team).[412]

Low sensitivity within electronic search strategies

Some degree of bias is clearly possible when a search is based exclusively on bibliographic databases, but in some cases the effects may be too small to influence the results of a systematic review; the power of contemporary search platforms to scan the vast numbers of records held in such databases also underlines the value of electronic searches. However, further bias may

arise if searches are designed without adequate sensitivity. At a simple level, searches are sometimes only applied to a single database (generally MEDLINE), or a limited set of databases, even though individual databases may only have partial coverage of the journals and other sources holding relevant studies.[413] Search results will then necessarily be constrained by the same criteria used to identify sources for inclusion on the databases themselves. The incorporation of excessively specific search terms (e.g. methodological terms) may also tend to reduce (rather than increase) sensitivity if the terms are used to limit (rather than expand) search results. Sensitivity may instead be maximised by favouring relatively generic terms,[414,415] albeit potentially at the cost of some reduction in precision. Similarly, the difficulties noted above in terms of the quality of database indexing[408] necessitate designing electronic strategies that combine text and index terms effectively. Language restrictions (most commonly to English) may also greatly reduce sensitivity where they are applied, depending on the amount of research published in the excluded languages. This may be a further source of bias where study outcome is linked to publication language (particularly biased publication of non-significant results in non-English-language journals): although some evidence indicates the effects on meta-analyses may be small, comprehensive searching without language restriction remains an important safeguard.[153,416,417]

In addressing a particular research question, it is important that systematic reviewers attempt to reduce the potential impact of dissemination bias on subsequent analyses. Accepting that there is a trade-off between search sensitivity and precision,[2] literature searches should therefore draw on as many different searching modalities as are necessary to identify relevant studies. In particular, searches may need to be extended beyond standard bibliographic databases to identify material from, for example, conference abstracts, other forms of grey literature (such as reports by companies, governments or regulatory bodies), and personal or research group data (which may only be disseminated locally).[204,409] Electronic search strategies should also be designed to ensure that the desired level of sensitivity is achieved (thus reducing biases associated with particular databases, search terms, language restrictions and so on), for example by developing sensitive search filters.[418] However, it is clear that reviewers face particular obstacles in attempting to

search *comprehensively*, in relation to both the grey literature, and ongoing/unpublished trials.

Grey literature and non-English-language studies

The ease with which conference abstracts can be identified is not always clear, because research organisations and societies may have different policies on publishing this material as official proceedings (i.e. in society journals).[419] In many cases, conference proceedings are well indexed on databases such as MEDLINE, the Cochrane Library, and possibly SIGLE (System for Information on Grey Literature in Europe), but it may be necessary to determine the best means of identifying abstracts in each individual case before a full search is performed. For example, reviewers could check with relevant societies and organisations to determine how conference abstracts are processed and design search strategies accordingly. Company and regulatory authority reports can be an additional source of useful unpublished data. For example, FDA reviews were used in one study to identify 10 unpublished FDA trial reports.[135] Although assessment of methodological quality was problematic due to poorer reporting, including this unpublished work did not appear to confound effect estimates by introducing 'small study bias'.[420]

Even well-designed and non-language-restricted searches run on MEDLINE and EMBASE may miss a large number of non-English-language studies. One potential solution to this is to search the Cochrane Library. The Cochrane Collaboration has an extensive programme of hand-searching that covers a wide range of journals to ensure that controlled trials from a wide range of sources (including non-English-language journals and conference abstracts) are identified and correctly indexed. Additionally searching of language-specific databases (such as LILACS – Latin American and Caribbean Health Sciences Literature) may be appropriate, but relies on the reviewer having some knowledge of the relevant language(s) to ensure that the correct terms are entered.

Locating ongoing or unpublished trials

Ongoing trials can be defined as any trials that have started but where the results are not yet

available or only interim results are reported, although it is not always straightforward to distinguish between ongoing studies and unpublished completed studies.[421] Ongoing trials should be considered seriously in systematic reviews because of a possible time lag bias or 'pipeline effect' where the speed of publication depends on the direction and strength of the trial results.[113] Large-scale trials often follow early small trials, and results of early published small trials may be overturned by more convincing evidence from later large-scale trials.[403–405] In addition, ongoing trials may be designed particularly to answer important clinical or policy questions that have not been investigated in previous trials.[422] Awareness of ongoing trials will be helpful in making recommendations about when a systematic review needs to be updated and about the need for further research.[423]

Trial registers and grey literature are important sources of information about ongoing trials. A study in 2003 that assessed six commonly used trial registers found that most registers provided sufficient information for reviewers to decide the relevance of identified ongoing trials.[421] However, it is sometimes difficult to know whether ongoing trials identified from different sources (registers) are the same trials or belong to the same multicentre trials, which may be resolved by the wide endorsement of the ISRCTN.

Carefully tailored internet searches[424,425] and email surveys[426,427] may provide useful means of identifying such trials. However, direct communication with trial investigators can be difficult to establish and maintain.[428–430] More targeted approaches to investigators, based on hand-searching of conference abstracts and review articles, may help to ensure that reliable contacts are established.[431] Difficulties can also be encountered in relation to requests for unpublished material from completed but only partially published trials; only 50% of study authors responded to such requests in one study.[432]

A survey in 1993 found that about 31% of published meta-analyses included unpublished data.[116] The unpublished trials were often identified by surveying individuals, organisations or pharmaceutical companies.[228,229,433–437] The number of questionnaires needed to elicit information for each unpublished study ranged from one to five in surveys without restrictions on the study area.

However, if the purpose was to obtain unpublished studies for a meta-analysis, the number of questionnaires needed for one unpublished study was 173 in the study by Shadish and colleagues'.[436] In many meta-analyses, there might be no unpublished trials identified by surveying potential authors, research funding agents and industry. For example, in a systematic review of near patient testing, no unpublished data were obtained by sending questionnaires to 194 academics and 152 commercial companies.[438]

The inclusion of unpublished data may not necessarily reduce the bias in meta-analysis, if the unpublished data are provided by interested sources such as pharmaceutical companies.[139] Unless one can decide whether the identified unpublished trials represent all unpublished trials and decide the proportion of identified unpublished trials in all unpublished trials, the potential publication bias cannot be convincingly solved by locating unpublished trials.

Assessing the risk of publication bias

Some study characteristics were found to be related to the risk of publication bias, such as observational studies, small sample size and small effect size (see Chapter 5). In addition, a comprehensive assessment of study quality is important to detect other potential biases, including selection bias, performance bias, attrition bias and detection bias.[439] A very important consideration is the conflict of interest of research and funding sources, particularly for deciding the possible direction of bias due to selective publication and reporting of results. The risk of bias may be great if all trials are funded by a single body with explicit or implicit reasons for favouring a particular finding. Conversely, when similar results are obtained from trials funded by sponsors with different competing interests, the risk of bias due to funding bodies may be less. In an article regarding false-positive findings from published studies, Ioannidis provided a list of circumstances where a finding is less likely to be true,[440] which may also be associated with a high risk of publication bias. Such situations include small sample size, small effect sizes, a large number of tested relationships, greater flexibility in designs or data analyses, great financial and other interests and prejudices, and hot scientific fields.[440]

Funnel plot and related statistical methods

Because of a larger random error, the results from smaller studies will be more widely spread around the average effect as compared with the results from larger studies. A plot of sample size versus treatment effect from individual studies in a meta-analysis should be shaped like a funnel if there is no publication bias or small-study effects.[441] To help the interpretation of funnel plot asymmetry, Peters *et al.* have recently proposed contour-enhanced funnel plots by adding contour lines indicating conventional milestones in levels of statistical significance (e.g. for $p < 0.01$, $p < 0.05$).[442] A contour-enhanced funnel plot makes it possible to ascertain whether areas where studies are perceived to be missing are in areas of statistical significance and hence whether publication bias is the likely cause of the asymmetry.

If the chance of publication is greater for trials with statistically significant results, the shape of the funnel plot may become skewed. In a funnel plot, the treatment effects from individual studies are often plotted against their standard errors (or the inverse of the standard errors) instead of the corresponding sample sizes (*Figure 11*). The use of standard errors has some advantages because the statistical significance is determined not only by the sample size but also by the level of variation in the outcome measured, or the number of events in the case of categorical data.[443]

Light and Pillemer described two ways in which the shape of the funnel plot can be modified when studies with statistically significant results are more likely to be published.[441] Firstly, assume that the true treatment effect is zero. Then the results of small studies can be statistically significant only when they are far away from zero, either positive or negative. If studies with significant results are published and studies with results around zero are not published, the funnel plot may not be obviously skewed but there will be an empty area around zero (see *Figure 11, 1-B*). These polarised results (significant negative or positive results) may cause many debates; however, the overall estimate obtained by combining all studies is unlikely to be biased.

Secondly, when the true treatment effect is small or moderate but not zero, small studies reporting a small effect size will not be statistically significant and therefore are less likely to be published, while small studies reporting a large effect size may

be statistically significant and more likely to be published. Consequently, there will be a lack of small studies with small effect in the funnel plot, and the funnel plot will be skewed with a larger effect among smaller studies and a smaller effect among larger studies (see *Figure 11, 2-B*). This will result in an overestimation of the treatment effect in a meta-analysis.

The selection of a study for publication may be a function of many variables, such as sample size, level of statistical significance, extent or direction of difference between comparison groups, and design quality of a study. If the publication of a study is associated with the direction of the results, the extent of publication bias may be much greater than when publication is associated with only the level of statistical significance. *Figure 11, 1-C* and *1-D* are the funnel plots in which the selection is a function of statistical significance and sample size when the true treatment effect is zero. *Figure 11, 2-A* to *2-D* show the funnel plots of the results of computer simulation under different selection assumptions when there is a small treatment effect (true odds ratio 0.8).

Limitations of funnel plot for detecting publication bias

For a funnel plot to be useful there needs to be a range of studies with varying sizes. The funnel plot is an informal and subjective method for assessing potential small-study effect; different people may interpret the same plot differently. The visual impression of a funnel plot may change depending on which measure of trial magnitude (SE, variance, sample size, etc.) or which outcome scale (e.g. risk difference, relative risk, odds ratio) is used.[444]

It should be stressed that a skewed funnel plot may be caused by factors other than publication bias. Possible sources of asymmetry include different intensity of intervention, differences in underlying risk, poor methodological design of small studies, inadequate analysis, fraud, choice of effect measure, and chance.[149] Clinical heterogeneity as a source of funnel plot asymmetry can be illustrated using data adopted from the meta-analysis by Hofmeyr *et al.* of calcium supplementation in pregnancy for the prevention of pre-eclampsia.[445] The funnel plot (*Figure 12*) is visually asymmetric, showing a tendency for large trials to be associated with a smaller treatment effect. However, it has been noted that the effect of calcium supplementation was greater for women with a high baseline risk of hypertension (*Figure 13*).[445]
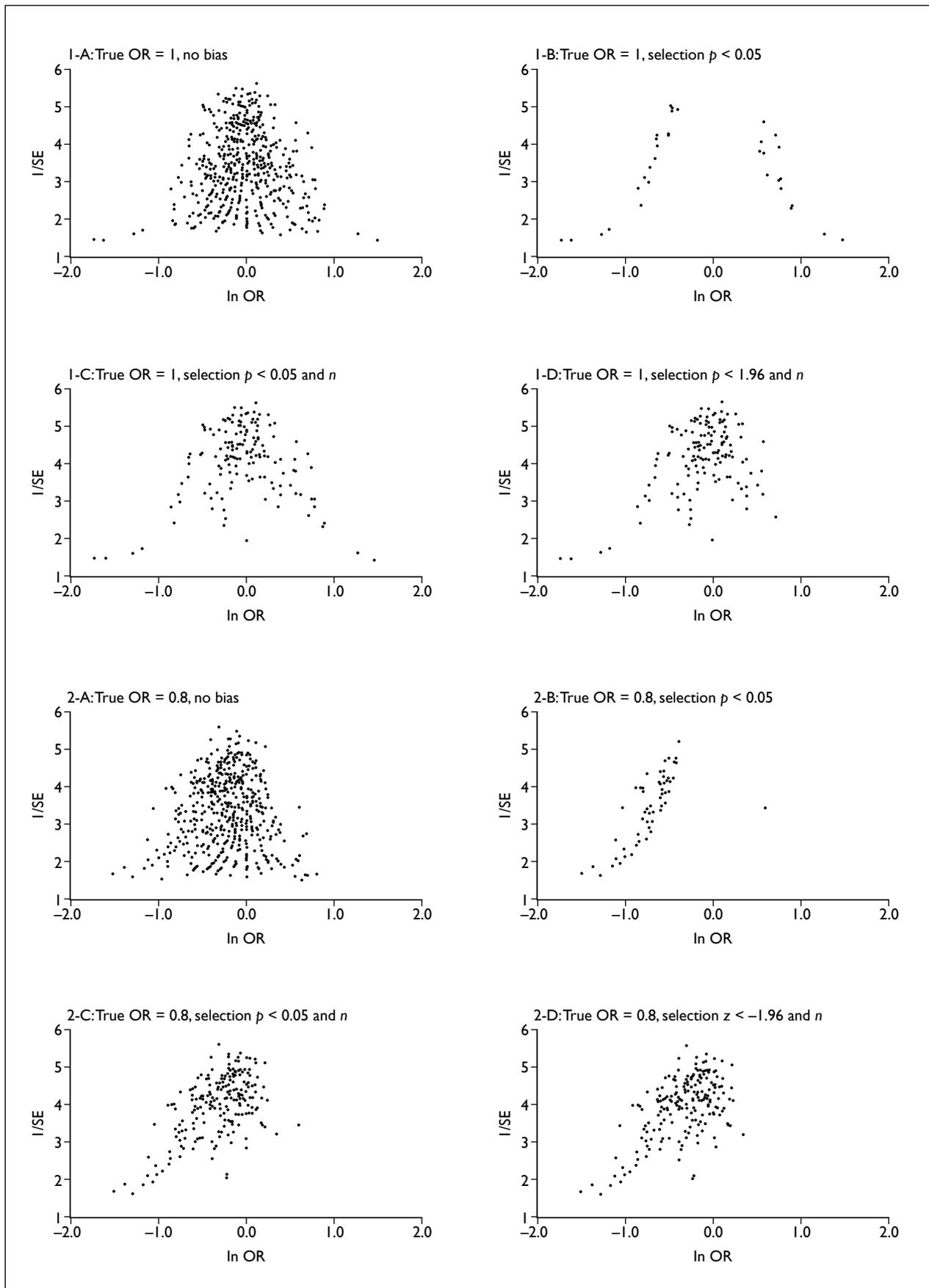
FIGURE 11 *Funnel plots of the results of computer simulations, under different assumptions about biased selection.*
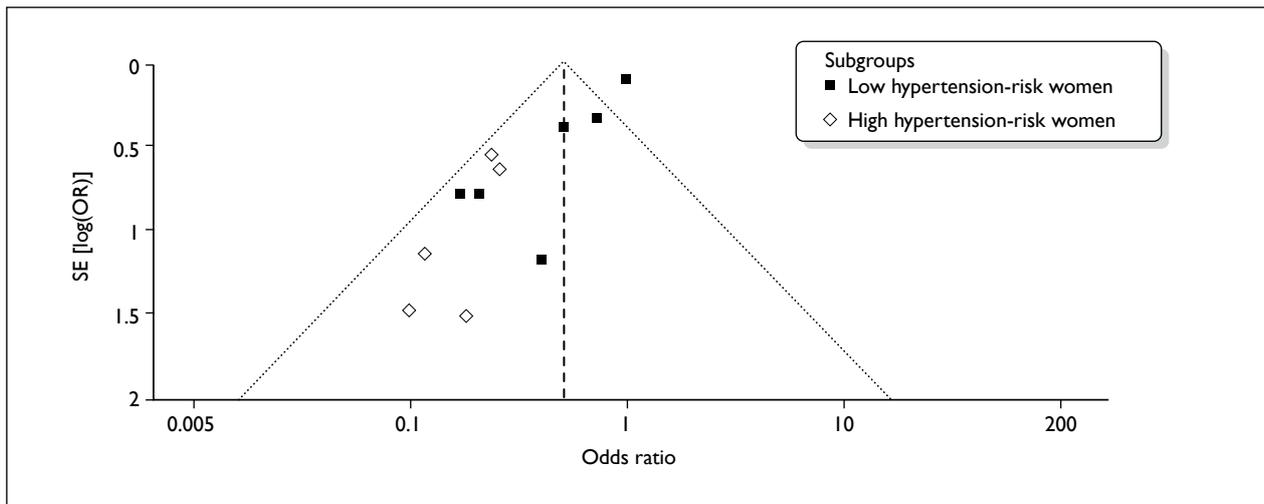
FIGURE 12 *Funnel plot of odds ratio (on log scale) against inverse of standard error. Meta-analysis of calcium supplementation for the prevention of pre-eclampsia. Results of funnel plot asymmetry: Begg's test p = 0.537; Egger's test p < 0.0001; Peters' test: p = 0.002.*

Study or subgroup	Calcium Events	Calcium Total	Piacebo Events	Piacebo Total	Weight	Odds ratio M-H, Random, 95% CI	Odds ratio M-H, Random, 95% CI
1.2.1 Low hypertension-risk women							
Belizan 1991[446]	15	579	23	588	13.5%	0.65 (0.34 to 1.27)	
CPEP 1997[447]	158	2163	168	2173	19.2%	0.94 (0.75 to 1.18)	
Crowther 1999[448]	10	227	23	229	12.1%	0.41 (0.19 to 0.89)	
Lopez-Jaramillo 1989[449]	2	55	12	51	5.4%	0.12 (0.03 to 0.58)	
Purwar 1996[450]	2	97	11	93	5.5%	0.16 (0.03 to 0.73)	
Villar 1987[451]	1	25	3	27	2.8%	0.33 (0.03 to 3.44)	
WHO 2006[452]	171	4151	186	4161	19.3%	0.92 (0.74 to 1.14)	
Subtotal (95% CI)		7297		7322	77.7%	0.65 (0.45 to 0.93)	
Total events	359		426				

Heterogeneity: $\tau^2 = 0.10$; $\chi^2 = 16.61$, df = 6 ($p = 0.01$); $I^2 = 64\%$
Test for overall effect: z = 2.38 ($p = 0.02$)

Study or subgroup	Calcium Events	Calcium Total	Piacebo Events	Piacebo Total	Weight	Odds ratio M-H, Random, 95% CI	Odds ratio M-H, Random, 95% CI
1.2.2 High hypertension-risk women							
Lopez-Jaramillo 1990[453]	0	22	8	34	1.9%	0.07 (0.00 to 1.27)	
Lopez-Jaramillo 1997[454]	4	125	21	135	8.5%	0.18 (0.06 to 0.54)	
Niromanesh 2001[455]	1	15	7	15	2.9%	0.08 (0.01 to 0.79)	
Sanchez-Ramos 1994[456]	4	29	15	34	7.2%	0.20 (0.06 to 0.71)	
Villar 1990[457]	0	90	3	88	1.8%	0.13 (0.01 to 2.65)	
Total (95% CI)		281		306	22.3%	0.16 (0.08 to 0.33)	
Total events	9		54				

Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 0.86$, df = 4 ($p = 0.93$); $I^2 = 0\%$
Test for overall effect: z = 4.94 ($p < 0.00001$)

Study or subgroup	Calcium Events	Calcium Total	Piacebo Events	Piacebo Total	Weight	Odds ratio M-H, Random, 95% CI	Odds ratio M-H, Random, 95% CI
Total (95% CI)		7578		7628	100.0%	0.42 (0.28 to 0.65)	
Total events	368		480				

Heterogeneity: $\tau^2 = 0.23$; $\chi^2 = 37.36$, df = 11 ($p = 0.0001$); $I^2 = 71\%$
Test for overall effect: z = 4.01 ($p < 0.0001$)

0.005 0.1 1 10 200
Favours treatment | Favours control

FIGURE 13 *Calcium supplementation in pregnant women for preventing pre-eclampsia: meta-analysis of 12 randomised controlled trials. Data from Hofmeyr* et al.[444] *Copyright Cochrane Collaboration, reproduced with permission.*

Pregnant women included in smaller trials tended to have a higher baseline risk of hypertension, as compared with those in larger trials. Therefore, the discrepancy in results between smaller and larger trials in this case may be partially explained by the different baseline risk of hypertension and other patient characteristics (dietary calcium intake).[445]

Terrin *et al.* asked 41 medical researchers visually to assess funnel plots of simulated meta-analyses (each included 10 trials).[458] They found that 44% of the funnel plots showed moderate to very high asymmetry when publication bias did not exist, whereas 34% of the funnel plots showed no clear asymmetry even when publication bias did exist. That is, the shape of a funnel plot may be asymmetric purely by chance without selection bias, and a symmetrical funnel plot cannot exclude the existence of publication bias. It was concluded that 'researchers who assess for publication bias using the funnel plot may be misled by its shape'.[458]

Statistical tests for funnel plot asymmetry

It is often difficult for people to decide visually whether a funnel plot is asymmetrical or not. Therefore, some statistical methods have been developed to examine formally the skewness of a funnel plot. The 2000 *HTA* report on publication bias[2] included two methods for testing funnel plot asymmetry: the rank correlation test by Begg and Mazumdar (1994)[340] and a linear regression method by Egger *et al.* (1997).[149] Since then, some new or modified tests have been proposed, all designed to test whether studies with smaller sample size or greater variation in results tend to report greater treatment effects in meta-analysis (*Table 2*).[8,9,149,340,444,459–462] Almost simultaneously, many recent studies have been conducted to compare the performance of these tests.[8,9,459–468] Different tests often lead to different conclusions in terms of the funnel plot asymmetry. All the proposed tests have some important limitations, including low statistical power to identify funnel plot asymmetry when it exists, and inflated rates of type I error when funnel plot asymmetry does not exist. The performance of tests for funnel plot asymmetry is particularly poor when heterogeneity in meta-analysis is large.[469]

Ioannidis and Trikalinos (2007)[470] suggested that statistical tests for funnel plot asymmetry may be appropriate only if the following four criteria are met:

- no significant heterogeneity
- $I^2 < 50\%$
- 10 or more studies (with statistically significant results in at least one study)
- ratio of the maximal to minimal variance across studies > 4.

A survey of 846 independent meta-analyses from the *Cochrane Database of Systematic Reviews* found that only 12% of the meta-analyses met all the above four criteria.[470] The number of studies was fewer than 10 in 74% of the meta-analyses and none of the studies was statistically significant in 34% of the meta-analyses. About 30% of the meta-analyses had statistically significant heterogeneity.[470] However, it should be noted that the above criteria were not based on convincing empirical evidence and further simulation studies may help to investigate how valid these criteria are.

In the recently updated *Cochrane Handbook for Systematic Reviews of Interventions*,[469] Sterne *et al.* have provided some recommendations about the use of statistical tests for funnel plot asymmetry. The main points of their recommendations are summarised below.

- The tests for funnel plot asymmetry should not be used in meta-analyses that include less than 10 studies.
- To test funnel plot asymmetry, studies included in a meta-analysis should have different sizes, although it is not clear when the difference in study sizes are sufficient.
- Egger's test can be used for continuous outcomes measured by mean differences.
- For dichotomous outcomes measured by odds ratio, tests proposed by Harbord *et al.*,[459] Peters *et al.*[9] or Rucker *et al.*[461] can be used in the absence of significant heterogeneity ($\iota^2 < 0.1$).
- Arcsine random-effect regression test[461] should be used when both treatment effect and heterogeneity are large (e.g. $\iota^2 > 0.1$).
- Funnel plot testing strategy should be specified in advance, and only one test should be used.

Other tests for funnel plot asymmetry included in *Table 2* are not recommended in the updated *Cochrane Handbook for Systematic Reviews*[469] mainly because of low power[8,340,460,462] or lack of statistical evaluation.[444] There is very limited empirical evidence on the performance of the available tests for dichotomous outcomes measured as risk ratios or risk differences, and for continuous outcomes measured by standardised mean differences.[469]

Trim and fill method

The trim and fill method is a rank-based data augmentation technique, designed to estimate the number of missing studies and to provide an estimate of the treatment effect by adjusting for funnel plot asymmetry in a meta-analysis.[471,472] Briefly, the asymmetrical outlying part of the funnel is firstly 'trimmed off' after estimating how many studies are in the asymmetrical part. Then the symmetrical remainder is used to estimate the 'true' centre of the funnel. Finally, the 'true' mean and its variance are estimated based on the 'filled' funnel plot in which the trimmed studies and their missing 'counterparts' symmetrical about the centre are replaced.[471]

An early simulation study reported that the trim and fill method estimates the point estimate of the overall effect size approximately correctly and the coverage of the confidence interval is substantially improved, compared with ignoring publication bias.[471] However, further simulation studies found that the trim and fill method may perform poorly with high false-positive findings when heterogeneity in meta-analysis is large.[473,474] Because of the existence of clinical heterogeneity, as with all other funnel plot-based tests, the trim and fill method may provide a misleading estimate by spuriously adjusting for bias that actually does not exist.

Other statistical and modelling methods

Fail-safe N methods

Several methods have been proposed to estimate the number of unpublished studies in a meta-analysis.[11,475–479] The first and most commonly used is Rosenthal's fail-safe N, designed to estimate the number of unpublished studies required, with zero effect on average ($z = 0$), to overturn a significant result ($p < 0.05$) in a meta-analysis.[11] If the number of unpublished studies with null results required to overturn the statistically significant result is large, and therefore unlikely to exist, the impact of publication bias is considered to be ignorable and thus the results obtained from published studies are held to be robust.

The plausible number of unpublished studies may be hundreds in some areas or only a few in others. Therefore, the estimated fail-safe N should be considered in proportion to the number of published studies (K). Rosenthal suggested that the fail-safe N may be considered as being unlikely if it is greater than a tolerance level of '$5K + 10$'.[11]

There are problems with the fail-safe N method.[480] Firstly, the method overemphasises the importance of statistical significance. Secondly, it may be misleading when the unpublished studies have an average effect that is in the opposite direction to the observed meta-analysis. If the unpublished studies reported contrary results compared with those in the published studies, the number of unpublished studies required to overturn a significant result would be smaller than that estimated, assuming an average effect of zero in unpublished studies. In addition, the interpretation of estimated fail-safe N may be misleading because it is often difficult to decide the plausible number of unpublished studies. Becker has suggested that 'the failsafe N should be abandoned in favour of other more informative analyses.'[481]

Recently, a weight function method of sensitivity analysis proposed by Copas and Jackson[482] has been used to estimate a range of possible numbers of unpublished studies in a meta-analysis. This method is discussed below.

Sophisticated modelling methods

The impact of missing studies may also be assessed by using more sophisticated modelling methods. Many of the sophisticated modelling methods were discussed in depth in the 2000 *HTA* report[2] and in a review article by Sutton *et al.*[483] These methods are usually based on weighted distribution theory derived from both classical[338,484–492] or Bayesian[493–497] perspectives. There are two aspects to the selection models that use weighted distribution theory: an effect size model, which specifies what the distribution of the effect size estimate would be if there were no selection, and the selection model, which specifies how this effect size distribution is modified by the selection process.[498] In some methods the nature of the selection process is predefined by the researcher, while in others it is dictated by the available data.

The appropriate application of modelling methods to test publication bias usually requires a large number of studies (e.g. 100 or more) in a meta-analysis.[499] However, the number of studies was fewer than 10 in most published meta-analyses.[470] In addition, it is difficult if not impossible empirically to verify the validity of assumed selection processes, since the true mechanisms and extent of biased publication or reporting are usually unknown.[482] Therefore, it has been

recognised that weight function models should be used to conduct sensitivity analyses, rather than to provide a single 'correct' estimate by adjusting for the assumed selection bias.[482,499]

A sensitivity analysis method using weight function for assessing the impact of publication bias has been proposed by Copas *et al.*[482,489,500] The probability of study selection is assumed to be associated with estimated effect sizes and corresponding standard errors. Then a range of plausible values for inestimable parameters can be tested using the model to provide a range of corresponding estimates on the size of bias or the number of unpublished studies, which can be used to indicate the possible impact of selection bias under different assumptions.[482]

Vevea and Woods recently proposed a new weight function model for the purpose of sensitivity analysis.[499] Similar to the sensitivity analysis approach proposed by Copas,[482] Vevea's new model can be used to conduct sensitivity analyses using a range of assumed weight function parameters (such as moderate or severe, one- or two-tailed selection), rather than to provide 'a best guess' at the true effect size.[499]

Many proposed modelling methods require a large number of studies and therefore may not be appropriate for use in typical meta-analyses. Sensitivity analysis approaches proposed by Copas *et al.*[482,489,500] and by Vevea and Woods[499] could be used even when the number of studies is not large. Unfortunately, the complexity of these methods means that they have mostly been used only by statistical experts in method studies. Further development of user-friendly software is required to bring the methods into more mainstream use.

Methods for detecting outcome reporting bias

The existence of outcome reporting bias is suspected when some eligible studies could not be included in a meta-analysis due to a lack of data on an outcome. Outcome reporting bias is not a problem for studies that had not measured the outcome of interest. Therefore, it is crucial, although difficult, to investigate whether the outcome was measured or not in studies in which the outcome has not been reported. Published studies can compared with their protocols when available, or authors of published studies may be contacted to request clarifications about or data on the unreported outcomes.[439]

Williamson and Gamble proposed a method to investigate the possible impact of outcome reporting bias by imputing data for unreported outcomes.[95] More recently, they compared their imputation method and the sensitivity analysis method proposed by Copas and Jackson[482] in the assessment of outcome reporting bias.[501] Results of simulation indicate that outcome reporting bias may be overadjusted by using the imputation method as compared to the Copas method.[501]

Other statistical methods

Bennett *et al.* proposed that a capture-recapture method may be used to assess the risk of publication bias.[407] The performance of the capture-recapture method has not been properly investigated by simulations.

Ioannidis and Trikalinos proposed a test for an excess of significant findings.[502] The exploratory test can be used to estimate the number of the expected studies with statistically significant results according to certain assumptions. Then the number of expected significant studies is compared with the number of observed studies with statistically significant results.[502] Publication and related bias may be one of the reasons for the excess of significant studies.

Fixed or random-effects models

In meta-analysis, larger studies will give greater weight than smaller studies when results are quantitatively combined.[503] This procedure may have an advantage in reducing the impact of publication bias because less weight is given to smaller studies that are more vulnerable to publication bias.

There are two statistical models that can be used to combine results of individual studies in a meta-analysis: the fixed-effects model or the random-effects model.[504] In the fixed-effect model it is assumed that all individual studies are measuring a single value of the true effect and that the observed difference between the studies is due to sampling error. The precision (e.g. the inverse of within-study variance) of individual results is employed as the weight for each study to estimate the pooled result in meta-analysis using the fixed-effect model. By contrast, the random-effects model assumes that individual studies are measuring a distribution of effects. In addition to the variation within studies,

the variation between studies is also incorporated into a meta-analysis using the random-effects model. The differences between the fixed-effects model and the random-effects model are often ignorable when heterogeneity is small.[505] When there is large heterogeneity between individual studies, the confidence interval estimated by using the random-effects model will be wider than that by using the fixed-effect model.

Jackson investigated the impact of publication bias on estimates of between-study variance (ι^2-statistic) in meta-analysis.[506] The results of mathematical analysis demonstrate that publication bias may increase or decrease the between-study variance in meta-analysis.

The weights used to combine individual studies are the inverse of within-study variances in the fixed-effect model, and are the inverse of total variance (i.e. within-study variance plus between-study variance) in the random-effects model. Therefore, by giving relatively larger weights to smaller studies, the random-effects model may be more vulnerable to publication bias than the fixed-effect model.[507] For this reason, it has been recommended that the result of the random-effects model should be compared with the result of the fixed-effect model when there is heterogeneity in meta-analysis.[469] If the pooled effect size by the random-effects model is greater than that by the fixed-effect model, underlying causes will need to be investigated, to exclude the possibility of publication bias.

Updating systematic reviews

Updating of systematic reviews may reduce the impact of publication bias because of the following reasons. First, publication of studies with negative or less favourable results may have a longer delay than studies with positive or favourable results.[113] Secondly, large-scale confirmatory trials are usually conducted and published after the publication of early small trials. Small trials may be more vulnerable to biased selection for publication, and the conclusions based on limited evidence from early small trials may be overturned by more convincing evidence from later large-scale trials.[403–405]

Jadad *et al.*[508] compared 36 Cochrane reviews with 39 meta-analyses published in paper-based journals (randomly selected sample) published

in 1995. They found that, within 2 years after publication, 18 of the 36 Cochrane reviews had been updated versus only one of the 39 reviews published in paper-based journals. Possible reasons given for a very low update rate among paper-based reviews included editors of such journals not being sufficiently interested in publishing updated versions of previously published systematic reviews, authors not being aware of such interest, or authors lacking the interest or resources to update previously published reviews.[508]

French *et al.* examined the effect of updating Cochrane systematic reviews from 1998 to 2002.[509] They found that 137 (38%) of the 362 completed reviews published in 1998 were updated and had included new studies by 2002. Among the 119 reviews that included new studies with comparable results, statistical significance of the primary outcome was changed from significant ($p < 0.05$) to non-significant ($p > 0.05$) in five and from non-significant to significant in six reviews. There was no mention of any changes in the direction of the estimates of treatment effects. French *et al.* concluded that 'a priority-setting approach to the updating of Cochrane systematic reviews may be more appropriate than a time-based approach'.[509]

A recent study by Shojania *et al.* used 100 randomly selected systematic reviews of conventional therapy published from 1995 to 2005 to investigate important changes in evidence by updating.[510] They defined quantitative signals for updating as any 'changes in statistical significance or relative changes in effect magnitude of at least 50%', and qualitative signals for updating as 'new information about harm and caveats about the previously reported findings that would affect clinical decision making'. Important changes in evidence were observed in 57% of the reviews. Updating was required for important changes in evidence by 15% of reviews within 1 year after publication, and by 23% of reviews within 2 years after their publication. Multivariate analysis suggested that systematic reviews with cardiovascular topics and heterogeneity may need to be frequently updated.[510]

Shojania *et al.*[510] also reported details about important changes in evidence for seven selected systematic reviews. Compared with findings from the original seven reviews, important changes in evidence suggested that treatments of interest were less beneficial in four cases, and more beneficial in another two cases, and less harmful in one.[510] This limited evidence indicates that estimated treatment

TABLE 2 *Alternative statistical tests for funnel plot asymmetry*

Study (year)	Tests	Comments
Begg and Mazumdar (1994)[340]	Rank correlation test of the association between standardised effect size and its variance	It suffers from a lack of power and the possibility of funnel plot asymmetry cannot be ruled out when the test is non-significant, particularly when the number of studies is small
Egger et al. (1997)[149]	The method is based on a regression analysis of Galbraith's radial plot.[515] The standard normal deviate (SND) is defined as the ln OR divided by its standard error (SE). The SND is then regressed against the estimate's precision (the inverse of the SE), weighted by the inverse of the variance. The intercept of the regression line provides a measure of funnel plot asymmetry	The test is unbiased for continuous outcomes. For binary data, Egger's test is biased due to the intrinsic association of the SND and its precision.[516] Egger's test is more powerful than Begg's rank correlation test, but has high false-positive rates, particularly when the treatment effect and/or the number of studies is large[464]
Tang and Liu (2000)[444]	A sample size-based linear regression in which the estimate of the effect size is regressed against the square root of the average number of participants in the two trial groups, weighted by the sample size	Tang and Liu (2000) mentioned that the interpretation of the intercept alpha and its p value is the same as that of Egger's method. The power and type I error of the method have not been properly investigated
Macaskill et al. 2001[460]	Linear regression of the estimated treatment effects (dependent variable) and corresponding sample size (Nt), weighted by the inverse of the variance of the logit of the pooled proportion (using the marginal total)	The correlation between the weight and treatment effect is reduced. The statistical power of the test is low
Deeks et al. (2005)[8]	Linear regression of ln OR against $1/(ESS)^{1/2}$, weighting by ESS, where effective sample size $ESS = 4 \cdot N0 \cdot N1/Nt$	The test is developed for the evaluation of diagnostic odds ratio (DOR). The power is likely to be low, particularly when heterogeneity across a study is large
Harbord et al. (2006)[459]	Modified Egger's test: a regression of $Z/\sqrt{V}$ against $\sqrt{V}$, where Z is the efficient score, and V the score variance	With large heterogeneity (e.g. $\iota^2 > 0.1$), the test has the problems of high false-positive rate and low power similar to Egger's and Macaskill's methods
Peters et al. (2006)[9]	Linear regression of estimated treatment effect against $1/\sqrt{Nt}$; weight used as in Macaskill's test	The test is superior to Egger's test in terms of more appropriate type I error rates. As with other statistical tests for funnel plot asymmetry, the statistical power is low when heterogeneity is large (e.g. $\iota^2 > 0.1$)
Schwarzer et al. (2007)[462]	Rank correlation test, based on observed and expected cell frequencies, and the variance of the observed cell frequencies, in 2×2 tables	The test is developed for meta-analysis with sparse binary data. The power of the test is low
Rucker et al. (2008)[461]	The test is based on the arcsine transformation to stabilise the variance of binomial random variables. Then arcsine transformed statistics can be used to replace variables used in Begg's test, Egger's test, or a random-effects regression analysis	Compared with other tests, arcsine transformed random-effects regression has improved power when both effect size and heterogeneity are large. The test is relatively conservative with small sample size and in the absence of heterogeneity

effects may be reduced or increased by updating systematic reviews.

Updating a meta-analysis involves repetitive statistical testing and the risk of type I errors will increase.[511] Under certain circumstances, type I errors due to repetitive tests for meta-analysis updating may be greater than publication bias.[512] To adjust for random error risk, Brok *et al.* recommended the use of trial sequential analysis (TSA) in meta-analysis.[511] However, possible type II errors and other biases are also important and the risk of type I error should not be a reason for not updating systematic reviews.

Summary

Many methods have been suggested for preventing, testing or adjusting for publication bias. The available methods could be classified as methods for preventing publication bias (discussed in Chapter 6) and methods for dealing with publication bias in systematic reviews (discussed in this chapter). In addition, it is possible to classify the available methods according to the stage of a literature review: to prevent publication bias before a literature review (e.g. prospective registration of trials), to detect publication bias during a literature review (e.g. locating unpublished studies, funnel plot and related tests, sensitivity analysis modelling), or to minimise the impact of publication bias after a literature review (e.g. confirmatory large-scale trials, updating the systematic review).

The recent development of clinical trial registration and electronic publication of results from clinical trials will facilitate the identification and location of ongoing or unpublished clinical trials. However, for non-trial studies, including basic laboratory research, epidemiological studies and even early stage clinical studies, publication and result reporting bias seems still at large as before.

Funnel plot and related statistical tests have been widely used in systematic reviews to assess the possibility of publication bias. Unfortunately, the interpretation of results of funnel plot-related tests was often too simplistic and likely misleading.[513,514] Therefore, detailed recommendations have been recently proposed about when and how to use the funnel plot-related statistical tests in meta-analysis,[469,470] which may facilitate more cautious interpretation of funnel plot asymmetry. However, the current recommendations about tests for funnel plot asymmetry are based on very limited and fast changing empirical evidence, and they may have to be revised when new evidence emerges.

Many sophisticated modelling methods have not been widely used in systematic reviews, possibly because of their complexity and lack of user-friendly software. The main development of sophisticated modelling methods perhaps is the more general recognition that these methods should be used to conduct sensitivity analyses, rather than to provide an estimate of the 'true' effect size by adjusting for assumed selection bias. However, it is unclear how useful such sensitivity analyses are when the results of meta-analyses are used for decision-making in practice.

We concluded previously in the 2000 *HTA* report that all statistical methods 'are by nature indirect and exploratory, and often based on certain strict assumptions that can be difficult to justify in the real world'; and 'the attempt at identifying or adjusting for publication bias in a systematic review should be mainly used for the purpose of sensitivity analyses'.[2] The updated review indicates that the above conclusions are still held.

Chapter 9
Survey of published systematic reviews

In our previous *Health Technology Assessment* (*HTA*) report,[2] 193 systematic reviews from the Database of Abstracts of Reviews of Effectiveness (DARE) were used to identify any evidence of dissemination bias and to illustrate the methods used in systematic reviews for dealing with publication bias. However, the systematic reviews included in DARE were probably, on average, of higher quality than those from general bibliographic databases and hence the representativeness of the reviews assessed was questionable. In addition, reviews on effectiveness of health-care interventions and accuracy of diagnostic technologies were not assessed separately. The problem of dissemination bias might be different between the two types of systematic reviews. In the current updated review we have obtained a sample of systematic reviews from a general bibliographic database (MEDLINE) and classified these reviews into the following categories: (1) systematic reviews of studies on effects of health-care interventions, (2) systematic reviews of studies on accuracy of diagnostic tests, (3) systematic reviews of epidemiological studies on association between risk factors and health outcomes, and (4) systematic reviews of genetic studies on association between genes and disease. We also assessed a sample of systematic reviews that explicitly discussed publication bias.

Assessment of randomly selected reviews

We searched MEDLINE for systematic reviews published in 2006 (see Chapter 2 for methods) and randomly selected 100 systematic reviews of studies of effectiveness of interventions, 50 reviews of studies of diagnostic accuracy, 100 reviews of epidemiological studies on risk factors and health outcomes, and 50 reviews of studies of gene-disease associations (Appendix 17). The reviews were assessed independently by two reviewers using a data extraction form, tailored to the type of review being assessed (Appendix 4).

Main characteristics of included reviews

The 100 treatment effectiveness reviews comprised 54 reviews of pharmaceutical interventions, 10 reviews of psycho-educational interventions, 11 reviews of surgical interventions, and 25 reviews of mixed or other interventions. The median number of individual studies included in each review was 14 (range 2 to 198).

The tests or techniques investigated in the 50 reviews of diagnostic accuracy included laboratory tests ($n = 21$), imaging techniques ($n = 28$), electrical tests ($n = 5$), clinical examinations ($n = 10$) and other tests ($n = 7$) (several reviews assessed more than one test or technique). The median number of studies included in the 50 reviews was 20 (range 4 to 213).

Risk factors investigated in 100 reviews of epidemiological studies included lifestyle ($n = 31$), environmental ($n = 17$), biomedical ($n = 45$), mental ($n = 6$) and other factors ($n = 18$). Cancer ($n = 24$) and cardiovascular diseases ($n = 20$) were common health outcomes considered in these reviews. The median number of studies included in each review was 20 (range 3 to 200).

In 50 reviews of gene-disease association, diseases investigated included cancer ($n = 13$), cardiovascular disease ($n = 4$), diabetes ($n = 3$) and mental diseases ($n = 15$). The median number of studies per review was 13 (range 3 to 86).

Among the 300 systematic reviews 83 were narrative systematic reviews in which the results of the primary studies were summarised but not statistically combined, and 217 were meta-analyses, in which statistical methods were used to combine the results of two or more primary studies (*Table 3*). There were 16 (16%) Cochrane reviews amongst the 100 treatment reviews.

Systematic literature search

Similar to the findings reported in the 2000 *HTA* report, MEDLINE (74% to 95%) and checking

TABLE 3 *Types of reviews – narrative or meta-analyses*

	Narratives (%)	**Meta-analyses (%)**	**Total (%)**
Treatment	40 (40%)	60 (60%)	100
Diagnostic	9 (18%)	41 (82%)	50
Risk factor	32 (32%)	68 (68%)	100
Genetic	2 (4%)	48 (96%)	50
Total	83 (28%)	217 (72%)	300

reference lists of retrieved studies (42% to 73%) were most commonly used to search literature (*Figure 14*). EMBASE was now searched in about half of the treatment and diagnostic reviews (50% and 54%, respectively), compared with only 17% in the previous reviews. There was a considerable increase in the utilisation of the Cochrane Library, from only 5% in reviews published in 1996 to 58% in treatment reviews, 46% in diagnostic reviews, 24% in risk-factor reviews and 6% in genetic reviews. The search of the CINAHL (Cumulative Index to Nursing and Allied Health Literature) database increased from 8% in 1996 to 24% in treatment reviews, 20% in diagnostic reviews, and 18% in risk factor reviews in 2006.

Prospective registers of clinical trials and other studies were searched for unpublished or ongoing studies in 18 treatment reviews, three diagnostic reviews and two risk-factor reviews (*Table 4*). The UK National Research Register and ClinicalTrials.gov are the two commonly searched prospective registers.

Language restriction

The current review examined whether language restrictions (e.g. included studies were limited to English-language ones only) were applied by the review authors. It was found that 35% of the reviews showed language restriction, with the majority of them being restricted to English language. Thirty-five percent of the reviews did not explicitly report whether any language restrictions were applied or not. The proportion of reviews that explicitly stated no language restriction was 39% in treatment reviews, 42% in diagnostic reviews, and 20% in risk-factor reviews and genetic reviews (*Table 5*).

Non-English-language studies were explicitly searched for in 45% treatment reviews, 52% diagnostic reviews, 34% risk-factor reviews and 22% genetic reviews. Overall, 39% of the reviews explicitly searched for non-English-language studies, compared with 19% in the reviews published in 1996. However, only 15% of the reviews actually included non-English-language studies (*Table 5*). Authors did not always explicitly mention that they had searched for non-English-language studies, even though non-English studies were indeed listed in the review.

Treatment reviews and diagnostic reviews were more likely to have no language restrictions, and more frequently searched for or included non-English-language literature compared with

TABLE 4 *Searching of prospective registers of trials or other studies in reviews*

Study register	Treatment review	Diagnostic review	Risk-factor review	Genetic review
Physician Data Query clinical trial register	2			
UK National Research Register	8	3	2	
ClinicalTrials.gov	8		1	
Current Controlled Trials	3	1	1	
Other/unclear	5			
Number of reviews searched study registers	18/100	3/50	2/100	0/50

Sources searched	Treatment n = 100 (%)	Diagnostic n = 50 (%)	Risk factors n = 100 (%)	Genetic n = 50 (%)
MEDLINE	95 (95)	47 (94)	95 (95)	37 (74)
EMBASE	50 (50)	27 (54)	37 (37)	7 (14)
PsycINFO	11 (11)	2 (4)	23 (23)	2 (4)
Cochrane	58 (58)	23 (46)	24 (24)	3 (6)
References	71 (71)	42 (42)	73 (73)	23 (46)
CINAHL	24 (24)	10 (20)	18 (18)	0 (0)
Handsearch	30 (30)	17 (34)	5 (5)	1 (2)
Experts/Authors	33 (33)	18 (36)	20 (20)	2 (4)
Company	19 (19)	0 (0)	1 (1)	0 (0)
Proceedings	35 (35)	8 (16)	11 (11)	3 (6)
SIGLE	0 (0)	4 (8)	1 (1)	0 (0)
Other	60 (60)	18 (36)	50 (50)	9 (18)

FIGURE 14 *Databases searched for literature in 300 reviews.*

risk-factor and genetic reviews. The proportion of reviews in which non-English-language studies were explicitly searched for or included was 47% in treatment reviews, 56% in diagnostic reviews, 39% in risk-factor reviews and 28% in genetic reviews (*Table 5*). The current findings indicate an improvement compared with previous findings, where only about 30% of the reviews searched for or included non-English-language studies.[2]

Grey literature and unpublished studies

The Third International Conference on Grey Literature has defined grey literature as 'that which is produced on all levels of governmental, academic, business and industry in print and electronic formats, but which is not controlled by commercial publishers'.[114] In this review, we have attempted to separate grey literature and other unpublished studies. The commonly used

TABLE 5 *Language restriction in 300 reviews*

Language restriction	Treatment *n*=100 (%)	Diagnostic *n*=50 (%)	Risk factor *n*=100 (%)	Genetic *n*=50 (%)
No restriction	39 (39%)	21 (42%)	20 (20%)	10 (20%)
Restricted to English	23 (23%)	15 (30%)	31 (31%)	11 (22%)
Restricted to two or more languages	6 (6%)	4 (8%)	14 (14%)	1 (2%)
Unclear	32 (32%)	10 (20%)	35 (35%)	28 (56%)
Non-English-language studies searched for	45 (45%)	26 (52%)	34 (34%)	11 (22%)
Non-English-language studies included	14 (14%)	14 (28%)	10 (10%)	6 (12%)
Non-English-language studies searched for or included	47 (47%)	28 (56%)	39 (39%)	14 (28%)

methods to identify grey literature were searching conference abstracts, meeting proceedings and grey literature-specific databases like SIGLE and LILACS. Checking the reference list of the reviews indicates that conference abstracts were frequently included. Grey literature was explicitly sought in 50% of treatment reviews, 30% of diagnostic reviews, 32% of risk-factor reviews, and only 8% of genetic reviews (*Table 6*). Overall, 34% of the 300 reviews explicitly searched for grey literature, although only 13% included them.

To identify other unpublished studies, the commonly used method was through contacting authors or experts, and pharmaceutical companies. Of the 300 reviews, 27% explicitly searched for other unpublished studies and only 8% actually included them (*Table 6*). When grey literature and other unpublished studies were combined, the proportion of reviews that explicitly searched for grey or unpublished studies was 58% for treatment reviews, 36% for diagnostic reviews, 35% for risk-factor reviews and 10% for genetic reviews (*Table 6*). In addition, *Table 6* also shows that grey literature and unpublished studies were more likely to be included in treatment reviews than diagnostic or risk-factor reviews.

The previous *HTA* report showed that only about 35% of reviews explicitly searched for or included studies that were unpublished or presented as abstracts.[2] In reviews published in 2006, this was 61% for treatment reviews, 36% for diagnostic reviews, 41% for risk-factor reviews and 20% for genetic reviews (*Table 6*).

TABLE 6 *Grey literature and unpublished studies in reviews*

	Treatment *n*=100 (%)	Diagnostic *n*=50 (%)	Risk factor *n*=100 (%)	Genetic *n*=50 (%)
Grey literature				
Searched for	50 (50%)	15 (30%)	32 (32%)	4 (8%)
Included	17 (17%)	5 (10%)	12 (12%)	5 (10%)
Other unpublished studies				
Searched for	49 (49%)	7 (14%)	20 (20%)	5 (10%)
Included	14 (14%)	2 (4%)	2 (2%)	6 (12%)
Grey literature or unpublished studies				
Searched for	58 (58%)	18 (36%)	35 (35%)	5 (10%)
Included	24 (24%)	6 (12%)	12 (12%)	8 (16%)
Searched for or included	61 (61%)	18 (36%)	41 (41%)	10 (20%)

Consideration of outcome reporting bias

Outcome reporting bias is related to the incomplete reporting within published studies and occurs when studies with multiple outcomes selectively report only some of the measured outcomes. We examined whether outcome reporting bias was considered and/or reported in our sample of 300 reviews. We found that outcome reporting bias was explicitly mentioned in 18% of treatment reviews, 14% of diagnostic reviews, 3% of risk-factor reviews and 16% of genetic reviews.

Methods used to test for publication bias

Available tests for publication bias were not used in the majority of treatment reviews (79%), diagnostic reviews (76%) and risk-factor reviews (69%) (*Table 7*). Compared with other reviews, publication bias was more likely to be tested in genetic reviews, possibly due to perceived high risk of bias in such reviews. The most commonly used methods for testing the association between sample sizes and treatment effects were funnel plots complemented by other related methods (Egger's and Begg's test). Egger's test was explicitly used in 45 reviews and Begg's test in 24 reviews. Funnel plot and other related methods were used in 26% of the 300 reviews, compared with less than 6% in the 193 reviews published in 1996. In contrast to reviews published in 1996, the fail-safe N method was used in far fewer reviews (7% versus 1%). All other statistical methods to test publication bias were only rarely used. In the reviews that explicitly tested for publication bias, 23% of the 21 treatment reviews, 42% of the 12 diagnostic reviews, 52% in the 31 risk-factor reviews, and 48% in the 27 genetic reviews showed some evidence of the existence or absence of publication bias.

Consideration of publication bias

In accordance with the findings of the previous report,[2] publication bias was discussed or mentioned more often in the meta-analyses than in the narrative reviews (*Table 8*). The possibility of potential publication bias was discussed more often in the genetic reviews (70%) than in treatment reviews (32%), diagnostic reviews (48%) and risk-factor reviews (42%) (*Table 8*).

When conclusions of authors of reviews were classified as positive, not positive and unclear, we found positive conclusions in 61% of treatment reviews, 35% of diagnostic reviews, 91% of risk-factor reviews and 32% of genetic reviews. Because of the small number of available reviews, it is not clear whether authors' conclusions were associated with whether or not publication bias was explicitly tested or considered in a review (*Table 8*). Similarly, there was no clear trend to show that publication bias was more or less likely to be explicitly tested or discussed in meta-analyses that reported statistically significant results (*Table 8*).

Assessors' judgement

Efforts taken to minimise publication bias

As part of our study, we used two assessors to assess independently the efforts taken by review authors to minimise publication bias within the selected sample of reviews. The assessors' judgements were based on the following measures: literature searching approach used, consideration of outcome reporting bias, reporting of any missing outcomes and methods used to deal with the missing outcomes, and discussion of publication bias and any methods used to deal with publication bias. For each review, two assessors independently scored

TABLE 7 *Methods used to deal with publication bias in reviews*

	Treatment	Diagnostic	Risk factor	Genetic
Not used	79 (79%)	38 (76%)	69 (69%)	23 (46%)
Funnel plot and related methods	15 (15%)	9 (18%)	27 (27%)	26 (52%)
Egger's test	9 (9%)	2 (4%)	16 (16%)	18 (36%)
Begg's test	5 (5%)	0	8 (8%)	11 (22%)
Trim-fill method	0	0	2 (2%)	0
Fail-safe N	2 (2%)	0	1 (1%)	0
Modelling	0	0	0	1 (2%)
Other	2 (2%)	1 (2%)	5 (5%)	5 (10%)

TABLE 8 *Consideration of publication bias and results or conclusions of reviews*

	Treatment reviews			Diagnostic reviews			Risk factor reviews			Genetic reviews		
	Total	Tested %	Discussed %	Total	Tested %	Discussed %	Total	Tested %	Discussed %	Total	Tested %	Discussed %
Authors' conclusions												
Positive	61	26.2%	32.8%	35	28.6%	48.6%	91	34.1%	45.1%	32	56.3%	71.9%
Not positive	12	25.0%	33.3%	9	1/9	4/9	5	0/5	1/5	17	47.1%	64.7%
Uncertain	27	7.4%	29.6%	6	1/6	3/6	4	0/4	0/4	1	1/1	1/1
Meta-analysis results												
Not conducted	40	2.5%	15.0%				32	6.3%	21.9%	2	0/2	1/2
Significant	49	36.7%	44.9%				66	43.9%	53.0%	33	57.6%	72.7%
Non-significant	11	18.2%	36.4%				2	0/2	0/2	15	53.3%	66.7%
Total	100	21.0%	32.0%	50	24.0%	48.0%	100	31.0%	42.0%	50	54.0%	70.0%

Actual numbers are given when denominators are < 10.

the level of efforts taken by the review authors to reduce publication bias (based on the assumption that all other assessors differed in a constant way from the first assessor). Efforts taken to minimise bias were 'sufficient' if the review attempted to search and probably include non-English-language studies, grey literature and unpublished studies, considered outcome reporting bias, the issue of publication bias and reported any missing outcome data. 'Partial' efforts to minimise bias were when the review searched at least two of the three, i.e. non-English-language studies and/or grey literature and/or unpublished studies, may or may not have considered outcome reporting bias and publication bias, and may or may not have reported missing outcome data. Efforts taken to minimise bias were 'insufficient' when no attempts were made by the review authors to search for non-English-language studies, grey literature or unpublished studies, they did not consider outcome reporting bias or publication bias, and did not report missing outcome data. These judgements were then converted to scores as follows: insufficient = 0; partial = 1 and sufficient = 2, and pooled together for each assessor.

Table 9 shows the results of assessors' judgement about whether the review authors' efforts to minimise publication bias were sufficient or not. There was a fair agreement between assessors for the treatment reviews ($\kappa = 0.30$) and risk factor reviews ($\kappa = 0.35$). Moderate agreement was seen for genetic ($\kappa = 0.43$) and diagnostic reviews ($\kappa = 0.40$). The rate of agreement between assessors was 55% for treatment reviews, 68% for diagnostic and risk factor reviews, and 74% for genetic reviews. Based on the agreed judgement by two independent assessors, efforts to minimise publication bias were less likely to be insufficient in treatment reviews (18%) compared with diagnostic reviews (30%), risk factor reviews (55%) and genetic reviews (56%) (*Table 9*).

Risk of publication bias

Two assessors also independently assessed the possibility that review authors' conclusions might be invalid because of possible publication and related biases. The judgement for assessing the potential risk of publication bias was based on the efforts taken to minimise publication bias, discussion of publication bias, methods used to deal with publication bias and finally the authors' conclusion. The assessment was subjective without proper validation of the criteria. Two assessors independently scored the perceived

risk of publication bias in systematic reviews. Risk of bias was marked as 'high' if the efforts taken to minimise publication bias were partial or insufficient, publication bias was not discussed and the authors' conclusions were positive. Risk of bias was 'moderate' if partial efforts were taken to minimise bias, publication bias was probably considered, and the author's conclusions might have been positive with cautious interpretation. Risk of bias was 'low' if partial or sufficient efforts were taken to minimise bias, publication bias was considered with some methods used to deal with it, and the author's conclusions were negative. These judgements were converted to the following scores: low = 0, moderate = 1 and high = 2.

Table 10 shows the results of assessors' assessment of the risk of publication bias in reviews. Agreement between assessors was poor for treatment reviews ($\kappa = 0.17$), fair for genetic and diagnostic reviews ($\kappa = 0.29$ and $\kappa = 0.25$ respectively), and moderate for risk factor reviews ($\kappa = 0.44$). The rate of agreement between the two assessors was 53% for treatment reviews, 68% for diagnostic reviews, 73% for risk factor reviews and 60% for genetic reviews. According to the agreed judgement by two independent assessors, the rate of moderate to high risk of publication bias was relatively lower in treatment reviews (48%) in comparison with diagnostic reviews (64%), risk factors reviews (71%) and genetic reviews (58%).

Assessment of reviews that explicitly tested for publication bias

A random sample was obtained of 50 reviews, published from 2000 to 2008 in MEDLINE, that explicitly tested for or considered publication bias. Data extraction of these 50 reviews was independently conducted by two reviewers and any disagreements were resolved by discussion between the two reviewers or by a third reviewer. Three reviews were further excluded as the full publication could not be obtained; hence a total of 47 reviews were analysed.

Of the 47 reviews included, 18 (38%) evaluated effects of treatment intervention, 16 (34%) studied the association between various risk factors and disease, seven (15%) studied the association between a specific gene and disease, three (6%) evaluated the accuracy of diagnostic tests, and the remaining three (6%) had other objectives. We analysed if non-English-language

TABLE 9 *Assessors' judgement about efforts taken in reviews to minimise the risk of publication bias*

Treatment reviews

Assessor-1	Assessor-2			
	Insufficient	Partial	Sufficient	Total
Insufficient	18	4	1	23
Partial	24	34	8	66
Sufficient	0	8	3	11
Total	42	46	12	100

Diagnostic reviews

Assessor-1	Assessor-2			
	Insufficient	Partial	Sufficient	Total
Insufficient	15	3	0	18
Partial	9	19	1	29
Sufficient	0	3	0	3
Total	24	25	1	50

Risk factor reviews

Assessor-1	Assessor-2			
	Insufficient	Partial	Sufficient	Total
Insufficient	55	5	0	60
Partial	19	12	4	35
Sufficient	3	1	1	5
Total	77	18	5	100

Genetic reviews

Assessor-1	Assessor-2			
	Insufficient	Partial	Sufficient	Total
Insufficient	28	12	0	40
Partial	1	8	0	9
Sufficient	0	0	1	1
Total	29	20	1	50

studies, grey literature and unpublished studies were searched for and included in these reviews. Non-English-language studies were explicitly searched for in 38% of reviews including four narrative reviews and 14 meta-analyses. None of the four narrative reviews included non-English-language studies. The remaining reviews (40%) were unclear regarding the search for or inclusion of non-English-language studies. Grey literature was searched for in 14% of the reviews and 13% included them. A majority of the reviews (85%) did not clearly mention searching for or including grey literature. The analysis showed that 72% of the reviews did not mention searching for or including unpublished studies. Only nine (19%) exclusively searched for unpublished studies and only two (4%) included them.

The assessment of the sources searched in these reviews is consistent with the findings of the 300 reviews assessed in the first part of this chapter. MEDLINE (96%) was the most commonly searched database. Other sources used to identify literature were checking of reference lists of identified studies (81%), EMBASE (64%), specialised and other databases (53%), Cochrane Library (47%), contacting authors or experts (30%), hand searching journals (17%), CINAHL (17%), PsycINFO (15%), conference proceedings (13%), SIGLE (6%), and contacting pharmaceutical companies (6%). Four percent of the reviews did not state the sources searched. Language restriction was not clearly stated in 36% of the reviews, and 36% of the reviews were restricted to one or more

TABLE 10 *Assessors' judgement about risk of publication bias in reviews*

Treatment reviews

Assessor-1	Assessor-2			
	Low	Moderate	High	Total
Low	5	7	1	13
Moderate	12	35	16	63
High	0	11	13	24
Total	17	53	30	100

Diagnostic reviews

Assessor-1	Assessor-2			
	Low	Moderate	High	Total
Low	2	1	0	3
Moderate	3	29	9	41
High	0	3	3	6
Total	5	33	12	50

Risk factor reviews

Assessor-1	Assessor-2			
	Low	Moderate	High	Total
Low	2	0	4	6
Moderate	3	14	20	37
High	0	0	57	57
Total	5	14	81	100

Genetic reviews

Assessor-1	Assessor-2			
	Low	Moderate	High	Total
Low	1	0	0	1
Moderate	3	13	10	26
High	2	5	16	23
Total	6	18	26	50

languages (more commonly the English language 23%). Outcome reporting bias was not considered in 87% of the reviews, and only 11% of the reviews reported missing outcome data.

The analysis of reviews for discussion of publication bias and the methods used to test publication bias showed that publication bias was discussed in 44 (94%) of the reviews, and that funnel plot was the most commonly used method to detect publication bias (75%). This method was then followed by Egger's test (49%) and Begg's test (32%). Methods like identifying unpublished trials, the fail-safe N method and trim and fill method were only rarely used (2% each). Of the 44 reviews in which authors explicitly considered risk of publication bias, 19

(43%) reviews had a high risk of bias, eight (18%) had a moderate risk of bias and 17 (39%) had a low risk of bias.

We further assessed the association of discussion of publication bias with the review authors' conclusions and found that 39 reviews that discussed publication bias had significant results with considerable uncertainty. Another four reviews that discussed publication bias had non-significant results and only one review had significant results. This indicates that in most of the reviews the authors have interpreted the results with caution when there is any consideration or existence of publication bias.

Two assessors independently assessed the efforts taken to minimise publication bias within the selected sample, and the risk of publication bias. In assessing whether the efforts taken to minimise publication bias were sufficient, partial or insufficient, there was a fair agreement between assessors ($\kappa = 0.32$). The judgement for assessing the potential risk of publication bias was based on the efforts taken to minimise publication bias, discussion of publication bias, methods used to deal with publication bias and finally the authors' conclusion. The 47 reviews showed a poor agreement between assessors for the risk of publication bias ($\kappa = 0.09$).

Summary

A survey of systematic reviews of studies of treatment efficacy and diagnostic accuracy published in 1996 concluded that the issue of publication bias was largely being ignored in systematic reviews, and very few of them actually used any methods to deal with publication bias. However, in the current survey of reviews published in 2006, there was some improvement in the methods used to deal with publication bias. Reviews of health-care interventions (therapeutic or diagnostic) are making greater efforts to locate and/or include non-English-language studies (47% versus 30%), and grey literature or unpublished studies (53% versus 35%). A thorough literature search while conducting a systematic review may reduce the possibility of excluding unpublished studies, those published in non-English languages or as grey literature. It is always advisable to search more than one electronic database as many journals are indexed in only one of the commonly used databases.[517]

Compared with the previous sample of reviews, there was an increase in the use of available methods to test for publication bias in recent reviews (22% versus 17%). However, the proportion of reviews in which publication bias was explicitly discussed remained the same between recent treatment and diagnostic reviews and the previous sample (37% versus 36%).

The previous assessment recognised that the problems of publication and related biases were more often dealt with in meta-analysis than in narrative reviews. This finding is unchanged in the updated review and which could merely be a reflection of marked heterogeneity within the sample. Assessment of the narrative reviews showed an overall lack of efforts taken to reduce or minimise publication bias in all four categories of reviews.

Funnel plot and related statistical tests (including Egger's test and Begg's test) are common methods used to detect publication bias in systematic reviews, particularly in risk factor reviews. The fail-safe N method was used in some previous reviews but it was much less likely to be used in recent reviews (7% versus 1%). All other methods are not, or very rarely, used in the 300 general reviews and in the 44 reviews in which publication bias was explicitly tested.

Non-English-language studies and grey literature or unpublished studies were more likely to be explicitly searched for in treatment and diagnostic reviews, compared with reviews of epidemiological studies (50% versus 35%, and 53% versus 34%, respectively). Conversely, publication bias was less likely to be tested and discussed in treatment and diagnostic reviews than in epidemiological reviews (22% versus 39%, and 37% versus 51%, respectively). These differences between reviews of intervention studies and reviews of observational studies are possibly due to different approaches taken by authors in different fields to deal with perceived problems of publication bias. In a recent study that examined the frequency and determinants of full publication of studies of diagnostic accuracy submitted as abstracts at international stroke meetings, it was found that 76% of 160 abstracts were subsequently published in full and that clinical utility of results or other study characteristics did not predict their publication. However, this study was unable to assess the extent of a possible bias in the selection of abstracts for presentation.[57]

When assessors were asked to assess independently the level of efforts taken to minimise publication bias and the risk of publication bias in reviews, the rate of agreement was on an average 64% and 63% respectively. Based on the agreed judgement, reviews of treatment effect were more likely to have insufficient efforts to minimise publication bias, but less likely to have moderate or high risk of publication bias, compared with reviews of diagnostic accuracy or risk factors (including gene-disease association). According to data from 44 reviews in which risk of publication bias was explicitly considered by authors, 43% of reviews had a high risk, 18% had a moderate risk and 39% had a low risk of publication bias.

The assessment of reviews was challenging in many ways. Most of the variables in the data extraction form were assessed subjectively as 'yes', 'no' or 'unclear' and hence information may have been lost. For example, many studies reported that non-English-language studies were included, but to what extent they were searched for was unclear. The extent of searching for studies in languages other than English may vary, from having no language restriction in a PubMed search to running searches in specific non-English language databases. The same applies for grey literature and unpublished studies. The risk of publication bias was assessed from several perspectives: completeness of literature search, findings of any efforts to detect publication bias, and results of meta-analysis. This assessment was qualitative and the criteria have not been properly validated. However, we reported the results of the assessment of risk of bias to illustrate difficulties in any such attempts.

Chapter 10

Discussion

Available evidence on publication bias

The updated review confirmed findings from the previous *HTA* report that studies with significant or important results were more likely to be published than those with non-significant or 'unimportant' results. It appears that publication bias occurs mainly before the presentation of findings at conferences and the submission of manuscripts to journals. However, factors associated with publication bias remain unclear. The existence of outcome reporting bias has been demonstrated by recently published empirical studies. There is limited evidence indicating that harm outcomes and subjectively assessed outcomes may be more vulnerable to reporting bias than efficacy outcomes and objectively assessed outcomes.

Studies with significant or important results were, on average, published earlier than studies with non-significant results, although the new evidence was less clear than was suggested in the previous report. Any time lag bias is likely to occur before manuscript submission for journal publication.[81] Substantial new evidence on grey literature and language bias was identified in this updated review. Grey literature or non-English-language studies on average reported smaller treatment effects than studies that were formally published or studies that were published in English. However, the direction and extent of bias was usually unpredictable. There is limited evidence indicating that the risk of language bias may be particularly high in some areas of research such as complementary and alternative medicine. The updated review also identified limited new evidence on citation bias, duplicate publication bias, place of publication bias, database bias and media attention bias.

As a direct consequence of publication and related biases, estimates based on published studies may be misleading. For example, publication and related bias may result in an overestimation of treatment effects or an underestimation of adverse effects. In this updated review, the consequences of publication and related biases were separately discussed for basic (animal and laboratory) research, observational studies and clinical trials.

Biased publication of results of basic research may explain negative results from subsequent clinical trials. Contradictory findings from epidemiological studies may be partly due to publication and related biases. The consequences of publication bias in clinical trials may be more serious, resulting in the use of less cost-effective, or ineffective, or even harmful interventions in clinical practice. This updated review identified a few new cases that indicated the detrimental impact of publication and related biases.

This updated review confirmed findings from the previous *HTA* report that the most common reason for publication bias was that investigators failed to write up or submit studies with non-significant results (see *Figure 7*). However, it should be recognised that investigators' decision to write up an article and then submit it may be affected by pressure from research sponsors, instruction from journal editors, and requirements of the research award system. Clearly, commercial and other competing interests of research sponsors and investigators may influence the profile of dissemination of research findings.

Limitations of evidence studies on publication bias

The most important evidence on publication bias comes from cohort studies showing that the publication of studies is associated with the strength or direction of the results. However, the definition of publication status and classification of study results are often different in empirical studies of publication bias. For time lag bias, time to publication could be measured starting from different time points (e.g. approval by REC, recruitment of participants, completion of follow-up) during the process of research. Therefore, bias may be introduced in studies of publication bias because of inevitable subjectivity in the choice of definitions and methods.

Large cohort studies on publication and related biases usually included cases that were highly diverse in terms of research questions, designs and other study characteristics. Many factors (e.g. sample size, design, research question and

investigators' characteristics) may be associated with both study results and the possibility of publication. Adjusted analyses by some factors may be conducted but it was generally impossible to exclude the impact of confounding factors on the observed association between study results and formal publication. However, confounding factors may not be a problem in many single case studies that provided empirical evidence on publication and related biases. But, evidence from case studies is susceptible to bias due to selective reporting.[123]

There are several high-quality empirical studies that were less selective and in which the impact of confounding factors could be controlled. For example, Egger *et al.* (2003)[3] and Moher *et al.* (2003)[4] used multiple meta-analyses to investigate grey literature and language bias (see Chapter 3). The results of trials published in English and those published in non-English languages was compared within each meta-analysis that aimed to answer the same clinical question. In empirical studies by Chan *et al.*,[6,7] outcomes reported in published papers were compared with outcomes specified in protocols within each trial, so that the observed outcome reporting bias is unlikely to be due to confounding factors. However, generalisability is still an issue even for findings from these good quality empirical studies.

Studies of publication bias themselves may be as vulnerable as other studies to the selective publication of significant or striking findings.[1,518] Dubben and Beck-Bornholdt (2005) used the funnel plot approach, and found no evidence of publication bias in studies of publication bias.[519] They acknowledged that the analysis was handicapped by insufficient power (with only 26 included studies) and also by the diverse definitions of publication bias in the primary studies. Song *et al.* pointed out that the study had other more important limitations so that dissemination bias of studies on publication bias could not be safely excluded.[520] Funnel plot analysis was used in Chapter 3 to detect small study effects in cohort studies of publication bias (see *Figure 6*), and the plot was not statistically significantly asymmetric. However, there is still reason to suspect the existence of publication or reporting bias in studies of publication bias. We identified a large number of reports of full publication of meeting abstracts, and the association between study results and full publication had not been reported in most of these reports. It is often unclear whether this association had not been examined, or was not reported because the association proved to be non-

significant. In addition, we have mentioned earlier that single case studies that provided empirical evidence on publication bias may be biased because of selective reporting of striking findings.

The existence of publication and related biases was usually confirmed by comparing results of published studies with those of unpublished studies. However, the actual impact of such bias is best investigated by a comparison of the result of published studies with the result of a combination of published and unpublished studies. In most cases, the actual impact of publication and related biases was non-significant in a systematic review that combined evidence from all relevant studies.[3]

How to deal with publication bias?

The consequences of publication bias were previously overlooked by many leading experts.[2] According to Beveridge, research with non-significant results 'clutters up the journals and does more harm than good to the author's reputation in the minds of the discerning.'[521] A book about ethics in the dissemination of new knowledge even recommended that 'it is preferable to publish positive research findings, because they advance knowledge'.[522]

The importance of 'negative' findings from research has now been generally recognised. A wide public awareness of detrimental consequences of publication bias has promoted recent efforts to prevent and reduce publication bias. For example, regulatory authorities, journals and research sponsors have taken action to improve the current situation because of several high-profile cases of incomplete disclosure of negative results of trials sponsored by pharmaceutical companies (see Chapter 7).

According to the stage of a literature review, measures to combat publication and related bias can be classified as those before, during or after a literature review (see *Figure 1*).[2] *Table 11* shows various methods that can be used to deal with different types of publication and related biased. For example, methods that can be used to combat the non-publication of 'negative' findings include prospective registration of studies, disclosure of data from unpublished studies, searching for and inclusion of unpublished studies, and assessment of risk of publication bias in systematic reviews.

The recent development of information technology and electronic publication provides great technical potential to overcome some limitations of conventional printed journals. Publication bias may be reduced by publication in electronic media with unlimited space, direct electronic linkage between references, timely publication, and cost-effective dissemination and archiving. In addition, electronic open-access databases maintained by regulatory bodies, research societies or research sponsors are increasingly important sources of published and unpublished studies.

It seems still reasonable to claim that 'the ideal solution to publication bias is the prospective, universal registration of all studies at their inception'.[2] Voluntary registration of trials is usually incomplete. The most important development was initiated by the compulsory policy of trial registration adopted by the International Committee of Medical Journals in 2004.[10] Efforts so far have focused on the registration, publication and disclosure of confirmatory phase III/IV trials due to the perceived immediate consequences. In spite of the greater risk of publication bias, there have been considerable difficulties facing the prospective registration of and disclosure of data from unpublished basic research, observational studies and early stage exploratory clinical trials.

Trials registers will only be helpful in reducing publication bias if systematic reviewers include the registries in their search strategies and results of trials are accessible. According to findings presented in Chapter 9 (see *Table 4*), only 18% of the treatment reviews and few reviews of diagnostic and epidemiological studies searched prospective research registers.

Certain types of dissemination bias, such as database bias, duplicate publication bias, citation bias and media attention bias, could be dealt with by following approaches adopted in standard systematic reviews (*Table 11*). The risk of publication bias may be assessed in a systematic review according to certain risk factors associated with publication bias, although the method has not been adequately investigated in empirical studies (Chapter 8). Funnel plot and related statistical methods have been widely used to assess publication bias in systematic reviews. Because the interpretation of a funnel plot can often be misleading, some recommendations have been recently proposed about when, and how to use the funnel plot and related statistical tests.[469,470] However, these recommendations were based on

limited and fast-changing evidence and have not been empirically validated.

Many complex statistical methods have been developed to detect or even adjust for assumed publication bias in meta-analysis. But they have never or very rarely been used in practice, possibly because of their complexity and the lack of user-friendly software. More importantly, the usefulness of any statistical methods, simple or complex, may be very limited in typical systematic reviews or meta-analyses (Chapter 9). It is now generally recognised that sophisticated modelling methods may be used to conduct sensitivity analyses, rather than to provide an adjusted estimate, although the usefulness of such sensitivity analyses is still unclear.

Dealing with publication bias in published systematic reviews

The 2000 *HTA* report on publication bias surveyed a sample of 193 systematic reviews published in 1996, and concluded that the issue of publication bias was largely being ignored, and methods to deal with publication bias were rarely used in these reviews.[2] This updated review found that in 300 systematic reviews published in 2006 there have been some improvements in dealing with publication bias (Chapter 9). Compared with reviews published in 1996, recently published reviews made greater efforts to locate and include grey literature or unpublished studies and studies published in non-English languages. In addition, more recently published reviews used methods to assess publication bias in systematic reviews.

The previous report found that for publications in 1996 the problems of publication bias were more often dealt with in meta-analyses than in narrative reviews. This phenomenon is also observed in systematic reviews published in 2006, which may be due to lack of methods that can be used in narrative reviews.

We observed some differences between different categories of systematic reviews published in 2006. Grey literature, unpublished studies or non-English-language studies were more likely to be searched for in reviews of treatment efficacy or diagnostic accuracy than in reviews of epidemiological studies. However, the risk of publication bias was less likely to be tested in reviews of treatment and diagnosis compared

TABLE 11 *Publication-related biases and methods to deal with these biases*

Methods/approaches	Non-publication of 'negative' results	Time lag bias	Outcome reporting bias	Grey literature bias	Language bias	Database bias / Duplicate bias / Citation bias / Media bias
Prospective registration of studies, publication of research protocols	✓	✓	✓	✓	✓	✓
Right to publication	✓	✓	✓			
Open access policy/regulation, improved research funders' guidelines	✓		✓	✓		
Endorsement of sound reporting guidelines for journal publication			✓			
Disclosure of unpublished studies or data	✓	✓	✓	✓		
Systematic literature review – Searching for and including grey literature, unpublished studies/data, and non-English-language studies				✓	✓	✓
Assessing risk of publication bias in systematic reviews – Considering risk factors, funnel plot and related tests	✓	✓	✓	✓	✓	
Contacting authors for missing data or clarification			✓		✓	✓
Individual patient data meta-analysis			✓	✓		
Updating systematic reviews		✓				
Confirmatory studies						

with reviews of epidemiological studies (Chapter 8). These differences between different types of reviews may be caused by the different availability of sources of grey literature or unpublished studies, and perceived risk of publication bias in different types of primary studies. For example, many initiatives have been taken to prevent biased publication of clinical trials and there are some good sources of grey literature and unpublished trials. At the same time, the limitations of available methods to test publication bias in systematic reviews have been more widely recognised. Therefore, the authors of reviews of treatment efficacy may focus their efforts on the completeness of literature search, rather than on the assessment of publication bias. However, there have been no great efforts to prevent publication bias in epidemiological studies. No good databases of unpublished epidemiological studies could be used by the authors of reviews of epidemiological studies. In view of the great risk of publication bias, authors of reviews of epidemiological studies may have to rely on the available methods to test publication bias, even if the results of such tests are often difficult to interpret.

Implications for researchers and decision-makers

- There is little doubt that dissemination of research findings is likely to be a biased process, although the actual impact of such bias is often uncertain, depending on specific circumstances. Therefore, the potential problem of publication and related bias should be taken into consideration by all who are involved in evidence-based decision-making.
- Decision-makers, research funders and RECs at the national and international level should continue to support the development of prospective research registration, and the implementation of research open-access policy.
- Practical and sound reporting guidelines should be endorsed by journals, and authors should report all measured outcomes in their studies.
- Whenever possible, a thorough literature search should be conducted in systematic reviews to identify all relevant studies. Registers of clinical trials and available databases of unpublished studies should be routinely searched for relevant clinical trials.
- The impact of grey literature or studies published in languages other than English may be non-significant in many cases. However, the exclusion of grey literature or non-English-language studies may introduce bias in a systematic review, particularly in the field of complementary medicine. Therefore, systematic reviews should not routinely exclude unpublished studies or conference abstracts. The quality of unpublished studies or abstracts should be assessed using the same criteria as for formally published studies.
- Outcome reporting bias has been confirmed by new evidence and should be seriously considered in systematic reviews. When relevant studies cannot be included owing to a lack of data on relevant outcomes, original authors should be contacted to clarify whether the outcome was actually measured, and to obtain data on missing outcomes.
- Funnel plot and related statistical tests can be used to detect 'small study effect'. However, it is usually impossible to separate the influence of factors other than publication bias on the observed association between the estimated effects and sample sizes across studies in meta-analysis. The inappropriate interpretation of

the funnel plot and its related tests may be reduced by following recent recommendations in the updated *Cochrane Handbook for Systematic Reviews*.[469]
- The risk of publication bias should be qualitatively assessed according to suspected factors associated with publication bias, including small sample size, small effect size, the shape of a funnel plot, the potential number of studies that may have been conducted, conflicting interests of investigators or research sponsors, and any other direct or indirect evidence. The estimated risk of publication bias should be incorporated into the review's conclusions.
- Large-scale confirmatory studies become necessary after a systematic review has reported a clinically significant finding, but publication bias cannot be safely excluded.

Recommendations for future research

- Further empirical research is needed to evaluate the effect of prospective registration of studies, open-access policy and improved publication guidelines in the prevention of research dissemination bias.
- The role of the developments in computer science and information technology for the prevention and reduction of research dissemination bias needs to be investigated by further research.
- There is still a lack of evidence about the impact of publication bias on health decision-making and the outcomes of patient management. Further research is required in this area.
- Many systematic reviews still have to depend upon studies identified retrospectively from the published literature, particularly in systematic reviews of basic research and observational studies. Further research is required to develop methods that can be used qualitatively or narratively to assess the risk of publication bias in systematic reviews.
- Many available statistical methods to test publication bias have never, or very rarely, been used in systematic reviews. Further research should focus on the practical application of these statistical methods.

Acknowledgements

References

1. Begg CB, Berlin JA. Publication bias; a problem in interpreting medical data. *J Roy Stat Soc A* 1988;**151**:445–63.

2. Song F, Eastwood AJ, Gilbody S, Duley L, Sutton AJ. Publication and related biases. *Health Technol Assess* 2000;**4**:1–115.

3. Egger M, Juni P, Bartlett C, Holenstein F, Sterne J. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. *Health Technol Assess* 2003;**7**:1–76.

4. Moher D, Pham B, Lawson ML, Klassen TP. The inclusion of reports of randomised trials published in languages other than English in systematic reviews. *Health Technol Assess* 2003;**7**:1–90.

5. Chan A-W, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. *BMJ* 2005;**330**:753.

6. Chan A-W, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004;**291**:2457–65.

7. Chan A-W, Krleza-Jeri K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research [see comment]. *Can Med Assoc J* 2004;**171**:735–40.

8. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed [see comment]. *J Clin Epidemiol* 2005;**58**:882–93.

9. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Comparison of two methods to detect publication bias in meta-analysis. *JAMA* 2006;**295**:676–80.

10. DeAngelis CD, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, *et al.* Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *JAMA* 2004;**292**:1363–4.

11. Rosenthal R. The 'file drawer problem' and tolerance for null results. *Psychol Bull* 1979;**86**:638–41.

12. Begg CB, Berlin JA. Publication bias and dissemination of clinical research. *J Natl Cancer Inst* 1989;**81**:107–15.

13. Dickersin K. The existence of publication bias and risk factors for its occurrence. *JAMA* 1990;**263**:1385–9.

14. Smith R. Editorial: What is publication? *BMJ* 1999;**318**:142.

15. Edwards S, Liford RJ, Kiauka S. Different types of systematic review in health services research. In Black N, Brazier H, Fitzpatrick R, Reeves B, editors. *Health services research methods: a guide to best practice.* London: BMJ Books; 1998.

16. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**:557–60.

17. Sohn D. Publications bias and the evaluation of psychotherapy efficacy in reviews of the research literature. *Clin Psychol Rev* 1996;**16**:147–56.

18. Sterling T. Publication decisions and their possible effects on inferences drawn from tests of significance – or vice versa. *Am Stat Assoc J* 1959;**54**:30–4.

19. Sterling TD, Rosenbaum WL, Weinkam JJ. Publication decisions revisited – the effect of the outcome of statistical tests on the decision to publish and vice-versa. *Am Stat* 1995;**49**:108–12.

20. Dickersin K, Min YI, Meinert CL. Factors influencing publication of research results. Follow up of applications submitted to two institutional review boards. *JAMA* 1992;**267**:374–8.

21. Dickersin K, Min. YI. NIH clinical trials and publication bias. *Online J Curr Clin Trials* 1993. Doc. No. 50(3).

22. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet* 1991;**337**:867–72.

23. Ioannidis J. Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *JAMA* 1998;**279**:281–6.

24. Stern JM, Simes RJ. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *BMJ* 1997;**315**:640–5.

25. Dickersin K. How important is publication bias? A synthesis of available data. *Aids Educ Prev* 1997;**9**(1 SA):15–21.

26. Bardy AH. Bias in reporting clinical trials. *Brit J Clin Pharmaco* 1998;**46**:147–50.

27. Cronin E, Sheldon T. Factors influencing the publication of health research. *Int J Technol Assess* 2004;**20**:351–5.

28. Decullier E, Lheritier V, Chapuis F. Fate of biomedical research protocols and publication bias in France: retrospective cohort study [see comment]. *BMJ* 2005;**331**:19.

29. Decullier E, Chapuis F. Impact of funding on biomedical research: a retrospective cohort study. *BMC Public Health* 2006;**6**:165.

30. Misakian AL, Bero LA. Publication bias and research on passive smoking: comparison of published and unpublished studies. *JAMA* 1998;**280**:250–3.

31. Wormald R, Bloom J, Evans J, Oldfled K. Publication bias in eye trials. In *Second International Conference Scientific Basis of Health Science and 5th Cochrane Colloquium*. Amsterdam, The Netherlands: Cochrane Collaboration; 1997.

32. Zimpel T, Windeler J. [Publications of dissertations on unconventional medical therapy and diagnosis procedures – a contribution to 'publication bias']. *Forsch Komp Klas Nat* 2000;**7**:71–4.

33. Blumle A, Antes G, Schumacher M, Just H, von Elm E. Clinical research projects at a German medical faculty: follow-up from ethical approval to publication and citation by others. *J Med Ethics* 2008;**34**:e20.

34. Druss BG, Marcus SC. Tracking publication outcomes of National Institutes of Health grants. *Am J Med* 2005;**118**:658–63.

35. Cooper H, Charlton K. Finding the missing science: the fate of studies submitted for review by a human subjects committee. *Psychol Methods* 1997;**2**:447–52.

36. Hahn S, Williamson PR, Hutton JL. Investigation of within-study selective reporting in clinical research: follow-up of applications submitted to a local research ethics committee. *J Eval Clin Pract* 2002;**8**:353–9.

37. Hall R, de Antueno C, Webber A. Publication bias in the medical literature: a review by a Canadian Research Ethics Board. *Can J Anaesth* 2007;**54**:380–8.

38. Pich J, Carne X, Arnaiz J-A, Gomez B, Trilla A, Rodes J. Role of a research ethics committee in follow-up and publication of results [see comment]. *Lancet* 2003;**361**:1015–6.

39. von Elm E, Rollin A, Blumle A, Huwiler K, Witschi M, Egger M. Publication and non-publication of clinical trials: longitudinal study of applications submitted to a research ethics committee. *Swiss Med Wkly* 2008;**138**:197–203.

40. Dickersin K, Min YI. Publication bias: the problem that won't go away. *Ann N Y Acad Sci* 1993;**703**:135–46.

41. Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan AW, Cronin E, *et al.* Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS ONE* 2008;**3**(8):e3081.

42. Lee K, Bacchetti P, Sim I. Publication of clinical trials supporting successful new drug applications: a literature analysis. *PLoS Med* 2008;**5**(9):e191.

43. Melander H, Ahlqvist-Rastad J, Meijer G, Beermann B. Evidence b(i)ased medicine – selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. *BMJ* 2003;**326**:1171–3.

44. Rising K, Bacchetti P, Bero L. Reporting bias in drug trials submitted to the Food and Drug Administration: review of publication and presentation. *PLoS Med* 2008;**5**(11):e217; discussion e217.

45. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* 2008;**358**:252–60.

46. van Luijn JCF, Stolk P, Gribnau FWJ, Leufkens HGM. Gap in publication of comparative information on new medicines. *Brit J Clin Pharmacol* 2008;**65**:716–22.

47. Callaham ML, Wears RL, Weber EJ, Barton C, Young G. Positive-outcome bias and other limitations in the outcome of research abstracts submitted to a scientific meeting. *JAMA* 1998;**280**:254–7. Erratum appears in *JAMA* 1998;**280**:1232.

48. Chalmers I, Adams M, Dickersin K, Hetherington J, Tamow-Mordi W, Meinert C, *et al.* A cohort study of summary reports of controlled trials. *JAMA* 1990;**263**:1401–5.

49. Cheng K, Preston C, Ashby D, O'Hea U, Smyth RL. Time to publication as full reports of abstracts of randomized controlled trials in cystic fibrosis. *Pediatr Pulm* 1998;**26**:101–5.

50. DeBellefeuille C, Morrison CA, Tannock IF. The fate of abstracts submitted to a cancer meeting: factors which influence presentation and subsequent publication. *Ann Oncol* 1992;**3**:187–91.

51. Landry VL. The publication outcome for the papers presented at the 1990 ABA conference. *J Burn Care Rehabil* 1996;**17**:23A–6A.

52. Loep M, Kleijnen J. Full publication of abstracts initially published in the *Netherlands Journal of Medicine*, 1999. Data obtained by personal communication with the author, Jo Kleijnen, Centre for Reviews and Dissemination, University of York, UK, June 1999.

53. Petticrew M, Gilbody S, Song F. Lost information? The fate of papers presented at the 40th Society for Social Medicine Conference. *J Epidemiol Commun H* 1999;**53**:442–3.

54. Scherer RW, Dickersin K, Langenberg P. Full publication of results initially presented in abstracts. A meta-analysis *JAMA* 1994;**272**:158–62. Published erratum appears in *JAMA* 1994;**272**:1410.

55. Scherer RW, Langenberg P, von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database Syst Rev* 2007;**2**:MR000005. DOI:10.1002/14651858.MR000005.pub3.

56. Akbari-Kamrani M, Shakiba B, Parsian S. Transition from congress abstract to full publication for clinical trials presented at laser meetings. *Laser Med Sci* 2008;**23**:295–9.

57. Brazzelli M, Lewis SC, Deeks JJ, Sandercock PAG. No evidence of bias in the process of publication of diagnostic accuracy studies in stroke submitted as abstracts. *J Clin Epidemiol* 2009;**62**:425–30.

58. Castillo J, Garcia Guasch R, Cifuentes I. Fate of abstracts from the Paris 1995 European Society of Anaesthesiologists meeting. *Eur J Anaesthesiol* 2002;**19**:888–93.

59. Delamere F, Williams H. To what extent are conference abstracts reporting randomised controlled trials of skin diseases published subsequently? [abstract]. *XIII Cochrane Colloquium, Oct 22–26 2005; Melbourne, Australia*.

60. Eloubeidi MA, Wade SB, Provenzale D. Factors associated with acceptance and full publication of GI endoscopic research originally published in abstract form [see comment]. *Gastrointest Endosc* 2001;**53**:275–82.

61. Evers JL. Publication bias in reproductive research. *Hum Reprod* 2000;**15**:2063–6.

62. Glick N, MacDonald I, Knoll G, Brabant A, Gourishankar S. Factors associated with publication following presentation at a transplantation meeting. *Am J Transplant* 2006;**6**:552–6.

63. Ha TH, Yoon DY, Goo DH, Chang SK, Seo YL, Yun EJ, *et al.* Publication rates for abstracts presented by Korean investigators at major radiology meetings. *Korean J Radiol* 2008;**9**:303–11.

64. Halpern SH, Palmer C, Angle P, Tarshis J. Published abstracts in obstetrical anesthesia: full publication rates and data reliability. *Anesthesiology* 2001;**94**:A69.

65. Harris IA, Mourad M, Kadir A, Solomon MJ, Young JM. Publication bias in abstracts presented to the annual meeting of the American Academy of Orthopaedic Surgeons. *J Orthopaed Surg (Hong Kong)* 2007;**15**:62–6.

66. Harris IA, Mourad MS, Kadir A, Solomon MJ, Young JM. Publication bias in papers presented to the Australian Orthopaedic Association Annual Scientific Meeting. *Aust NZ J Surg* 2006;**76**:427–31.

67. Hashkes PJ, Uziel Y. The publication rate of abstracts from the 4th Park City Pediatric Rheumatology meeting in peer-reviewed journals: what factors influenced publication? *J Rheumatol* 2003;**30**:597–602.

68. Kiroff GK. Publication bias in presentations to the Annual Scientific Congress. *Aust NZ J Surg* 2001;**71**:167–71.

69. Klassen TP, Wiebe N, Russell K, Stevens K, Hartling L, Craig WR, *et al.* Abstracts of randomized controlled trials presented at the society for pediatric research meeting: an example of publication bias [comment]. *Arch Pediat Adol Med* 2002;**156**:474–9.

70. Krzyzanowska MK, Pintilie M, Tannock IF. Factors associated with failure to publish large randomized trials presented at an oncology meeting. *JAMA* 2003;**290**:495–501.

71. Peng PH, Wasserman JM, Rosenfeld RM. Factors influencing publication of abstracts presented at the AAO-HNS Annual Meeting. *Otolaryng Head Neck* 2006;**135**:197–203.

72. Sanossian N, Ohanian AG, Saver JL, Kim LI, Ovbiagele B. Frequency and determinants of nonpublication of research in the stroke literature. *Stroke* 2006;**37**:2588–92.

73. Smith WA, Cancel QV, Tseng TY, Sultan S, Vieweg J, Dahm P. Factors associated with the full publication of studies presented in abstract form at the annual meeting of the American Urological Association. *J Urol* 2007;**177**:1084–8; discussion 1088–9.

74. Timmer A, Hilsden RJ, Cole J, Hailey D, Sutherland LR. Publication bias in gastroenterological research – a retrospective cohort study based on abstracts submitted to a scientific meeting. *BMC Medical Research Methodology* 2002;**2**(1):7.

75. Vecchi S, Belleudi V, Amato L, Davoli M, Perucci CA. Does direction of results of abstracts submitted to scientific conferences on drug addiction predict full publication? *BMC Med Res Methodol* 2009;**9**:23.

76. Zamakhshary M, Abuznadah W, Zacny J, Giacomantonio M. Research publication in pediatric surgery: a cross-sectional study of papers presented at the Canadian Association of Pediatric Surgeons and the American Pediatric Surgery Association. *J Pediatr Surg* 2006;**41**:1298–301.

77. Zaretsky Y, Imrie K. The fate of phase III trial abstracts presented at the American Society of Hematology [abstract]. *Blood* 2002;**100**:185a.

78. Lee KP, Boyd EA, Holroyd-Leduc JM, Bacchetti P, Bero LA. Predictors of publication: characteristics of submitted manuscripts associated with acceptance at major biomedical journals. *Med J Aust* 2006;**184**:621–6.

79. Lynch JR, Cunningham MRA, Warme WJ, Schaad DC, Wolf FM, Leopold SS. Commercially funded and United States-based research is more likely to be published; good-quality studies with negative outcomes are not. *J Bone Joint Surg Am* 2007;**89**:1010–8.

80. Okike K, Kocher MS, Mehlman CT, Heckman JD, Bhandari M. Publication bias in orthopaedic research: an analysis of scientific factors associated with publication in the Journal of Bone and Joint Surgery (American Volume). *J Bone Joint Surg Am* 2008;**90**:595–601.

81. Olson CM, Rennie D, Cook D, Dickersin K, Flanagin A, Hogan JW, *et al.* Publication bias in editorial decision making. *JAMA* 2002;**287**:2825–8.

82. Dickersin K, Olson CM, Rennie D, Cook D, Flanagin A, Zhu Q, *et al.* Association between time interval to publication and statistical significance. *JAMA* 2002;**287**:2829–31.

83. Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. *Cochrane Database Syst Rev* 2009;**1**:MR000006. DOI:10.1002/14651858.MR000006.pub3.

84. Dickersin K, Ssemanda E, Mansell C, Rennie D. What do the JAMA editors say when they discuss manuscripts that they are considering for publication? Developing a schema for classifying the content of editorial discussion. *BMC Med Res Methodol* 2007;**7**:44.

85. Calnan M, Smith GD, Sterne JAC. The publication process itself was the major cause of publication bias in genetic epidemiology. *J Clin Epidemiol* 2006;**59**:1312–8.

86. Hahn S, Garner P, Williamson P. Are systematic reviews taking heterogeneity into account? An analysis from the Infectious Diseases Module of the Cochrane Library. *J Eval Clin Pract* 2000;**6**:231–3.

87. Phillips CV. Publication bias in situ. *BMC Med Res Methodol* 2004;**4**:20.

88. Pocock SJ, Hughes MD, Lee RJ. Statistical problems in the reporting of clinical trials. A survey of three medical journals. *N Engl J Med* 1987;**317**:426–32.

89. Tannock IF. False-positive results in clinical trials: multiple significance tests and the problem of unreported comparisons. *J Natl Cancer Inst* 1996;**88**:206–7.

90. West RRJ, D.A. Publication bias in statistical overview of trials: example of psychological rehabilitation following myocardial infarction [abstract]. *The 2nd International Conference on the Scientific Basis of Health Services and 5th Cochrane Colloquium, 8–12 October 1997, Amsterdam*.

91. Ghersi D, Clarke M, Simes J. Selective reporting of the primary outcomes of clinical trials: a follow-up study [abstract]. *14th Cochrane Colloquium, 23–26 October 2006, Dublin, Ireland*.

92. McCormack K, Grant A, Scott N. Value of updating a systematic review in surgery using individual patient data. *Br J Surg* 2004;**91**:495–9.

93. Bekkering GE, Harris RJ, Thomas S, Mayer AM, Beynon R, Ness AR, *et al.* How much of the data published in observational studies of the association between diet and prostate or bladder cancer is usable for meta-analysis? *Am J Epidemiol* 2008;**167**:1017–26.

94. Furukawa TA, Watanabe N, Omori IM, Montori VM, Guyatt GH. Association between unreported outcomes and effect size estimates in Cochrane meta-analyses. *JAMA* 2007;**297**:468–70.

95. Williamson PR, Gamble C. Identification and impact of outcome selection bias in meta-analysis. *Stat Med* 2005;**24**:1547–61.

96. Scharf O, Colevas AD. Adverse event reporting in publications compared with sponsor database for cancer clinical trials. *J Clin Oncol* 2006;**24**:3933–8.

97. Kyzas PA, Loizou KT, Ioannidis JPA. Selective reporting biases in cancer prognostic factor studies [see comment]. *J Natl Cancer I* 2005;**97**:1043–55.

98. Kavvoura FK, Liberopoulos G, Ioannidis JPA. Selection in reported epidemiological risks: an empirical assessment. *PLoS Medicine/Public Library of Science* 2007;**4**(3):e79.

99. Williamson P, Gamble C, Jacoby A, Altman D. Understanding the process and impact of within-study selective reporting bias [abstract]. *14th Cochrane Colloquium, 23–26 October 2006, Dublin, Ireland.*

100. Jadad AR, Rennie D. The randomized controlled trial gets a middle-aged checkup. *JAMA* 1998;**279**:319–20.

101. Simes RJ. Confronting publication bias: a cohort design for meta analysis. *Stat Med* 1987;**6**:11–29.

102. Liebeskind DS, Kidwell CS, Sayre JW, Saver JL. Evidence of publication bias in reporting acute stroke clinical trials. *Neurology* 2006;**67**:973–9.

103. Soares H, Kumar A, Clarke M, Djulbegovic B. How long does it take to publish a high quality trial in oncology? [abstract]. *13th Cochrane Colloquium, 22–26 October 2005, Melbourne, Australia.*

104. Min YI, Dickersin K. Rate of full publication and time to full publication of observational studies [abstract]. *Am J Epidemiol* 2005;**161**(Suppl.):S76.

105. Callaham ML, Weber E, Young G, Wears R, Barton C. Time to publication of studies was not affected by whether results were positive [comment]. *BMJ* 1998;**316**:1536.

106. Rothwell PM, Robertson G. Meta-analyses of randomised controlled trials [letter]. *Lancet* 1997;**350**:1181–2.

107. Song F, Gilbody S. Bias in meta-analysis detected by a simple, graphical test. Increase in studies of publication bias coincided with increasing use of meta-analysis [comment]. *BMJ* 1998;**316**:471.

108. Gehr BT, Weiss C, Porzsolt F. The fading of reported effectiveness. A meta-analysis of randomised controlled trials. *BMC Med Res Methodol* 2006;**6**:25.

109. Vaitkus PT, Brar C. *N*-acetylcysteine in the prevention of contrast-induced nephropathy: publication bias perpetuated by meta-analyses [see comment]. *Am Heart J* 2007;**153**:275–80.

110. Jennions MD, Moller AP. Relationships fade with time: a meta-analysis of temporal trends in publication in ecology and evolution. *Proceedings of the Royal Society of London Series B: Biological Sciences* 2002;**269**(1486):43–48.

111. Leimu R, Koricheva J. Cumulative meta-analysis: a new tool for detection of temporal trends and publication bias in ecology. *P R Soc B* 2004;**271**:1961–6.

112. Ioannidis JP, Trikalinos TA. Early extreme contradictory estimates may appear in published research: the Proteus phenomenon in molecular genetics research and randomized trials. *J Clin Epidemiol* 2005;**58**:543–9.

113. Hopewell S, Clarke M, Stewart L, Tierney J. Time to publication for results of clinical trials. *Cochrane Database Syst Rev* 2007;**2**:MR000011. DOI:10.1002/14651858.MR000011.pub3.

114. Auger CP. *Information sources in grey literature.* 4th edn. London/Melbourne/Munich/New Providence, NJ: Bowker-Saur; 1998.

115. Helmer D. Grey literature. In *Etext on Health Technology Assessment (HTA) information resources.* URL: www.nlm.nih.gov/archive//2060905/nichsr/ehta/chapter10.html. 2004.

116. Cook DJ, Guyatt GH, Ryan G, Clifton J, Buckingham L, Willan A, *et al.* Should unpublished data be included in meta analyses? Current convictions and controversies. *JAMA* 1993;**269**:2749–53.

117. Taus C, Pucci E, Giuliani G, Telaro E, Pistotti V. The use of 'grey literature' in a sub-set of neurological Cochrane reviews. *7th Cochrane Colloquium, 5–9 October 1999, Rome, Italy.*

118. Tetzlaff J, Moher D, Pham B, Altman D. Survey of views on including grey literature in systematic reviews [abstract]. *14th Cochrane Colloquium, 23–26 October 2006, Dublin, Ireland.*

119. McAuley L, Pham B, Tugwell P, Moher D. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? *Lancet* 2000;**356**:1228–31.

120. Burdett S, Stewart LA, Tierney JF. Publication bias and meta-analyses: a practical example. *Int J Technol Assess* 2003;**19**:129–34.

121. Ferguson D, Laupacis A, Salmi LR, McAlister FA, Huet C. What should be included in meta-analyses? An exploration of methodological issues using the ISPOT meta-analyses [review]. *Int J Technol Assess* 2000;**16**:1109–19.

122. Hopewell S. *Impact of grey literature on systematic reviews of randomised trials*. D Phil thesis. University of Oxford, 2004.

123. Hopewell S, McDonald S, Clarke M, Egger M. Grey literature in meta-analyses of randomized trials of health care interventions. *Cochrane Database Syst Rev* 2007;**2**:MR000010. DOI:10.1002/14651858. MR000010.pub3.

124. Glass GV, McGaw B, Smith ML. *Meta-analysis in social research*. London: Sage Publications; 1981.

125. Smith M. Publication bias and meta-analysis. *Eval Educ* 1980;**4**:22–4.

126. White KR. The relation between socioeconomic status and academic achievement. *Psychol Bull* 1982;**91**:461–81.

127. Simes RJ. Publication bias: the case for an international registry of clinical trials. *J Clin Oncol* 1986;**4**:1529–41.

128. Detsky AS, Baker JP, O'Rourke K, Goel V. Perioperative parenteral nutrition: a meta-analysis. *Ann Intern Med* 1987;**107**:195–203.

129. Devine EC. Empirical assessment of publication bias: lessons from two meta-analyses. *7th Cochrane Colloquium, 5–9 October 1999, Rome, Italy*.

130. MacLean C, Morton S, Straus W, Ofman J, Roth E, Shekelle P. Unpublished data from United States Food and Drug Administration New Drug Application Reviews: How do they compare to published data when assessing Nonsteroidal Antiinflammatory Drug (NSAm) associated dyspepsia? *7th Cochrane Colloquium, 5–9 October 1999, Rome, Italy*.

131. Jeng GT, Scott JR, Burmeister LF. A comparison of meta-analytic results using literature vs individual patient data. Paternal cell immunization for recurrent miscarriage. *JAMA* 1995;**274**:830–6.

132. Macleod MR, O'Collins T, Howells DW, Donnan GA. Pooling of animal experimental data reveals influence of study design and publication bias. *Stroke* 2004;**35**:1203–8.

133. Man-Son-Hing M, Wells G, Lau A. Quinine for nocturnal leg cramps: a meta-analysis including unpublished data. *J Gen Intern Med* 1998;**13**:600–6.

134. McLeod BD, Weisz JR. Using dissertations to examine potential bias in child and adolescent clinical trials. *J Consult Clin Psych* 2004;**72**:235–51.

135. MacLean CH, Morton SC, Ofman JJ, Roth EA, Shekelle PG, Southern California Evidence Based Practice Center. How useful are unpublished data from the Food and Drug Administration in meta-analysis? *J Clin Epidemiol* 2003;**56**:44–51.

136. Whittington CJ, Kendall T, Fonagy P, Cottrell D, Cotgrove A, Boddington E. Selective serotonin reuptake inhibitors in childhood depression: systematic review of published versus unpublished data. *Lancet* 2004;**363**:1341–5.

137. Wallace AE, Neily J, Weeks WB, Friedman MJ. A cumulative meta-analysis of selective serotonin reuptake inhibitors in pediatric depression: did unpublished studies influence the efficacy/safety debate? *J Child Adult Psychop* 2006;**16**:37–58.

138. Batt K, Fox Rushby JA, Castillo Riquelme M. The costs, effects and cost-effectiveness of strategies to increase coverage of routine immunizations in low- and middle-income countries: systematic review of the grey literature. *Bull WHO* 2004;**82**:689–96.

139. Davey-Smith G, Egger M. Meta-analysis: unresolved issues and future developments. *BMJ* 1998;**316**:221–5.

140. Winkmann G, Schlutius S, Schweim HG. Publication languages of impact factor journals and of medical bibliographic databanks. *Klinische Monatsblatter fur Augenheilkunde* 2002;**219**:65–71.

141. Vandenbrouche J. On not being born a native speaker of English. *BMJ* 1989;**298**:1461–2.

142. Herrera AJ. Language bias discredits the peer-review system. *Nature* 1999;**397**:467.

143. Coates R, Sturgeon B, Bohannan J, Pasini E. Language and publication in 'Cardiovascular Research' articles. *Cardiovasc Res* 2002;**53**:279–85.

144. Nylenna M, Riis P, Karlsson Y. Multiple blinded reviews of the same two manuscripts. Effects of referee characteristics and publication language. *JAMA* 1994;**272**:149–51.

145. Moher D, Fortin P, Jadad AR, Juni P, Klassen T, Le Lorier J, *et al.* Completeness of reporting of trials published in languages other than English: implications for conduct and reporting of systematic reviews. *Lancet* 1996;**347**:363–6.

146. Junker CA. Adherence to published standards of reporting: a comparison of placebo-controlled trials published in English or German. *JAMA* 1998;**280**:247–9.

147. Gregoire G, Derderian F, Le Lorier J. Selecting the language of the publications included in a meta-analysis: is there a Tower of Babel bias? *J Clin Epidemiol* 1995;**48**:159–63.

148. Egger M, Zellweger-Zahner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. *Lancet* 1997;**350**:326–9.

149. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;**315**:629–34.

150. Heres S, Wagenpfeil S, Hamann J, Kissling W, Leucht S. Language bias in neuroscience – is the Tower of Babel located in Germany? *Eur Psychiat* 2004;**19**:230–2.

151. Moher D, Pham B, Klassen TP, Schulz KF, Berlin JA, Jadad AR, *et al.* What contributions do languages other than English make on the results of meta-analyses? *J Clin Epidemiol* 2000;**53**:964–72.

152. Pham B, Klassen TP, Lawson ML, Moher D. Language of publication restrictions in systematic reviews gave different results depending on whether the intervention was conventional or complementary. *J Clin Epidemiol* 2005;**58**:769–76. Erratum appears in *J Clin Epidemiol* 2006;**59**:216.

153. Juni P, Holenstein F, Sterne J, Bartlett C, Egger M. Direction and impact of language bias in meta-analyses of controlled trials: empirical study [see comment]. *Int J Epidemiol* 2002;**31**:115–23.

154. Brooks TA. Private acts and public objects: an investigation of citer motivations. *J Am Soc Inform Sci* 1985;**36**:223–9.

155. Shadish WR, Tolliver D, Gray M, SenGupta SK. Author judgements about works they cite: three studies from psychology journals. *Soc Stud Sci* 1995;**25**:477–98.

156. Christensenszalanski JJJ, Beach LR. The citation bias – fad and fashion in the judgment and decision literature. *Am Psychol* 1984;**39**:75–8.

157. Gotzsche PC. Reference bias in reports of drug trials. *Br Med J (Clin Res Ed)* 1987;**295**:654–6.

158. Ravnskov U. Quotation bias in reviews of the diet-heart idea. *J Clin Epidemiol* 1995;**48**:713–9.

159. Hutchison BG, Oxman AD, Lloyd S. Comprehensiveness and bias in reporting clinical trials. Study of reviews of pneumococcal vaccine effectiveness. *Can Fam Physician* 1995;**41**:1356–60.

160. Song F, Landes DP, Glenny AM, Sheldon TA. Prophylactic removal of impacted third molars: an assessment of published reviews. *Br Dent J* 1997;**182**:339–46.

161. Callaham M, Wears RL, Weber E. Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals. *JAMA* 2002;**287**:2847–50.

162. Chapman S, Ragg M, McGeechan K. Citation bias in reported smoking prevalence in people with schizophrenia. *Aust NZ J Psychiat* 2009;**43**:277–82.

163. Kjaergard LL, Gluud C. Citation bias of hepato-biliary randomized clinical trials. *J Clin Epidemiol* 2002;**55**:407–10.

164. Nieminen P, Rucker G, Miettunen J, Carpenter J, Schumacher M. Statistically significant papers in psychiatry were cited more often than others. *J Clin Epidemiol* 2007;**60**:939–46.

165. Schmidt LM, Gotzsche PC. Of mites and men: reference bias in narrative review articles: a systematic review. *J Fam Practice* 2005;**54**:334–8.

166. Robins RW, Craik KH. Is there a citation bias in the judgment and decision literature. *Organ Behav Hum Dec* 1993;**54**:225–44.

167. Johnson C. Repetitive, duplicate, and redundant publications: a review for authors and readers. *J Manip Physiol Ther* 2006;**29**:505–9.

168. Jefferson T. Redundant publication in biomedical sciences: scientific misconduct or necessity? *Sci Eng* 1998;**4**:135–40.

169. Angell M, Relman AS. Redundant publication. *New Engl J Med* 1989;**320**:1212–3.

170. Fulginiti VA. Unfortunately, more on duplicate publication. *Am J Dis Child* 1985;**139**:865–6.

171. Lock S. Repetitive publication: a waste that must stop. *Br Med J* 1984;**288**:661–2.

172. Yank V, Barnes D. Consensus and contention regarding redundant publications in clinical research: cross-sectional survey of editors and authors. *J Med Ethics* 2003;**29**:109–14.

173. Errami M, Hicks JM, Fisher W, Trusty D, Wren JD, Long TC, *et al.* Deja vu: A study of duplicate citations in Medline. *Bioinformatics* 2007;btm574.

174. Tramer MR, Reynolds DJ, Moore RA, McQuay HJ. Impact of covert duplicate publication on meta-analysis: a case study. *BMJ* 1997;**315**:635–40.

175. Gotzsche PC. Multiple publication of reports of drug trials. *Eur J Clin Pharmacol* 1989;**36**:429–32.

176. Huston P, Moher D. Redundancy, disaggregation, and the integrity of medical research. *Lancet* 1996;**347**:1024–6.

177. Vandekerckhove P, O'Donovan PA, Lilford RJ, Harada TW. Infertility treatment: from cookery to science. The epidemiology of randomised controlled trials. *Br J Obstet Gynaecol* 1993;**100**:1005–36.

178. Martin J, Bainbridge D, Cheng D. Impact of duplicate publication in a meta-analysis of off-pump versus on-pump coronary artery bypass surgery [abstract]. *12th Cochrane Colloquium, 2–6 October 2004, Ottawa, Canada*.

179. Ben-Shlomo Y, Davey-Smith G. 'Place of publication' bias? [letter]. *BMJ* 1994;**309**:274.

180. Bero LA, Galbraith A, Rennie D. Sponsored symposia on environmental tobacco-smoke. *JAMA* 1994;**271**:612–7.

181. Penel N, Adenis A. Publication biases and phase II trials investigating anticancer targeted therapies. *Invest New Drug* 2009;**27**:287–8.

182. Pittler MH, Abbot NC, Harkness EF, Ernst E. Location bias in controlled clinical trials of complementary/alternative therapies. *J Clin Epidemiol* 2000;**53**:485–9.

183. Ottenbacher K, DiFabio RP. Efficacy of spinal manipulation/mobilization therapy. A meta-analysis. *Spine* 1985;**10**:833–7.

184. Vickers A, Goyal N, Harland R, Rees R. Do certain countries produce only positive results? A systematic review of controlled trials. *Control Clin Trials* 1998;**19**:159–66.

185. Tang JL, Zhan SY, Ernst E. Review of randomised controlled trials of traditional Chinese medicine. *BMJ* 1999;**319**:160–1.

186. Pan Z, Trikalinos TA, Kavvoura FK, Lau J, Ioannidis JPA. Local literature bias in genetic epidemiology: an empirical evaluation of the Chinese literature [see comment]. *PLoS Med* 2005;**2**:e334.

187. King DA. The scientific impact of nations. *Nature* 2004;**430**:311–6. Erratum appears in *Nature* 2004;**432**:8.

188. Felson DT. Bias in meta analytic research. *J Clin Epidemiol* 1992;**45**:885–92.

189. Dickersin K, Hewitt P, Mutch L, Chalmers I, Chalmers TC. Perusing the literature: comparison of MEDLINE searching with a perinatal trials database. *Control Clin Trials* 1985;**6**:306–17.

190. Gotzsche PC, Lange B. Comparison of search strategies for recalling double-blind trials from Medline. *Dan Med Bull* 1991;**38**:476–8.

191. Adams CE, Power A, Frederick K, Lefebvre C. An investigation of the adequacy of Medline searches for randomized controlled trials (Rcts) of the effects of mental health care. *Psychol Med* 1994;**24**:741–8.

192. Zielinski C. New equities of information in an electronic age. *BMJ* 1995;**310**:1480–1.

193. Nieminen P, Isohanni M. Bias against European journals in medical publication databases. *Lancet* 1999;**353**:1592.

194. Caulfield T. Biotechnology and the popular press: hype and the selling of science. *Trends Biotech* 2004;**22**:337–9.

195. Combs B, Slovic P. Causes of death: biased newspaper coverage and biased judgements. *Journalism Q* 1979;**56**:837–43.

196. Houn F, Bober MA, Huerta EE, Hursting SD, Lemon S, Weed DL. The association between alcohol and breast cancer: popular press coverage of research. *Am J Public Health* 1995;**85**:1082–6.

197. Koren G, Klein N. Bias against negative studies in newspaper reports of medical research. *JAMA* 1991;**266**:1824–6.

198. Wing S, Shy C, Wood J, Wolf S, Cragle D, Frome E. Mortality among workers at Oak Ridge National Laboratory. *JAMA* 1991;**265**:1397–402.

199. Jablon S, Hrubec Z, Boice JJ. Cancer in populations living near nuclear facilities: a survey of mortality nationwide and incidence in two states. *JAMA* 1991;**265**:1403–8.

200. Schwartz LM, Woloshin S, Baczek L. Media coverage of scientific meetings. Too much, too soon? *JAMA* 2002;**287**:2859–63.

201. Woloshin S, Schwartz LM. Media reporting on research presented at scientific meetings: more caution needed. *Med J Aust* 2006;**184**:576–80.

202. Whiteman MK, Cui Y, Flaws JA, Langenberg P, Bush TL. Media coverage of women's health issues: is there a bias in the reporting of an association between hormone replacement therapy and breast cancer? *J Women Health Gen-B* 2001;**10**:571–7.

203. Koper M, Bubela T, Caulfield T, Boon H. Media portrayal of conflicts of interest in herbal remedy clinical trials. *Health Law Rev* 2006;**15**:9–11.

204. Lefebvre C, Manheimer E, Glanville J. Searching for studies. In Higgins J, Green S, editors. *Cochrane*

Handbook for Systematic Reviews of Interventions Version 5.00 (Updated Feb 2008). The Cochrane Collaboration, 2008.

205. Perel P, Roberts I, Sena E, Wheble P, Briscoe C, Sandercock P, *et al.* Comparison of treatment effects between animal experiments and clinical trials: systematic review. *BMJ* 2007;**334**:197.

206. Hackam DG. Translating animal research into clinical benefit. *BMJ* 2007;**334**:163–4.

207. Skolbekken JA. The risk epidemic in medical journals. *Soc Sci Med* 1995;**40**:291–305.

208. Angell M, Kassirer JP. Clinical research – what should the public believe? *New Engl J Med* 1994;**331**:189–90.

209. Taubes G. Epidemiology faces its limits. *Science* 1995;**269**:164–9.

210. Chalmers I. Under-reporting research is scientific misconduct. *JAMA* 1990;**263**:1405–8.

211. Cowley AJ, Skene A, Stainer K, Hampton JR. The effect of lorcainide on arrhythmias and survival in patients with acute myocardial infarction – an example of publication bias. *Int J Cardiol* 1993;**40**:161–6.

212. The Cardiac Arrhythmia Suppression Trial (CAST) Investigators. Preliminary report: effect of encainide and flecainide on mortality in a randomised trial of arrhythmia suppression after myocardial infarction. *N Engl J Med* 1989;**321**:406–12.

213. The Cardiac Arrhythmia Suppression Trial II Investigators. Effect of the antiarrhythmic agent moricisine on survival after myocardial infarction. *N Engl J Med* 1992;**327**:227–33.

214. Topol EJ. Failing the public health – rofecoxib, Merck, and the FDA. *N Engl J Med* 2004;**351**:1707–9.

215. Curfman GD, Morrissey S, Drazen JM. Expression of concern: Bombardier *et al.*, 'Comparison of upper gastrointestinal toxicity of rofecoxib and naproxen in patients with rheumatoid arthritis.' *N Engl J Med* 2000;**343**:1520–8.[See editorial in *N Engl J Med* 2005;**353**:2813–4.]

216. Bombardier C, Laine L, Burgos-Vargas R, Davis B, Day R, Ferraz MB, *et al.* Response to expression of concern regarding VIGOR study. *N Engl J Med* 2006;**354**:1196–9.

217. Curfman GD, Morrissey S, Drazen JM. Expression of concern reaffirmed. *N Engl J Med* 2006;**354**:1193.

218. Brown TJ, Hooper L, Elliott RA, Payne K, Webb R, Roberts C, *et al.* A comparison of the cost-effectiveness of five strategies for the prevention of non-steroidal anti-inflammatory drug-induced gastrointestinal toxicity: a systematic review with economic modelling. *Health Technol Assess* 2006;**10**:iii–iv, xi–xiii, 1–183.

219. Psaty BM, Kronmal RA. Reporting mortality findings in trials of rofecoxib for Alzheimer disease or cognitive impairment: a case study based on documents from rofecoxib litigation. *JAMA* 2008;**299**:1813–7.

220. Garland EJ. Facing the evidence: antidepressant treatment in children and adolescents. *Can Med Assoc J* 2004;**170**:489–91.

221. Jureidini JN, Doecke CJ, Mansfield PR, Haby MM, Menkes DB, Tonkin AL. Efficacy and safety of antidepressants for children and adolescents. *BMJ* 2004;**328**:879–83.

222. Kondro W, Sibbald B. Drug company experts advised staff to withhold data about SSRI use in children. *Can Med Assoc J* 2004;**170**:783.

223. Dyer O. GlaxoSmithKline faces US lawsuit over concealment of trial results. *BMJ* 2004;**328**:1395.

224. Ioannidis JP. Effectiveness of antidepressants: an evidence myth constructed from a thousand randomized trials? *Philos Ethics Humanit Med* 2008;**3**:14.

225. Godlee F, Dickersin K. Bias, subjectivity, chance, and conflict of interest in editorial decisions. In Godlee F, Jefferson T, editors. *Peer review in health sciences*. London: BMJ Books; 1999. pp. 57–78.

226. MacCoun R. Biases in the interpretation and use of research results. *Annu Rev Psychol* 1998;**49**:259–87.

227. Cain DM, Detsky AS. Everyone's a little bit biased (even physicians). *JAMA* 2008;**299**:2893–5.

228. Dickersin K, Chan S, Chalmers TC, Sacks HS, Smith H, Jr. Publication bias and clinical trials. *Control Clin Trials* 1987;**8**:343–53.

229. Rotton J, Foos PW, Vanmeek L, Levitt M. Publication practices and the file drawer problem – a survey of published authors. *J Soc Behav Pers* 1995;**10**:1–13.

230. Weber EJ, Callaham ML, Wears RL, Barton C, Young G. Unpublished research from a medical specialty meeting: why investigators fail to publish. *JAMA* 1998;**280**:257–9.

231. Blumenthal D, Campbell EG, Anderson MS, Causino N, Louis KS. Withholding research results in academic life science. *JAMA* 1997;**277**:1224–8.

232. Camacho LH, Bacik J, Cheung A, Spriggs DR. Presentation and subsequent publication rates of phase I oncology clinical trials. *Cancer* 2005;**104**:1497–504.

233. Hartling L, Craig WR, Russell K, Stevens K, Klassen TP. Factors influencing the publication of randomized controlled trials in child health research [see comment]. *Arch Pediat Adol Med* 2004;**158**:983–7.

234. Hopewell S, Clarke M. Methodologists and their methods. Do methodologists write up their conference presentations or is it just 15 minutes of fame? *Int J Technol Assess* 2001;**17**:601–3.

235. Machan C, Ammenwerth E, Bodner T. Publication bias in medical informatics evaluation research: is it an issue or not? *St Heal T* 2006;**124**:957–62.

236. Sprague S, Bhandari M, Devereaux PJ, Swiontkowski MF, Tornetta P III, Cook DJ, *et al.* Barriers to full-text publication following presentation of abstracts at annual orthopaedic meetings. *J Bone Joint Surg Am* 2003;**85**-a(1):158–63.

237. Vuckovic Dekic L, Gajic Veljanoski O, Jovicevic Bekic A, Jelic S. Research results presented at scientific meetings: to publish or not? *Arch Oncol* 2001;**9**:161–3.

238. Shakiba B, Salmasian H, Yousefi-Nooraie R, Rohanizadegan M. Factors influencing editors' decision on acceptance or rejection of manuscripts: the authors' perspective. *Arch Iran Med* 2008;**11**:257–62.

239. Chalmers TC, Frank CS, Reitman D. Minimizing the three stages of publication bias. *JAMA* 1990;**263**:1392–5.

240. Rothwell PM, Slattery J, Warlow CP. A systematic review of the risks of stroke and death due to endarterectomy for symptomatic carotid stenosis. *Stroke* 1996;**27**:260–5.

241. Frank E. Authors' criteria for selecting journals. *JAMA* 1994;**272**:163–4.

242. McCambridge J. A case study of publication bias in an influential series of reviews of drug education. *Drug Alcohol Rev* 2007;**26**:463–8.

243. Tobler N. Meta-analysis of 143 adolescent drug prevention programs: quantitative outcome results of program participants compare to a control or comparison group. *J Drug Issues* 1986;**16**:537–67.

244. Tobler N, Stratton H. Effectiveness of school-based drug prevention programs: a meta-analysis of the research. *J Prim Prev* 1997;**18**:71–128.

245. Tobler N, Roona MR, Ochshorn P, Marshall DG, Streke AV, Stackpole KM. School-based adolescent drug prevention programs: 1998 meta-analysis. *J Prim Prev* 2000;**20**:275–336.

246. Lee K, Boyd E, Bero L. A look inside the black box: a description of the editorial process at three leading biomedical journals [abstract]. *12th Cochrane Colloquium, 2–6 October 2004; Ottawa, Canada*.

247. Luty J, Arokiadass SMR, Easow JM, Anapreddy JR. Preferential publication of editorial board members in medical specialty journals. *J Med Ethics* 2009;**35**:200–2.

248. Kupfersmid J, Fiala M. A survey of attitudes and behaviors of authors who publish in psychology and education journals [comments]. *Am Psychol* 1991;**46**:249–50.

249. Kerr S, Tolliver J, Petree D. Manuscript characteristics which influence acceptance for management and social science journals. *Acad Manage J* 1977;**20**:132–41.

250. Radford DR, Smillie L, Wilson RF, Grace AM. The criteria used by editors of scientific dental journals in the assessment of manuscripts submitted for publication. *Br Dent J* 1999;**187**:376–9.

251. Angell M. Negative studies. *N Engl J Med* 1989;**321**:464–6.

252. Kassirer JP, Campion EW. Peer review: crude and understudied, but indispensable. *JAMA* 1994;**272**:96–7.

253. Abby M, Massey MD, Galandiuk S, Polk HC. Peer review is an effective screening process to evaluate medical manuscripts. *JAMA* 1994;**272**:105–7.

254. Zelen M. Guidelines for publishing papers on cancer clinical trials: responsibilities of editors and authors. *J Clin Oncol* 1983;**1**:164–9.

255. Editor. Manuscript guideline. *Diabetologia* 1984;**25**:4A.

256. Davis RM, Mullner M. Editorial independence at medical journals owned by professional associations: a survey of editors. *Sci Eng Ethics* 2002;**8**:513–28.

257. Bailar JC, Patterson K. Journal peer review: the need for a research agenda. *N Engl J Med* 1985;**312**:654–7.

258. Hojat M, Gonnella JS, Caelleigh AS. Impartial judgment by the 'gatekeepers' of science: fallibility

and accountability in the peer review process. *Adv Health Sci Educ* 2003;**8**:75–96.

259. Mahoney MJ. Publication prejudices: an experimental study of confirmatory bias in the peer review system. *Cognitive Ther Res* 1977;**1**:161–75.

260. Ernst E, Resch KL. Reviewer bias – a blinded experimental study. *J Lab Clin Med* 1994;**124**:178–82.

261. Abbot NE, Ernst E. Publication bias: direction of outcome less important than scientific quality. *Perfusion* 1998;**11**:182–4.

262. Ector H, Aubert A, Stroobandt R. Review of the reviewer [editorial]. *Pacing Clin Electrophys* 1995;**18**:1215–7.

263. Blackburn JL, Hakel MD. An examination of sources of peer-review bias. *Psychol Sci* 2006;**17**:378–82.

264. Wager E, Parkin EC, Tamber PS. Are reviewers suggested by authors as good as those chosen by editors? Results of a rater-blinded, retrospective study. *BMC Med* 2006;**4**:13.

265. Rivara FP, Cummings P, Ringold S, Bergman AB, Joffe A, Christakis DA. A comparison of reviewers selected by editors and reviewers suggested by authors [see comment]. *J Pediatr* 2007;**151**:202–5.

266. Link AM. US and non-US submissions: an analysis of reviewer bias. *JAMA* 1998;**280**:246–7.

267. Opthof T, Coronel R, Janse MJ. The significance of the peer review process against the background of bias: priority ratings of reviewers and editors and the prediction of citation, the role of geographical bias. *Cardiovasc Res* 2002;**56**:339–46.

268. Gilbert JR, Williams ES, Lundberg GD. Is there gender bias in JAMA's peer review process? *JAMA* 1994;**272**:139–42.

269. Johansson EE, Risberg G, Hamberg K, Westman G. Gender bias in female physician assessments. Women considered better suited for qualitative research. *Scand J Prim Health Care* 2002;**20**:79–84.

270. Caelleigh AS, Hojat M, Steinecke A, Gonnella JS. Effects of reviewers' gender on assessments of a gender-related standardized manuscript. *4th International Congress on Peer Review in Biomedical Publication, 14–16 September 2001, Barcelona, Spain.*

271. Garfunkel JM, Ulshen MH, Hamrick HJ, Lawson EE. Effect of institutional prestige on reviewers' recommendations and editorial decisions. *JAMA* 1994;**272**:137–8.

272. Epstein WM. Confirmation response bias among social work journals. *Sci Techol Hum Values* 1990;**15**:9–38.

273. Campillo C. Publication bias in two Spanish medical journals. *The International Congress on Biomedical Peer Review and Global Communications, September 1997, Prague, Czech Republic.*

274. Justice AC, Berlin JA, Fletcher SW, Fletcher RH, Goodman SN. Do readers and peer reviewers agree on manuscript quality? *JAMA* 1994;**272**:117–9.

275. Rosenberg SA. Secrecy in medical research. *New Engl J Med* 1996;**334**:392–4.

276. Rennie D. Thyroid storm. *JAMA* 1997;**277**:1238–43.

277. Davidson RA. Source of funding and outcome of clinical trials. *J Gen Intern Med* 1986;**1**:155–8.

278. Rochon PA, Gurwitz JH, Simms RW, Fortin PR, Felson DT, Minaker KL, *et al.* A study of manufacturer-supported trials of nonsteroidal anti-inflammatory drugs in the treatment of arthritis. *Arch Intern Med* 1994;**154**:157–63.

279. Stelfox HT, Chua G, O'Rourke K, Detsky AS. Conflict of interest in the debate over calcium-channel antagonists. *New Engl J Med* 1998;**338**:101–6.

280. Smith R. Beyond conflict of interest: transparency is the key. *BMJ* 1998;**317**:291–2.

281. Opie L. Conflict of interest in the debate over calcium-channel antagonists. *N Engl J Med* 1998;**338**:1696–7; author reply 1697–8.

282. Detsky AS, Stelfox HT. Correspondence: conflict of interest in the debate over calcium-channel antagonists. *N Engl J Med* 1998;**338**:1698.

283. Meltzer JI. Correspondence: conflict of interest in the debate over calcium-channel antagonists. *N Engl J Med* 1998;**338**:1696.

284. Strandgaard S. Correspondence: conflict of interest in the debate over calcium-channel antagonists. *N Engl J Med* 1998;**338**:1697.

285. Bekelman JE, Li Y, Gross CP. Scope and impact of financial conflicts of interest in biomedical research: a systematic review. *JAMA* 2003;**289**:454–65.

286. Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ* 2003;**326**:1167–70.

287. Golder S, Loke YK. Is there evidence for biased reporting of published adverse effects data in

pharmaceutical industry-funded studies? *Brit J Clin Pharmaco* 2008;**66**:767–73.

288. Als Nielsen B, Chen W, Gluud C, Kjaergard LL. Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events? *JAMA* 2003;**290**:921–8.

289. Baker CB, Johnsrud MT, Crismon ML, Rosenheck RA, Woods SW. Quantitative analysis of sponsorship bias in economic studies of antidepressants [review]. *Brit J Psychiat* 2003;**183**:498–506.

290. Bhandari M, Busse JW, Jackowski D, Montori VM, Schunemann H, Sprague S, *et al.* Association between industry funding and statistically significant pro-industry findings in medical and surgical randomized trials [see comment]. *Can Med Assoc J* 2004;**170**:477–80.

291. Buchkowsky SS, Jewesson PJ. Industry sponsorship and authorship of clinical trials over 20 years [see comment]. *Ann Pharmacother* 2004;**38**:579–85.

292. Montgomery JH, Byerly M, Carmody T, Li B, Miller DR, Varghese F, *et al.* An analysis of the effect of funding source in randomized clinical trials of second generation antipsychotics for the treatment of schizophrenia. *Control Clin Trials* 2004;**25**:598–612.

293. Perlis CS, Harwood M, Perlis RH. Extent and impact of industry sponsorship conflicts of interest in dermatology research. *J Am Acad Dermatol* 2005;**52**:967–71.

294. Liss H. Publication bias in the pulmonary/allergy literature: effect of pharmaceutical company sponsorship [see comment]. *Isr Med Assoc J* 2006;**8**:451–4.

295. Ridker PM, Torres J. Reported outcomes in major cardiovascular clinical trials funded by for-profit and not-for-profit organizations: 2000–2005. *JAMA* 2006;**295**:2270–4.

296. Etter J-F, Burri M, Stapleton J. The impact of pharmaceutical company funding on results of randomized trials of nicotine replacement therapy for smoking cessation: a meta-analysis. *Addiction* 2007;**102**:815–22.

297. Huss A, Egger M, Hug K, Huwiler Muntener K, Roosli M. Source of funding and results of studies of health effects of mobile phone use: systematic review of experimental studies. *Environ Health Persp* 2007;**115**:1–4.

298. Lesser LI, Ebbeling CB, Goozner M, Wypij D, Ludwig DS. Relationship between funding source and conclusion among nutrition-related scientific articles. *PLoS Med* 2007;**4**(1):e5.

299. Jorgensen AW, Hilden J, Gotzsche PC. Cochrane reviews compared with industry supported meta-analyses and other meta-analyses of the same drugs: systematic review [review]. *BMJ* 2006;**333**:782–5.

300. Sawka AM, Thabane L. Effect of industry sponsorship on the results of biomedical research [comment]. *JAMA* 2003;**289**:2502–3.

301. Barden J, Derry S, McQuay HJ, Moore RA. Bias from industry trial funding? A framework, a suggested approach, and a negative result. *Pain* 2006;**121**:207–18.

302. Tulikangas PK, Ayers A, O'Sullivan DM. A meta-analysis comparing trials of antimuscarinic medications funded by industry or not. *BJU Int* 2006;**98**:377–80.

303. Ahmer S, Haider II, Anderson D, Arya P. Do pharmaceutical companies selectively report clinical trial data? *Pak J Med Sci* 2006;**22**:338–46.

304. Millstone E, Brunner E, White I. Plagiarism or protecting public health. *Nature* 1994;**371**:647–8.

305. Nathan DG, Weatherall DJ. Academia and industry: lessons from the unfortunate events in Toronto. *Lancet* 1999;**353**:771–2.

306. Nathan DG, Weatherall DJ. Academic freedom in clinical research. *N Engl J Med* 2002;**347**:1368–71.

307. Shuchman M. Drug company threatens legal action over Canadian guidelines. *BMJ* 1999;**319**:1388.

308. Skolnick AA. Drug firm suit fails to halt publication of Canadian Health Technology Report. *JAMA* 1998;**280**:683–4.

309. Wise J. Research suppressed for seven years by drug company. *BMJ* 1997;**314**:1145.

310. McCarthy M. Company sought to block paper's publication (news). *Lancet* 2000;**356**:1659.

311. Dat NV. Letters: Outcomes of a trial of HIV-1 immunogen in patients with HIV infection. *Lancet* 2001;**285**:2191.

312. Goldberg BS, Stricker RB. Letters: Outcomes of a trial of HIV-1 immunogen in patients with HIV infection. *Lancet* 2001;**285**:2192.

313. Gotch FM. Letters: Outcomes of a trial of HIV-1 immunogen in patients with HIV infection. *Lancet* 2001;**285**:2192.

314. Khahn JO, Lagakos S. Letters: Outcomes of a trial of HIV-1 immunogen in patients with HIV infection. *Lancet* 2001;**285**:2193–4.

315. Lurie P, Wolfe SM. Letters: Outcomes of a trial of HIV-1 immunogen in patients with HIV infection. *Lancet* 2001;**285**:2193.

316. Poretz DM. Letters: Outcomes of a trial of HIV-1 immunogen in patients with HIV infection. *Lancet* 2001;**285**:2192–3.

317. Lauritsen K, Kavelund T, Larsen LS, Rask-Madsen J. Withholding unfavourable results in drug company sponsored clinical trials. *Lancet* 1987;**i**:1091.

318. Symmonds M, Matheson NJ, Harnden A. Guidelines on neuraminidase inhibitors in children are not supported by evidence. *BMJ* 2004;**328**:227.

319. van Heteren G. Wyeth suppresses research on pill, programme claims [see comment]. *BMJ* 2001;**322**:571.

320. van Veldhuisen DJ, Poole Wilson PA. The underreporting of results and possible mechanisms of 'negative' drug trials in patients with chronic heart failure. *Int J Cardiol* 2001;**80**:19–27.

321. Wilmshurst P. Policing the drug industry. *Lancet* 1986;**ii**:1280–1.

322. Wilmshurst P. An investigation by the ABPI (Association of the British Pharmaceutical Industry). *Lancet* 1987;**i**:104.

323. Dieppe PA, Ebrahim S, Martin RM, Juni P. Lessons from the withdrawal of rofecoxib. *BMJ* 2004;**329**:867–8.

324. Gottlieb S. Researchers deny any attempt to mislead the public over JAMA article on arthritis drug. *BMJ* 2001;**323**:301.

325. Hrachovec J. Publication bias with cetirizine in atopic dermatitis: safe but ineffective? [comment]. *J Allergy Clin Immun* 2002;**110**:818; author reply 818.

326. Silverstein FE, Faich G, Goldstein GL. Letter: Reporting of 6-month vs 12-month data in a clinical trial of celecoxib. *JAMA* 2001;**286**:2399.

327. Wright JM, Perry JL. Letter: Reporting of 6-month vs 12-month data in a clinical trial of celecoxib. *JAMA* 2001;**286**:2398–9.

328. Applegate WB, Furberg CD, Byington RP, Grimm R, Jr. The Multicenter Isradipine Diuretic Atherosclerosis Study (MIDAS). *JAMA* 1997;**277**:297; author reply 297–8.

329. Lenzer J. Alteplase for stroke: money and optimistic claims buttress the 'brain attack' campaign. *BMJ* 2002;**324**:723–9.

330. Reines SA, Block GA, Morris JC, Liu G, Nessly ML, Lines CR, *et al.* Rofecoxib: no effect on Alzheimer's disease in a 1-year, randomized, blinded, controlled study. *Neurology* 2004;**62**:66–71.

331. Steinman MA, Bero LA, Chren MM, Landefeld CS. Narrative review: the promotion of gabapentin: an analysis of internal industry documents. *Ann Intern Med* 2006;**145**:284–93.

332. Alasbali T, Smith M, Geffen N, Trope GE, Flanagan JG, Jin Y, *et al.* Discrepancy between results and abstract conclusions in industry- vs nonindustry-funded studies comparing topical prostaglandins [see comment]. *Am J Ophthalmol* 2009;**147**:33–38.e2.

333. Dong BJ, Hauck WW, Gambertoglio JG, Gee L, White JR, Bubp JL, *et al.* Bioequivalence of generic and brand-name levothyroxine products in the treatment of hypothyroidism. *JAMA* 1997;**277**:1205–13.

334. CCOHTA. *Canadian Coordinating Office For Health Technology Assessment: Annual Report 1997–1998*. Ottawa: CCOHTA; 1998.

335. Metcalfe S, Burgess C, Laking G, Evans J, Wells S, Crausaz S. Trastuzumab: possible publication bias. *Lancet* 2008;**371**:1646–8.

336. Panahloo Z. Data on neuraminidase inhibitors were made available. *BMJ* 2004;**328**:523.

337. Thal LJ, Ferris SH, Kirby L, Block GA, Lines CR, Yuen E, *et al.* A randomized, double-blind study of rofecoxib in patients with mild cognitive impairment. *Neuropsychopharmacology* 2005;**30**:1204–15.

338. Hedges LV. Estimation of effect size under nonrandom sampling: the effects of censoring studies yielding statistically insignificant mean differences. *J Educ Stat* 1984;**9**:61–85.

339. Lane DM, Dunlap WP. Estimating effect size: bias resulting from the significance criterion in editorial decisions. *Br J Math Statist Psychol* 1978;**31**:107–12.

340. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994;**50**:1088–101.

341. Borm GF, den Heijer M, Zielhuis GA. Publication bias was not a good reason to discourage trials with low power. *J Clin Epidemiol* 2009;**62**:47.e1–10.

342. Berlin JA, Begg CB, Louis TA. An assessment of publication bias using a sample of published clinical trials. *J Am Stat Assoc* 1989;**84**:381–92.

343. Irwig L, Macaskill P, Glasziou P, Fahey M. Meta-analytic methods for diagnostic test accuracy. *J Clin Epidemiol* 1995;**48**:119–30.

344. Newcombe RG. Discussion of the paper by Begg and Berlin: Publication bias: a problem in interpreting medical data. *J R Stat Soc A* 1988;**151**:448–9.

345. Julian D. Meta-analysis and the meta-epidemiology of clinical research. Registration of trials should be required by editors and registering agencies [letter]. *BMJ* 1998;**316**:311.

346. Chalmers I, Altman D. How can medical journals help prevent poor medical research? Some opportunities presented by electronic publishing. *Lancet* 1999;**353**:490–3.

347. *The Lancet*. URL: http://www.thelancet.com.

348. Horton R. Medical editors trial amnesty. *Lancet* 1997;**350**:756.

349. ICMJE. Uniform requirements for manuscripts submitted to biomedical journals: writing and editing for biomedical publication. International Council of Medical Journal Editors, 2007. URL: http://www.icmje.org/.

350. Altman DG, Simera I, Hoey J, Moher D, Schulz K. EQUATOR: reporting guidelines for health research. *Lancet* 2008;**371**:1149–50.

351. Moher D, Simera I, Schulz KF, Hoey J, Altman DG. Helping editors, peer reviewers and authors improve the clarity, completeness and transparency of reporting health research. *BMC Med* 2008;**6**:13.

352. Simera I, Altman DG, Moher D, Schulz KF, Hoey J. Guidelines for reporting health research: the EQUATOR network's survey of guideline authors. *PLoS Med* 2008;**5**(6):e139.

353. Fletcher RH, Fletcher SW. The effectiveness of editorial peer review. *Peer Review in Health Sciences* 1999:45–55.

354. Godlee F, Gale CR, Martyn CN. Effect on the quality of peer review of blinding reviewers and asking them to sign their reports: a randomized controlled trial. *JAMA* 1998;**280**:237–40.

355. Laband DN, Piette MJ. A citation analysis of the impact of blinded peer review. *JAMA* 1994;**272**:147–9.

356. van-Rooyen S, Godlee F, Evans S, Smith R, Black N. Effect of blinding and unmasking on the quality of peer review. A randomized trial. *JAMA* 1998;**280**:234–7.

357. Yankauer A. How blind is blind review? *Am J Public Health* 1991;**81**:843–5.

358. Jefferson T, Alderson P, Wager E, Davidoff F. Effects of editorial peer review. A systematic review. *JAMA* 2002;**287**:2784–6.

359. Cooper RJ, Gupta M, Wilkes MS, Hoffman JR. Conflict of interest disclosure policies and practices in peer-reviewed biomedical journals. *J Gen Intern Med* 2006;**21**:1248–52.

360. Cain DM, Loewenstein G, Moore DA. The dirt on coming clean: perverse effects of disclosing conflicts of interest. *J Legal Stud* 2005;**34**:1–25.

361. Miller FG, Brody H. Viewpoint: professional integrity in industry-sponsored clinical trials. *Acad Med* 2005;**80**:899–904.

362. Bero L. The electronic future: what might an online scientific paper look like in five years' time? *BMJ* 1997;**315**:1692.

363. Fletcher RH, Fletcher SW. The future of medical journals in the western world. *Lancet* 1998;**352**(SII):30–3.

364. Huth EJ. Is the medical world ready for electronic journals? [editorial]. *Online J Curr Clin Trials* 1992(7).

365. Huth EJ. Electronic publishing in the health sciences. *Bull PAHO* 1995;**29**:81–7.

366. Song F, Eastwood A, Gilbody S, Duley L. The role of electronic journals in reducing publication bias. *Med Inform Internet* 1999;**24**:223–9.

367. Sim I, Rennels G. A trial bank model for the publication of clinical trials. *Proceedings of the Annual Symposium on Computer Applications in Medical Care*; 1995. American Medical Informatics Association.

368. Delamothe T, Mullner M, Smith R. Pleasing both authors and readers. A combination of short print articles and longer electronic ones may help us do this. *BMJ* 1999;**318**:888–9.

369. Fox T. *Crisis in communication: the functions and future of medical journals*. London: The Athlone Press; 1965.

370. Pfeffer C, Olsen BR. Editorial: Journal of negative results in biomedicine. *J Negative Results Biomed* 2002;**1**:2.

371. Roberts I. An amnesty for unpublished trials. *BMJ* 1998;**317**:763–4.

372. Abbasi K. Compulsory registration of clinical trials. *BMJ* 2004;**329**:637–8.

373. Al-Marzouki S, Roberts I, Evans S, Marshall T. Selective reporting in clinical trials: analysis of trial protocols accepted by The Lancet [see comment]. *Lancet* 2008;**372**:201.

374. Boissel JP, Ad Hoc Working Party of the International Collaborative Group on Clinical Trial Registries. Position paper and consensus recommendation on clinical trial registries. *Clin Trial Meta-Anal* 1993;**28**:255–66.

375. Verstraete M. International Committee on Thrombosis and Haemostasis, Basil, 8–11 September 1974. Subcommittee on Clinical Investigations: Registry of prospective clinical trials. *Thromb Diath Haemorrh* 1975;**33**:655–63.

376. Easterbrook PJ. Directory of registries of clinical trials. *Stat Med* 1992;**11**:345–423.

377. Haug C, Gotzsche PC, Schroeder TV. Registries and registration of clinical trials. *N Engl J Med* 2005;**353**:2811–2.

378. Zarin DA, Ide NC, Tse T, Harlan WR, West JC, Lindberg DA. Issues in the registration of clinical trials. *JAMA* 2007;**297**:2112–20.

379. Chalmers I. Government regulation is needed to prevent biased under-reporting of clinical trials [comment]. *BMJ* 2004;**329**:462.

380. Chalmers I. From optimism to disillusion about commitment to transparency in the medico-industrial complex [see comment]. *J R Soc Med* 2006;**99**:337–41.

381. Dickersin K, Garcia Lopez F. Regulatory process effects clinical trial registration in Spain. *Control Clin Trial* 1992;**13**:507–12.

382. Chollar S. A registry for clinical trials [news]. *Ann Intern Med* 1998;**128**:701–2.

383. Gulmezoglu AM, Pang T, Horton R, Dickersin K. WHO facilitates international collaboration in setting standards for clinical trial registration. *Lancet* 2005;**365**:1829–31.

384. Ghersi D, Pang T. En route to international clinical trial transparency. *Lancet* 2008;**372**:1531–2.

385. Laine C, Horton R, DeAngelis CD, Drazen JM, Frizelle FA, Godlee F, *et al.* Clinical trial registration: looking back and moving ahead. *Lancet* 2007;**369**:1909–11.

386. Zarin DA, Tse T, Ide NC. Trial registration at ClinicalTrials.gov between May and October 2005. *N Engl J Med* 2005;**353**:2779–87.

387. Krleza Jeric K, Chan AW, Dickersin K, Sim I, Grimshaw J, Gluud C. Principles for international registration of protocol information and results from human trials of health related interventions: Ottawa statement (part 1) *BMJ* 2005;**330**:956–8. Erratum appears in *BMJ* 2005;**330**:1258.

388. Sim I, Detmer DE. Beyond trial registration: a global trial bank for clinical trial reporting. *PLoS Med* 2005;**2**(11):e365.

389. Chalmers I. TGN1412 and The Lancet's solicitation of reports of phase I trials [comment]. *Lancet* 2006;**368**:2206–7.

390. Choi BCK, Frank J, Mindell JS, Orlova A, Lin V, Vaillancourt ADMG, *et al.* Vision for a global registry of anticipated public health studies. *Am J Public Health* 2007;**97**(Suppl. 1):S82–7.

391. Ramsey S, Scoggins J. Commentary: practicing on the tip of an information iceberg? Evidence of underpublication of registered clinical trials in oncology [see comment]. *Oncologist* 2008;**13**:925–9.

392. Abraham J, Lewis G. Secrecy and transparency of medicines licensing in the EU. *Lancet* 1998;**352**:480–2.

393. Bardy AH. Freedom of information. *Lancet* 1998;**352**:1229.

394. Roberts I, Li-Wan-Po A, Chalmers I. Intellectual property, drug licensing, freedom of information, and public health. *Lancet* 1998;**352**:726–9.

395. Mayor S. Opening the lid on open access. *BMJ* 2008;**336**:688–9.

396. Lenzer J. US Congress and European Research Council insist on open access to research results. *BMJ* 2008;**336**:176–7.

397. Craig ID, Plume AM, McVeigh ME, Pringle J, Amin M. Do open access articles have greater citation impact? A critical review of the literature. *J Informetrics* 2007;**1**:239–48.

398. Wren JD. Open access and openly accessible: a study of scientific publications shared via the internet [see comment]. *BMJ* 2005;**330**:1128.

399. Suber P. Open access, impact, and demand. *BMJ* 2005;**330**:1097–8.

400. Eckert CH. Bioequivalence of levothyroxine preparations: industry sponsorship and academic freedom. *JAMA* 1997;**277**:1200.

401. Mello MM, Clarridge BR, Studdert DM. Academic medical centers' standards for clinical-

trial agreements with industry. *N Engl J Med* 2005;**352**:2202–10.

402. Peto R, Collins R, Gray R. Large-scale randomized evidence: large, simple trials and overviews of trials. *J Clin Epidemiol* 1995;**48**:23–40.

403. Cappelleri JC, Ioannidis JPA, Schmid CH, de Ferranti SD, Aubert M, Chalmers TC, *et al.* Large trials vs meta-analysis of smaller trials. How do their results compare? *JAMA* 1996;**276**:1332–8.

404. LeLorier J, Gregoire G, Benhaddad A, Lapierre J, Derderian F. Discrepancies between meta-analyses and subsequent large randomized, controlled trials. *N Engl J Med* 1997;**337**:536–42.

405. Villar J, Carroli G, Belizan JM. Predictive ability of meta-analyses of randomised controlled. *Lancet* 1995;**345**:772–6.

406. Flournoy N, Olkin I. Do small trials square with large ones? *Lancet* 1995;**345**:741–2.

407. Bennett DA, Latham NK, Stretton C, Anderson CS. Capture-recapture is a potentially useful method for assessing publication bias. *J Clin Epidemiol* 2004;**57**:349–57.

408. Fedorowicz Z, Amin F, Eisinga A, Al Sayyad J. Handsearching for 'buried' randomized trials in Bahrain medical journals [abstract]. *13th Cochrane Colloquium, 22–26 October 2005, Melbourne, Australia*.

409. Peinemann F, Sauerland S, Lange S. Identification of unpublished studies contribute to a systematic review on negative pressure wound therapy [abstract]. *14th Cochrane Colloquium, 23–26 October 2006, Dublin, Ireland.*

410. Prentice VJ, Sayers MK, Milan S. Accessibility of trial data to EBM reviews – lessons for systematic reviewers and the pharmaceutical industry. First Symposium on Systematic Reviews: beyond the basics; 1998 Jan 8–9, Oxford, UK.

411. Lusher ALC. European collaboration to identify reports of controlled trials in specialized health care journals published in Western Europe – helping to overcome some of the barriers associated with trial identification The BIOMED Specialized Health Care Journals Project. *9th Cochrane Colloquium, 9–13 October 2001, Lyon, France.*

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

413. McDonald S. Improving access to the international coverage of reports of controlled trials in electronic databases: a search of the Australasian Medical Index. *Health Info Libr J* 2002;**19**:14–20.

414. Bagett R, Chiquette E, Anagnostelis B, Mulrow C. Locating reports of serious adverse drug reactions. *7th Cochrane Colloquium, 5–9 October 1999, Rome, Italy.*

415. Kassai B, Sonie S, Shah NR, Boissel J-P. Literature search parameters marginally improved the pooled estimate accuracy for ultrasound in detecting deep venous thrombosis. *J Clin Epidemiol* 2006;**59**:710–4.

416. Clarke M. Commentary: searching for trials for systematic reviews: what difference does it make? [comment]. *Int J Epidemiol* 2002;**31**:123–4.

417. Sterne JAC, Bartlett C, Jüni P, Egger M. Do we need comprehensive literature searches? A study of publication and language bias in meta-analyses of controlled trials. *3rd Symposium on Systematic Reviews: Beyond the Basics, July 2000 in Oxford.*

418. Robinson KA, Dickersin K. Development of a highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed. *Int J Epidemiol* 2002;**31**:150–3.

419. Kelly JA. Scientific meeting abstracts: significance, access, and trends. *B Med Libr Assoc* 1998;**86**:68–76.

420. Bennett DAJA. FDA: untapped source of unpublished trials. *Lancet* 2003;**361**:1402–3.

421. Song FJ, Fry Smith A, Davenport C, Bayliss S, Adi Y, Wilson JS, *et al.* Identification and assessment of ongoing trials in health technology assessment reviews. *Health Technol Assess* 2004;**8**:1–87.

422. Scott R. Answering the unanswered questions: ongoing trials of statins and antihypertensives in type 2 diabetes. *Acta Diabetol* 2002;**39**(Suppl. 2):S46–51.

423. Chalmers I. Using systematic reviews and registers of ongoing trials for scientific and ethical trial design, monitoring, and reporting. In Egger M, Davey Smith G, Altman DG, editors. *Systematic reviews in healthcare: meta-analysis in context*. London: BMJ Books; 2000. pp. 429–43.

424. Evsenbach G. Use of the world-wide-web to identify unpublished evidence for systematic reviews – the future role of the internet to improve information identification. *7th Cochrane Colloquium, October 1999, Rome, Italy.*

425. Eysenbach G, Tuische J, Diepgen TL. Evaluation of the usefulness of internet searches to identify unpublished clinical trials for systematic reviews. *Chinese J Evidence Based Med* 2002;**2**:196–200.

426. Reveiz L, Andres Felipe C, Egdar Guillermo O. Using e-mail for identifying unpublished and ongoing clinical trials and those published in non-indexed journals [abstract]. *12th Cochrane Colloquium: Bridging the Gaps; 2004 Oct 2–6, Ottawa, Ontario, Canada* 2004:177–178.

427. Reveiz 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):755–758.

428. Kober T. Obtaining information on clinical trials: a challenging dilemma for Cochrane reviewers [abstract]. *11th Cochrane Colloquium, 26–31 October 2003, Barcelona, Spain.*

429. Milton JLSGR. Well-known signatory does not affect response to a request for information from authors of clinical trials: a randomised controlled trial. *9th Cochrane Colloquium, 9–13 October 2001 Lyon, France.*

430. Higgins J, Soornro M, Roberts I, Clarke M. Collecting unpublished data for systematic reviews: a proposal for a randomised trial. *7th Cochrane Colloquium, 5–9 October 1999, Rome, Italy.*

431. Shukla V. The challenge of obtaining unpublished information from the drug industry [abstract]. *11th Cochrane Colloquium, 26–31 October 2003, Barcelona, Spain.*

432. Hadhazy V, Ezzo J, Berman B. How valuable is effort to contact authors to obtain missing data in systematic reviews. *7th Cochrane Colloquium, 5–9 October 1999, Rome, Italy.*

433. Coursol A, Wagner EE. Effect of positive findings on submission and acceptance rates: a note on meta-analysis bias. *Prof Psychol* 1986;**17**:136–7.

434. Greenwald AG. Consequences of prejudice against the null hypothesis. *Psychol Bull* 1975;**82**:1–20.

435. 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**:374–80.

436. Shadish WR, Doherty M, Montgomery LM. How many studies are in the file drawer – an estimate from the family marital psychotherapy literature. *Clin Psychol Rev* 1989;**9**:589–603.

437. Sommer B. The file drawer effect and publication rates in menstrual cycle research. *Psychol Women Quart* 1987;**11**:233–42.

438. McManus R, Wilson S, Delaney B, Fitzmaurice D, Hyde C, Tobias R, *et al*. Review of the usefulness of contacting other experts when conducting

a literature search for systematic reviews. *BMJ* 1998;**317**:1562–3.

439. Higgins J, Altman DG. Assessing risk of bias in included studies. In Higgins J, Green S, editors. *Cochrane handbook for systematic reviews of interventions version 5.0.0 (updated February 2008)*. The Cochrane Collaboration, 2008.

440. Ioannidis JP. Why most published research findings are false: author's reply to Goodman and Greenland. *PLoS Med* 2007;**4**(6):e215.

441. Light RJ, Pillemer DB. *Summing up: the science of reviewing research*. Cambridge, MA; London: Harvard University Press; 1984.

442. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *J Clin Epidemiol* 2008;**61**(10):991–6.

443. Sterne JA, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *J Clin Epidemiol* 2001;**54**:1046–55.

444. Tang JL, Liu JL. Misleading funnel plot for detection of bias in meta-analysis. *J Clin Epidemiol* 2000;**53**:477–84.

445. Hofmeyr GJ, Atallah AN, Duley L. Calcium supplementation during pregnancy for preventing hypertensive disorders and related problems. *Cochrane Database Syst Rev* 2006;3:CD001059.

446. Belizan JM, Villar J, Gonzalez L, Campodonico L, Bergel E. Calcium supplementation to prevent hypertensive disorders of pregnancy. *N Engl J Med* 1991;**325**:1399–405.

447. Levine RJ, Hauth JC, Curet LB, Sibai BM, Catalano PM, Morris CD, *et al*. Trial of calcium to prevent preeclampsia. *N Engl J Med* 1997;**337**:69–76.

448. Crowther CA, Hiller JE, Pridmore B, Bryce R, Duggan P, Hague WM, *et al*. Calcium supplementation in nulliparous women for the prevention of pregnancy-induced hypertension, preeclampsia and preterm birth: an Australian randomized trial. *Aust NZ J Obstet Gynaecol* 1999;**39**:12–18.

449. Lopez-Jaramillo P, Narvaez M, Weigel RM, Yepez R. Calcium supplementation reduces the risk of pregnancy-induced hypertension in an Andes population. *Br J Obstet Gynaecol* 1989;**96**:648–55.

450. Purwar M, Kulkarni H, Motghare V, Dhole S. Calcium supplementation and prevention of pregnancy induced hypertension. *J Obstet Gynaecol Res* 1996;**22**:425–30.

451. Villar J, Repke J, Belizan JM, Pareja G. Calcium supplementation reduces blood pressure during pregnancy: results of a randomized controlled clinical trial. *Obstet Gynecol* 1987;**70**:317–22.

452. Villar J, Abdel-Aleem H, Merialdi M, Mathai M, Ali M, Zavaleta N, *et al*. World Health Organization randomized trial of calcium supplementation among low calcium intake pregnant women. *Am J Obstet Gynecol* 2006;**194**:639–49.

453. Lopez-Jaramillo P, Narvaez M, Felix C, Lopez A. Dietary calcium supplementation and prevention of pregnancy hypertension. *Lancet* 1990;**335**:293.

454. Lopez-Jaramillo P, Delgado F, Jacome P, Teran E, Ruano C, Rivera J. Calcium supplementation and the risk of preeclampsia in Ecuadorian pregnant teenagers. *Obstet Gynecol* 1997;**90**:162–7.

455. Niromanesh S, Laghaii S, Mosavi-Jarrahi A. Supplementary calcium in prevention of pre-eclampsia. *Int J Gynecol Obstet* 2001;**74**:17–21.

456. Sanchez-Ramos L, Briones DK, Kaunitz AM, Delvalle GO, Gaudier FL, Walker KD. Prevention of pregnancy-induced hypertension by calcium supplementation in angiotensin II-sensitive patients. *Obstet Gynecol* 1994;**84**:349–53.

457. Villar J, Repke JT. Calcium supplementation during pregnancy may reduce preterm delivery in high-risk populations. *Am J Obstet Gynecol* 1990;**163**:1124–31.

458. Terrin N, Schmid CH, Lau J. In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *J Clin Epidemiol* 2005;**58**:894–901.

459. Harbord RM, Egger M, Sterne JAC. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Stat Med* 2006;**25**:3443–57.

460. Macaskill P, Walter SD, Irwig L. A comparison of methods to detect publication bias in meta-analysis [see comment]. *Stat Med* 2001;**20**:641–54.

461. Rucker G, Schwarzer G, Carpenter J. Arcsine test for publication bias in meta-analyses with binary outcomes. *Stat Med* 2008;**27**:746–63.

462. Schwarzer G, Antes G, Schumacher M. A test for publication bias in meta-analysis with sparse binary data. *Stat Med* 2007;**26**:721–33.

463. Pham B, Platt R, McAuley L, Klassen TP, Moher D. Is there a 'best' way to detect and minimize publication bias? An empirical evaluation. *Eval Health Prof* 2001;**24**:109–25.

464. Schwarzer G, Antes G, Schumacher M. Inflation of type I error rate in two statistical tests for the detection of publication bias in meta-analyses with binary outcomes. *Stat Med* 2002;**21**:2465–77.

465. Kromrey JD, Rendina Gobioff G. On knowing what we do not know: an empirical comparison of methods to detect publication bias in meta-analysis. *Educ Psychol Meas* 2006;**66**:357–73.

466. Hayashino Y, Noguchi Y, Fukui T. Systematic evaluation and comparison of statistical tests for publication bias. *J Epidemiol* 2005;**15**:235–43.

467. Saveleva E, Selinski S. Meta-analyses with binary outcomes: how many studies need to be omitted to detect a publication bias? *J Toxicol Env Health A* 2008;**71**:845–50.

468. Moreno SG, Sutton AJ, Ades AE, Stanley TD, Abrams KR, Peters JL, *et al*. Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study. *BMC Med Res Methodol* 2009;**9**:2.

469. Sterne J, Egger M, Moher D. Addressing reporting biases. In Higgins J, Green S, editors. *Cochrane handbook for systematic reviews of intervention. version 5.0.0 (updated February 2008)*. The Cochrane Collaboration; 2008. Chapter 10. Available from www.cochrane-handbook.org.

470. Ioannidis JPA, Trikalinos TA. The appropriateness of asymmetry tests for publication bias in meta-analyses: a large survey. *Can Med Assoc J* 2007;**176**:1091–6.

471. Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 2000;**56**:455–63.

472. Taylor S, Tweedie R. Trim and fill: a simple funnel plot based method of adjusting for publication bias in meta-analysis. *Unpublished manuscript.* 1998.

473. Terrin N, Schmid CH, Lau J, Olkin I. Adjusting for publication bias in the presence of heterogeneity. *Stat Med* 2003;**22**:2113–26. Erratum appears in *Stat Med* 2005;**24**:825–6.

474. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. *Stat Med* 2007;**26**:4544–62.

475. Orwin RG. A fail-safe N for effect size in meta-analysis. *J Educ Stat* 1983;**8**:157–9.

476. Klein S, Simes J, Blackburn GL. Total parenteral nutrition and cancer clinical trials. *Cancer* 1986;**58**:1378–86.

477. Ashworth SD, Osburn HG, Callender JC, Boyle KA. The effects of unrepresented studies on the robustness of validity generalization results. *Pers Psychol* 1992;**45**:341–60.

478. Gleser LJ, Olkin I. Models for estimating the number of unpublished studies. *Stat Med* 1996;**15**:2493–507.

479. Rosenberg MS. The file-drawer problem revisited: a general weighted method for calculating fail-safe numbers in meta-analysis. *Evolution* 2005;**59**:464–8.

480. Evans S. Statistician's comments on: 'Fail safe N' is a useful mathematical measure of the stability of results [letter; comment]. *BMJ* 1996;**312**:125.

481. Becker BJ. Failsafe N or file-drawer number. In Rothstein HR, Sutton AJ, Borenstein M, editors. *Publication bias in meta-analysis: prevention, assessment and adjustments.* Chichester, UK: John Wiley & Sons Ltd; 2005. Chapter 7.

482. Copas J, Jackson D. A bound for publication bias based on the fraction of unpublished studies. *Biometrics* 2004;**60**:146–53.

483. Sutton AJ, Song F, Gilbody SM, Abrams KR. Modelling publication bias in meta-analysis: a review. *Stat Methods Med Res* 2000;**9**:421–445.

484. Dear HBG, Begg CB. An approach for assessing publication bias prior to performing a meta-analysis. *Stat Sci* 1992;**7**:237–45.

485. Hedges LV. Modelling publication selection effects in meta-analysis. *Stat Sci* 1992;**7**:246–55.

486. Iyengar S, Greenhouse JB. Selection models and the file drawer problem. *Stat Sci* 1988;**3**:109–35.

487. Vevea JL, Hedges LV. A general linear model for estimating effect size In the presence of publication bias. *Psychometrika* 1995;**60**:419–35.

488. Rust RT, Lehmann DR, Farley JU. Estimating publication bias in metaanalysis. *J Marketing Res* 1990;**27**:220–6.

489. Copas J. What works? selectivity models and meta-analysis. *J R Stat Soc* 1999;**162**:95–109.

490. Baker R, Jackson D. Using journal impact factors to correct for the publication bias of medical studies. *Biometrics* 2006;**62**:785–92.

491. Bowden J, Thompson JR, Burton P. Using pseudo-data to correct for publication bias in meta-analysis. *Stat Med* 2006;**25**:3798–813.

492. Formann AK. Estimating the proportion of studies missing for meta-analysis due to publication bias. *Contemp Clin Trials* 2008;**29**:732–9.

493. Cleary RJ, Casella G. An application of Gibbs sampling to estimation in meta-analysis: accounting for publication bias. *J Educ Behav Stat* 1997;**22**:141–54.

494. Givens GH, Smith DD, Tweedie RL. Publication bias in meta-analysis: a Bayesian data-augmentation approach to account for issues exemplified in the passive smoking debate. *Stat Sci* 1997;**12**:221–50.

495. Larose DT, Dey KK. Modeling publication bias using weighted distributions in a Bayesian framework. *Comput Stat Data An* 1998;**26**:279–302.

496. Silliman NP. Nonparametric classes of weight functions to model publication bias. *Biometrika* 1997;**84**:909–18.

497. Silliman NP. Hierarchical selection models with applications in meta-analysis. *J Am Stat Assoc* 1997;**92**:926–36.

498. Hedges LV, Vevea JL. Estimating effect size under publication bias: Small sample properties and robustness of a random effects selection model. *J Educ Behav Stat* 1996;**21**:299–332.

499. Vevea JL, Woods CM. Publication bias in research synthesis: sensitivity analysis using a priori weight functions. *Psychol Methods* 2005;**10**:428–43.

500. Copas JB, Shi JQ. A sensitivity analysis for publication bias in systematic reviews. *Stat Methods Med Res* 2001;**10**:251–65.

501. Williamson PR, Gamble C. Application and investigation of a bound for outcome reporting bias. *Trials* 2007;**8**:9.

502. Ioannidis JP, Trikalinos TA. An exploratory test for an excess of significant findings. *Clin Trials* 2007;**4**:245–53.

503. Whitehead A, Whitehead J. A general parametric approach to the meta-analysis of randomized clinical trials. *Stat Med* 1991;**10**:1665.

504. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Contro Clin Trials* 1986;**7**:177–88.

505. Berlin JA, Laird NM, Sacks HS, Chalmers TC. A comparison of statistical methods for combining event rates from clinical trials. *Stat Med* 1989;**8**:141–51.

506. Jackson D. The implications of publication bias for meta-analysis' other parameter. *Stat Med* 2006;**25**:2911–21.

507. Greenland S. Invited commentary: a critical look at some popular meta-analytic methods. *Am J Epidemiol* 1994;**140**:290–6.

508. Jadad AR, Cook DJ, Jones A, Klassen TP, Tugwell P, Moher M, *et al.* Methodology and reports of systematic reviews and meta-analyses. A comparison of Cochrane reviews with articles published in paper-based journals. *JAMA* 1998;**280**:278–80.

509. French SD, McDonald S, McKenzie JE, Green SE. Investing in updating: how do conclusions change when Cochrane systematic reviews are updated? *BMC Med Res Methodol* 2005;**5**:33.

510. Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis [see comment]. *Ann Intern Med* 2007;**147**:224–33.

511. Brok J, Thorlund K, Gluud C, Wetterslev J. Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses. *J Clin Epidemiol* 2008;**61**:763–9.

512. Borm GF, Donders AR. Updating meta-analyses leads to larger type I errors than publication bias. *J Clin Epidemiol* 2009;**62**(8):825–30.

513. Ioannidis J, Trikalinos T. Appropriateness of asymmetry tests for publication bias in meta-analysis: large-scale survey [abstract]. *14th Cochrane Colloquium; 23–26 October 2006, Dublin, Ireland.*

514. Lau J, Ioannidis JP, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. *BMJ* 2006;**333**:597–600.

515. Galbraith RF. A note on graphical presentation of estimated odds ratios from several clinical trials. *Stat Med* 1988;**7**:889–94.

516. Irwig L, Macaskill P, Berry G, Glasziou P. Bias in meta-analysis detected by a simple, graphical test. Graphical test is itself biased [comment]. *BMJ* 1998;**316**:470; author reply 470–1.

517. McDonald S, Taylor L, Adams C. Searching the right database. A comparison of four databases for psychiatry journals. *Health Libr Rev* 1999;**16**:151–6.

518. Laupacis A. Methodological studies of systematic reviews: is there publication bias? *Arch Intern Med* 1997;**157**:357.

519. Dubben H-H, Beck-Bornholdt H-P. Systematic review of publication bias in studies on publication bias [see comment]. *BMJ* 2005;**331**:433–4.

520. Song F. Review of publication bias in studies on publication bias: studies on publication bias are probably susceptible to the bias they study [comment]. *BMJ* 2005;**331**:637–8.

521. Beveridge WIB. *The art of scientific investigation.* London: Mercury Books; 1961.

522. Meslin E. Toward an ethic in dissemination of new knowledge in primary care research. In Dunn E, editor. *Disseminating research/changing practice.* Thousand Oaks: Sage; 1994.

523. Anderson SJ. Some thoughts on the reporting of adverse events in phase II cancer clinical trials. *J Clin Oncol* 2006;**24**:3821–2.

524. Sutton AJ, Duval SJ, Tweedie RL, Abrams KR, Jones DR. Empirical assessment of the effect of publication bias on meta-analyses. *BMJ* 2000;**320**:1574–7.

525. Jennions MD, Moller AP. Publication bias in ecology and evolution: an empirical assessment using the 'trim and fill' method. *Biol Rev Camb Philos* 2002;**77**:211–22.

526. Mieog S, Ghersi D. Is timing important in systematic reviews of interventions for breast cancer? [abstract]. *13th Cochrane Colloquium; 22–26 October 2005, Melbourne, Australia.*

527. Bhandari M, Guyatt GH, Tong D, Adili A, Shaughnessy SG. Reamed versus nonreamed intramedullary nailing of lower extremity long bone fractures: a systematic overview and meta-analysis. *J Orthop Trauma* 2000;**14**:2–9.

528. Horn J, Limburg M. Calcium antagonists for ischemic stroke: a systematic review. *Stroke* 2001;**32**:570–6.

529. Marinovich L, Ghersi D, Lord S. Data maturity and systematic reviews of new health technologies [abstract]. *12th Cochrane Colloquium, 2–6 October 2004, Ottawa, Canada.*

530. Martin JLR, Perez V, Sacristan M, Alvarez E. Is grey literature essential for a better control of publication bias in psychiatry? An example from three meta-analyses of schizophrenia. *Eur Psychiatr* 2005;**20**:550–3.

531. Maguire M, Hutton J, Marson A. The impact of including grey and non-English literature on the combined effect estimates from non-randomized add-on anti-epileptic studies [abstract]. *14th Cochrane Colloquium, 23–26 October 2006, Dublin, Ireland.*

532. Moller AP, Thornhill R, Gangestad SW. Direct and indirect tests for publication bias: asymmetry and sexual selection. *Anim Behav* 2005;**70**:497–506.

533. Bartlett C, Sterne JA, Juni P, Egger M. Can we ignore the non-English literature? A study of language bias in meta-analyses of controlled trials [abstract]. *16th Annual Meeting of the International Society of Technology Assessment in Health Care, 18–21 June 2000, The Hague, The Netherlands.*

534. Frame SM. Rapid responses to: Article about Canadian guidelines on proton pump inhibitors was misleading. *BMJ* 2000;**320**:1143.

535. Hrachovec JB, Mora M. Letter: Reporting of 6-month vs 12-month data in a clinical trial of celecoxib. *JAMA* 2001;**286**:2398.

536. Reicin A, Shapiro D. Response to expression of concern regarding VIGOR study. *N Engl J Med* 2006;**354**:1196–9.

537. Perez EA, Suman VJ. Lack of publication bias related to results from trastuzumab study. *Lancet* 2008;**372**:626–7.

Appendix 1
Search strategies for electronic databases

MEDLINE
Search terms
1   *publications/
2   exp publication bias/
3   (bias$adj3 (publication$or disseminat$or language$or reporting or grey or gray or citation$or time delay or time lag or national or country or location or conference or abstract or duplicat$or multiple publication$)).tw,ot.
4   ((reference$or database$or index$) adj2 bias$).tw,ot.
5   (file adj drawer$).tw,ot.
6   (time adj2 (completion or publication)).tw,ot.
7   unpublished research.tw,ot.
8   (fail$adj2 publish$).tw,ot.
9   Or/1–8
10   Limit 9 to yr='1998–2007'

Cochrane Methodology Register
Search terms
1   'Study identification' or
2   'Information retrieval' or
3   'Unpublished data' or
4   'Missing data' or
5   'Updating and cumulative meta-analysis' or
6   'Prospective meta-analysis' or
7   'Small study effects' or
8   'Small trial bias' or
9   'Funding' or
10   'Outcome reporting bias' or
11   'Bias in review' or
12   (bias* NEAR/3 (publication* or disseminat* or language* or reporting or grey or gray or citation* or time delay or time lag or national or country or location or conference or abstract or reference* or index* or database* or duplicat* or multiple publication*)) in Title, Abstract or Keywords
13   from 1998 to 2007 in Cochrane Methodology Register

EMBASE

Search terms

1 (bias$adj3 (publication$or disseminat$or language$or reporting or grey or gray or citation$or time delay or time lag or national or country or location or conference or abstract or duplicat$or multiple publication$)).tw,ot.

2 ((reference$or database$or index$) adj2 bias$).tw,ot.

3 (file adj drawer$).tw,ot.

4 (time adj2 (completion or publication)).tw,ot.

5 unpublished research.tw,ot.

6 (fail$adj2 publish$).tw,ot.

7 Or/1–6

8 Limit 9 to yr='1998–2007'

AMED

Search terms

1 exp publications/

2 publication bias.tw,ti.

3 (bias$adj3 (publication$or disseminat$or language$or reporting or grey or gray or citation$or time delay or time lag or national or country or location or conference or abstract or duplicat$or multiple publication$)).tw,ti.

4 ((reference$or database$or index$) adj2 bias$).tw,ti.

5 (time adj2 (completion or publication)).tw,ti.

6 unpublished research.tw,ti.

7 (fail$adj2 publish$).tw,ti.

8 or/1–7

9 limit 8 to yr='1998–2007'

CINAHL

Search terms

1 exp publications/

2 publication bias.tw,ti.

3 (bias$adj3 (publication$or disseminat$or language$or reporting or grey or gray or citation$or time delay or time lag or national or country or location or conference or abstract or duplicat$or multiple publication$)).tw,ti.

4 ((reference$or database$or index$) adj2 bias$).tw,ti.

5 (time adj2 (completion or publication)).tw,ti.

6 unpublished research.tw,ti.

7 (fail$adj2 publish$).tw,ti.

8 or/1–7

9 limit 8 to yr='1998–2007'

MEDLINE search strategy – part II	
	Search terms
1	publication bias
2	reporting bias
3	OR/1–2
4	systematic review
5	meta-analysis
6	OR/4–5
7	loattrfull text [sb] AND loattrfree full text [sb] AND has abstract[text]
8	2000 [PDAT]: 2008 [PDAT]
9	English [lang]
10	AND/7–9

Appendix 2

Data extraction sheet for empirical studies

Author (year): _____
Title: _____
Source: _____
Study design & objectives: _____
Issues:

o	Existence/identifying	o	Causes/risk factors
o	Consequence	o	Other: _____

Categories:

o	Non-publication	o	Incomplete publication
o	Limited accessibility	o	Other: _____

Specific bias:

o	Publication bias	o	Grey literature bias
o	Language bias	o	Reporting bias
o	Abstract bias	o	Time delay bias
o	Database index bias	o	Citation bias
o	Duplicate bias	o	Media attention bias
o	Other: _____		

Areas:

o   general healthospecific health (e.g., obesity)   o   other: _____

Study results: o Significant/important bias   o Non-significant   o Can't tell
Details:

Original authors' conclusions:
_____

Evidence: o Direct   o Indirect
For indirect evidence, stop. For studies with direct evidence, continue:

Scientific rigorousness (hints: prospective or retrospective, sample selection bias)
_____

Sample representativeness (hints: research field, participants, outcomes, interventions; study designs)
_____

Appropriateness of data analysis (hints: consider objectives, available data and methods of data analysis)
_____

Appropriateness of interpretations (hints: limitations of the study should be taken into consideration)
_____

Overall study quality:
       o High (hint: without considerable concern on study validity)
       o Moderate (hint: with some concern on study validity)
       o Low (hint: with considerable concern on study validity)
       o Can't tell

Any other comments:

Appendix 3

Data extraction sheet for methodological studies

Author (year): _____

Title: _____

Source: _____

Study design:_____

Study objectives:

o New method   o Established method   o Evidence of usefulness/limitations

Methods:

o	Study registration	o	Literature search
o	Funnel plot	o	Statistical/modelling
o	Updating reviews	o	Publication process
o	Research ethics/policy	o	Confirmatory studies
o	Other: _____		

Purpose:

o Preventing bias   o Reducing bias   o Detecting bias   o Adjusting bias   o Other:_____

Stage of literature review:

   o Before literature review   o In literature review   o After literature review

What dissemination bias the method is relevant:

o	Publication bias	o	Grey literature bias
o	Language bias	o	Reporting bias
o	Abstract bias	o	Time delay bias
o	Database index bias	o	Citation bias
o	Duplicate bias	o	Media attention bias
o	Other: _____		

Main findings and conclusions:

_____

Resources required to use the method:

_____

Reviewer's commentary (study's validity, scientific rigorousness, method's usefulness and limitations, any empirical evidence provided):

Appendix 4

Data extraction sheet – systematic reviews of treatment

Data Extraction Sheet – Systematic Reviews of Treatment

Author (year): _____ Reviewer: _____Date: _____

Title: _____
Source: _____
Review Characteristics:
Participants: _____
Interventions (and control): _____
□ Drug   □ Surgical   □ Educational/behavioural   □ Alternative □ Other

Outcomes: _____
Total no of outcomes evaluated: _____Primary outcome(s) defined? □ Yes   □ No

Type of reviews: □ Narrative   □ Meta-analysis;   How many? _____

Designs of included studies:
□   RCTs/CCTs   (Study =   ; patients =   )
□   Other_____   (Study =   ; patients =   )

How were differences between studies investigated? □   NA
□   Narrative   □   Meta-regression / Subgroup
□   Statistical: $I^2 =$ ____ $P =$ ____
□   Other: _____

Sources searched to identify studies: □   Not stated
□   MEDLINE □   EMBASE   □   Psychlit   □   Cochrane   □ References
□   CINAHL   □   Handsearch □   Experts/authors   □   Company
□   Proceedings   □   Other   _____

Language restriction: □   Unclear   □   No   □   Yes, what language(s) included:
_____

Non-English language studies:
Searched: □   Unclear □   No   □   Yes If yes, search methods: _____
Included: □   Unclear □   No   □   Yes, how many _____
For   □   Main analysis;   □   sensitivity analysis?
Grey literature/conference abstracts:
Searched: □   Unclear □   No   □   Yes   If yes, search methods: _____
Included: □   Unclear □   No   □   Yes, how many _____
For   □   main analysis;   □   sensitivity analysis?
Other unpublished studies:
Searched: □   Unclear □   No   □   Yes   If yes, search methods: _____
Included: □   Unclear □   No   □   Yes, how many _____
For   □   main analysis;   □   sensitivity analysis?

Are all the relevant trials included in meta-analyses? □   Unclear   □   Yes
□ No, how many _____

Outcome reporting bias considered?   □   No   □   Yes

Were there missing outcome data?   □   Unclear   □   No   □   Yes

If yes, methods used to deal with missing data on outcomes: _____

Methods used for dealing with publication bias:
□   Not used   □   Identify unpublished studies
□   Prospective register   □   Fail-safe N

□ Funnel plot	□ Rank correlation (Begg method)
□ Egger's method	□ Large scale trials
□ Modelling	□ Other: _____

Details:

Issue of publication bias discussed?   □   No   □   Yes

Evidence on publication and related bias: □   Not available   □ Available
If available, details (such as results of published trials versus unpublished trials; or shape of Funnel plot or related methods)

Meta-analysis results:
□   Not applicable (no meta-analysis)
□   Statistically significant (at least one primary outcomes)
□   Non significant (primary outcomes)

Authors' conclusion:
□   Significant/positive (At least one intervention recommended; or significant difference found)
□   Non-significant/not important (No intervention recommended, or no significant differences)
□   Unclear (Not able to judge; neither positive nor negative; lack of evidence)
_____

Assessor's Judgement:

Efforts to minimise publication bias   □ Sufficient □ Partial   □ Insufficient

Risk of Publication Bias (Considering the possibility that authors' conclusion might be wrong because of possible publication and related biases)**:**
     □ High   □ Moderate   □ Low
Reasons, if any:

_____

Any other comments:

Appendix 5

Main characteristics of inception cohort studies of publication bias

Study	Cohort types Speciality	Study design Follow-up	Verification of publication status and results of unpublished studies	Publication rate	Definition of study results and other notes
Bardy 1998[26]	Clinical trials on medicinal products notified to the National Agency for Medicines in 1987, Finland Mixed speciality	Clinical trials Follow-up: 5–6 years	MEDLINE searched for publications Questionnaires sent to trial sponsors for trial results	Positive 47% (52/111) Inconclusive 33% (11/33) Negative 11% (5/44)	Positive: the drug better (or equivalent to in equivalent trials) than comparators, or the objective of the study supported or confirmed Inconclusive: exploratory studies or non-comparative or the risk–benefit was inconclusive Publication: published in journals included in MEDLINE
Cronin and Sheldon 2004[27]	Studies sponsored by the NHS R&D programme (the North Thames Regional Office) from 07/1995 to 12/1998 Mixed speciality	Mixed. Including quantitative (47%) and qualitative (53%) research Follow-up: > 2 years	Questionnaires sent to investigators 17% failed to respond	Quantitative or qualitative effect ($n = 70$) published in peer-reviewed journals: Showed an effect 76% (26/34) No effect 64% (23/36)	Methods used in Dickersin et al.[20] were adopted to classify findings
Decullier and Chapuis 2006[29]	Protocols submitted for funding to the G. Lyon regional scientific committee in 1997 Mixed speciality	Mixed: Ob = 51% CT = 25% Follow-up: 8 years	Questionnaires sent to investigators, up to 3 times 20% failed to respond	Completed studies: Important 70% (26/37) Less important 60% (6/10)	Investigators rated the importance of results from 1 to 10. Important results were those > 5
Decullier et al. 2005[28]	Biomedical research protocol approved by French RECs in 1994 Mixed: biomedical research	Mixed: Ob = 13% Exp (CT) = 87%. Follow-up: 5–7 years	Questionnaires sent to investigators, or from REC databases 31% failed to respond	Confirmatory 69% (129/188) Invalidating 19% (3/16) Inconclusive 32% (14/44)	Confirmatory: results confirming study hypothesis Invalidating: results invalidating study hypothesis Inconclusive: not confirming or invalidating
Dickersin and Min 1993[21]	NIH 1979 funded clinical trials that were completed by 1988 Mixed speciality	Clinical trials: CT = 100% Follow-up: 9 years	Contacting and telephone interview of investigators 26% failed to respond	Sig/important 98% (121/124) Non-significant 85% (63/74)	Significant results: $p < 0.05$ or deemed to be of 'great importance' Non-significant results: all other results

Study	Cohort types Speciality	Study design Follow-up	Verification of publication status and results of unpublished studies	Publication rate	Definition of study results and other notes
Dickersin et al. 1992[20]	Studies approved by IRBs at Johns Hopkins Health Institutions up to the end of 1980 Mixed speciality	Mixed: Ob = 37%/85% Exp = 17%/9% CT = 46%/6% Follow-up: > 7 years	Telephone interview of investigators 30% failed to provide adequate data	*Medicine & hospital:* Sig/important 89% (184/208) Non-significant 69% (93/134) *Public health:* Sig/important 71% (75/106) Non-significant 58% (38/66) *Clinical trials (both centres):* Sig/important 87% (84/96) Non-significant 72% (52/72)	Significant results: $p < 0.05$ or results considered to be of great importance Non-significant results: all other results Risk of publication bias may be underestimated by excluding studies due to lack of information Unpublished data for clinical trials obtained from Hopewell et al. 2009[83]
Easterbrook et al. 1991[22]	Studies approved by the Central Oxford REC between 1984 and 1987 Mixed speciality	Mixed: Ob = 30% Exp = 18% CT = 52% Follow-up: 3–6 years	Questionnaires sent to investigators, followed by a telephone interview 8% failed to respond or provide adequate data	*Fully published:* Significant 60% (93/154) Non-significant trend 35% (12/34) No difference 34% (33/97) *Published or presented:* Significant 85% (131/154) Non-significant trend 65% (22/34) No difference 56% (54/97)	Significant results: $p < 0.05$ Non-significant trend: difference with a p value of ≥ 0.05 Null: no difference Examined factors associated with publication (but not necessarily publication bias)
Ioannidis 1998[23]	RCTs conducted by two trialist groups (sponsored by the NIH) from 1986 to 1996 AIDS/HIV	Clinical trial: CT = 100% Follow-up: 1–10 years	Information obtained from a database of HIV trials sponsored by NIH. Supplemental data from investigators and staff responsible for the protocols	Positive 74% (20/27) Non-positive 41% (16/39)	Positive: statistically significant ($p < 0.05$) in favour of an experimental arm Non-positive: significantly in favour of the control arm or non-significant. The focus of the study was time lag bias. Data obtained from Hopewell et al.[113]
Misakian and Bero 1998[30]	Research on passive smoking funded by 76 organisations between 1981 and 1995 Health effects of passive smoking	Mixed: Exp = 23% Obs = 77% Follow-up: median 5 years	Semistructured telephone interview of investigators 17% failed to respond	Significant 85% (28/33) Non-significant 86% (18/21) Mixed 14% (1/7)	Statistically significant: $p \leq 0.05$ Mixed results: multiple primary outcomes at least one was statistically significant Cox regression analysis was used to estimate hazard ratio

Study	Cohort types Speciality	Study design Follow-up	Verification of publication status and results of unpublished studies	Publication rate	Definition of study results and other notes
Stern and Simes 1997[24]	Studies submitted to Royal Prince Alfred Hospital REC between 1979 and 1988 Mixed speciality	Mixed: Ob = 22% Exp = 22% CT = 56% (details available) Follow-up: 3–12 years	Questionnaires sent to investigators 30% failed to respond	*Quantitative studies* Significant 68% (99/146) Non-significant trend 20% (4/20) No difference 44% (23/52) *Qualitative studies* Striking 70% (19/27) Important/definite 59% (35/59) Negative/unimportant 53% (9/17) *Clinical trials* (n = 167) Quantitative trials: Significant 72% (55/76) Non-significant trend 20% (3/15) Null 38% (15/39) Qualitative trials: Striking 50% (3/6) Important/definite 61% (11/18) Negative/unimportant 69% (9/13)	*Quantitative studies* Significant: $p < 0.05$ Non-significant trend: $p \geq 0.05 \leq 0.10$ No difference: $p \geq 0.10$ Classification of qualitative studies based on principal investigators' judgement With data on time delayed publication. Qualitative studies similarly vulnerable to publication bias
Wormald *et al.* 1997[31]	Randomised trials processed through the Pharmacy of Moorfields Eye Hospital since 1963 Eye health	RCTs Follow-up: > 2 years	Retrospective review	Significant 93% (14/15) Non-significant 71% (15/21)	Significant: $p < 0.05$ Non-significant: $p \geq 0.05$ Published as a brief abstract, additional data from Dwan *et al.*[41]
Zimpel and Windeler 2000[32]	140 medical theses on complementary medical subjects Complementary medicine	Mixed Follow-up: > 5 years	Literature search and contacting investigators Response rate unclear	Positive 40% (43/107) Negative 28% (15/53)	Full publication in German, information obtained mainly from the abstract

CT, clinical trial; RCT, randomised controlled trial; Ob, observational study; Exp, experimental study.

Appendix 6

Main characteristics of included regulatory cohort studies of publication bias: trials submitted to regulatory authorities

Study	Cohort types Speciality	Study design Follow-up	Verification of publication status and results of unpublished studies	Publication rate	Definition of study results and other notes
Lee et al. 2008[42]	Trials supporting new drugs approved by FDA between 1998 and 2000 Mixed speciality	Clinical trials Follow-up: >5 years	Searches of PubMed and other databases Trial results obtained from FDA files	Statistically significant 66% (285/432) Not statistically significant 36% (52/144)	Statistical significance (primary outcome): $p<0.05$ or if an equivalency trial $p>0.05$ or a CI excluding the prespecified difference
Melander et al. 2003[43]	RCTs of five SSRIs submitted to the Swedish drug regulatory authority for marketing approval Mental: depression	Clinical trials Follow-up: unclear	Literature search for publications Study results based on submitted material	*Stand-alone or pooled publications:* Significant 100% (21/21) Non-significant 81% (17/21) *Stand-alone publications:* Significant 90% (19/21) Non-significant 29% (6/21)	Also provided data on duplicate, and selective outcome reporting bias
Rising et al. 2008[44]	Efficacy trials supporting new drug applications approved by FDA from 2001 to 2002 Mixed speciality	Clinical trials Follow-up: >4 years	Search of PubMed and Cochrane Library and contacting sponsors and authors Trial results obtained from FDA files	*Primary outcomes* Favourable 82% (102/124) Not favourable 66% (19/29) Unknown 60% (6/10) *Conclusion* Favourable 79% (90/114) Not favourable 64% (7/11) Unknown 57% (4/7)	Favourable: statistically significantly in favour of the new drug or equivalence was found Note: pooled publication not reported separately was considered 'not published'
Turner et al. 2008[45]	Trials of 12 antidepressants, submitted to FDA between 1987 and 2004 Mental: depression	Phase 2 and 3 clinical trials Follow-up: >2 years	Literature search and contacting sponsors for publications Trial results obtained from FDA files	Positive 97% (37/38) Questionable 50% (6/12) Negative 33% (8/24)	Trial results were classified according to the FDA's regulatory decisions

Appendix 7

Main characteristics of abstract cohort studies of publication bias: abstracts presented at conferences

Study	Cohort types Speciality	Study design Follow-up	Verification of publication status and results of unpublished studies	Publication rate	Definition of study results and other notes
Akbari-Kamrani et al. 2008[56]	Presented meeting abstracts Laser medicine and surgery	Clinical trials Follow-up: > 3 years	Literature search for full publications	Significant ($p < 0.05$): 51% (23/45) Non-significant 50% (22/44) Positive 43% (59/137) Not-positive 38% (3/8)	Positive: if stated the intervention has some beneficial effect. Negative: against the intervention. Equivocal: no statement about the effectiveness or the comparison groups considered the same
Brazzelli et al. 2009[57]	Presented meeting abstracts Stroke	Studies of diagnostic tests Follow-up: > 2 years	Literature search and contacting authors for full publications	*Clinical utility* Accurate 76% (107/141) Possibly or non-informative 68% (13/19) *Sensitivity (median 0.91)* Above median 77% (38/49) Below median 74% (34/46) Not given 75% (49/65) *Specificity (median 0.91)* Above median 71% (30/42) Below median 73% (27/37) Not given 79% (64/81)	Accurate: diagnostic accuracy of the test was high enough to recommend its use in clinical practice. Possibly useful: having a good sensitivity (and vice versa). Non-informative: the accuracy of the test was not good enough to recommend its use in clinical practice or equivalent to or not better than that of an existing alternative test
Callaham et al. 1998[47]	Submitted meeting abstracts Emergency medicine	Mixed: CT = 26% Follow-up: 5 years	Literature search and contacting authors for full publications	Positive 50% (77/153) Not positive 49% (36/74)	Positive results: beneficial results or $p < 0.05$. Publication rates from CMRD by Scherer et al.[55]
Castillo et al. 2002[58]	Presented meeting abstracts Anaesthesiology	Mixed: Ob = 69% RCT = 31%. Follow-up: 4–5 years	Literature search for full publications	Significant 44% (160/361) Non-significant 41% (23/56)	Significant: $p < 0.05$
Chalmers et al. 1990[48]	Summary trial reports published from 1940 to 1984, included in the Oxford Database of Perinatal Trials Perinatal	Clinical trials: CT = 100%. Follow-up: > 4 years	The Oxford Database of Perinatal Trials was searched to identify full publications	Positive 33% (32/98) Neutral/negative 41% (32/78)	Positive: the test treatment superior to the control Negative: the test treatment potentially harmful Neutral: no real difference

Study	Cohort types Speciality	Study design Follow-up	Verification of publication status and results of unpublished studies	Publication rate	Definition of study results and other notes
Cheng et al. 1998[49]	Abstracts from three international conferences over a 30-year period Cystic fibrosis	Clinical trials: CT = 100%. Follow-up: not reported	Literature search for full publications	Positive 38% (43/113) Negative 33% (14/42)	Positive: authors concluded the test treatment was superior to control or equally effective (in equivalence trials) Publication rates from CMRD by Scherer et al. log-rank tests showed no significant difference in time to publication between 'positive' or 'negative' results ($p = 0.54$)
De Bellefeuille et al. 1992[50]	Submitted meeting abstracts Clinical oncology	Mixed: CT = 48% Follow-up: 5 years	Literature search and contacting authors for full publications	Positive 74% (48/65) Negative 32% (10/31) Neutral/descriptive 56% (57/101)	Positive results: $p < 0.05$ or beneficial to interventions
Delamere and Williams 2005[59]	Conference abstracts Dermatology	Clinical trials: CT = 100% Follow-up: 3–5 years	Literature search for full publications	Positive 68% (15/22) Negative 0% (0/2) Neutral 17% (1/6)	Unclear definition of positive or negative results. Only abstract available
Eloubeidi et al. 2001[60]	Submitted meeting abstracts Gastrointestinal endoscopy	Mixed: RCT = 9% Follow-up: 4 years	Literature search for full publications	Significant 37% (36/98) Non-significant: 22% (77/353)	Positive: statistically significant $p < 0.05$ Multivariate adjusted OR: 0.97 (0.58–1.60); HR = 1.92 (1.28–2.87). Also with data on presentation acceptance
Evers 2000[61]	Presented meeting abstracts Reproductive	RCTs = 100% Follow-up: 4–8 years	Literature search for full publications	Significant 59% (41/69) Non-significant 46% (38/82)	Significant: $p < 0.05$
Glick et al. 2006[62]	Presented meeting abstracts Organ transplantation	Mixed: Ob = 81% Other = 13% RCT = 6% Follow-up 4–5 years	Literature search for full publications	Significant 52% (208/397) Not specified 59% (304/516) Non-significant 41% (95/234)	Significant: $p < 0.05$ Statistical significance was excluded from multivariate analysis due to >5% of data missing
Ha et al. 2008[63]	Presented meeting abstracts Radiology	Mixed Follow-up: >4 years	Literature search for full publications	Positive 29% (288/982) Negative 11% (13/115)	Positive outcomes: beneficial or statistically significant results Negative: non-positive results

131

Study	Cohort types Speciality	Study design Follow-up	Verification of publication status and results of unpublished studies	Publication rate	Definition of study results and other notes
Halpern et al. 2001[64]	Presented meeting abstracts Anaesthesiology	Mixed Follow-up: > 5 years	Literature search for full publications	Positive 35% (29/83) Not positive 19% (9/47)	Positive results: significant results. Lack of details. Unpublished data reported in Scherer et al.[55]
Harris et al. 2006[66]	Presented meeting abstracts Orthopaedics	Mixed: Ob=72% Other=26% RCT=2% Follow-up: 5 years	Literature search and contacting authors for full publications	Positive 34% (45/132) Negative 50% (5/10) Neutral 21% (12/58) Significant (p<0.05) 50% (12/24) Non-significant 28% (50/176)	Positive results: beneficial regardless of p values Negative results: against the intervention Neutral results: no opinion After adjusting for study setting (clinical or laboratory), neither statistical significance (p=0.2) nor direction of outcome (p=0.3) were significantly associated with publication
Harris et al. 2007[65]	Presented meeting abstracts Orthopaedics	Mixed: Ob=82% Other=12% RCT=6% Follow-up: 5 years	Literature search and contacting authors for full publications	Positive 61% (123/203) Negative 53% (18/34) Neutral 43% (35/81) Significant (p<0.05) 68% (69/101) Non-significant 49% (107/217)	Positive results: beneficial regardless of p values Negative results: against the intervention Neutral results: no opinion
Hashkes and Uziel 2003[67]	Presented meeting abstracts Paediatric rheumatology	Mixed: Basic=5% Ob=69% Analytic/CT=26% Follow-up: 3 years	Literature search and contacting authors for full publications	Positive 48% (54/112) Negative 14% (2/14) Neutral 27% (36/131)	Definition of positive, negative or neutral results not provided
Kiroff 2001[68]	Presented meeting abstracts Surgery	Mixed: RCT=4% CT=31% (subgroup – clinical trials) Follow-up: 3–5 years	Contacted authors for full publications	*All studies:* Positive 71% (98/139) Negative/inclusive 48% (76/159) *Clinical trials:* Positive 92% (11/12) Negative/inclusive 50% (4/8)	Significance or importance of the results based on information from authors, but lack of details A single investigator assessed meeting abstracts

Study	Cohort types / Speciality	Study design / Follow-up	Verification of publication status and results of unpublished studies	Publication rate	Definition of study results and other notes
Klassen et al. 2002[69]	Presented meeting abstracts / Paediatric	RCT = 100% / Follow-up: 5–8 years	Literature search for full publications	Favouring treatment 69% (162/235) / Non-favourable 50% (93/187)	Favouring treatment: overall conclusions favoured the intervention / With data on time to publication, and abstract bias
Krzyzanowska et al. 2003[70]	Presented meeting abstracts / Oncology	Large clinical trials (n>200): CT = 100% / Follow-up: <5 years	Literature search and contacting authors for full publications	Significant 81% (181/223) / Non-significant 68% (195/287) / Positive 81% (148/183) / Negative 70% (229/327)	Significant results: $p \leq 0.05$ / Positive results: $p \leq 0.05$ in favour of the experimental treatment / With data on time to publication
Landry 1996[51]	Presented meeting abstracts / Burn research	Mixed: CT = 27%. / Follow-up: 4 years	Literature search for full publications	Positive 41% (24/58) / Non-positive 18% (20/110)	Positive: $p<0.05$ or stated to be positive / Data not clearly presented. Publication rate from CMRD by Scherer et al.
Loep and Kleijnen 1999[52]	Abstracts initially published in a journal / Unclear	Clinical trials / Follow-up: >1 year	Literature search and contacting authors for full publications	Positive 81% (72/89) / Negative 81% (34/42)	Data from the 2000 *HTA* report on publication bias, based on unpublished manuscript
Peng et al. 2006[71]	Presented meeting abstracts / Otolaryngology: head and neck surgery	Mixed / Follow-up: >5 years	Literature search for full publications	Positive 61% (189/337) / Negative 50% (13/26)	Unclear definition of positive or negative results
Petticrew et al. 1999[53]	Presented meeting abstracts / Social medicine	Mixed: CT = 5% / Follow-up: 2 years	Literature search and contacting authors for full publications	Positive 50% (22/36) / Uncertain 56% (19/34) / Negative 57% (4/7)	Classification of results based on subjective assessment of the study results and the authors' conclusions
Sanossian et al. 2006[72]	Presented meeting abstracts / Stroke	Mixed: CT = 2% / Follow-up: 5 years	Literature search and contacting authors for full publications	Positive 62% (136/220) / Non-positive 62% (83/133)	Positive: beneficial or supported hypothesis or objective, and either $p<0.05$ or no statistical test reported / Adjusted publication rate: 64% for positive and 59% for non-positive results. Clinical trials 100% published
Scherer et al. 1994[54]	Presented meeting abstracts / Ophthalmology	Clinical trials: CT = 100% / Follow-up: 3 years	Literature search and contacting authors for full publications	Significant 72% (33/46) / Non-significant 59% (28/47)	Statistically significant $p<0.05$

Study	Cohort types Speciality	Study design Follow-up	Verification of publication status and results of unpublished studies	Publication rate	Definition of study results and other notes
Smith *et al.* 2007[73]	Presented meeting abstracts Urology	Mixed clinical research Follow-up: > 2 years	Literature search for full publications	Significant 47% (521/1120) Non-significant 43% (86/202)	Positive results were those showing statistically significant results ($p < 0.05$) regardless of the direction
Timmer *et al.* 2002[74]	A random sample of abstracts submitted to a conference Gastroenterology	Controlled clinical trials (39%), other clinical research (40%) and basic studies (21%) Follow-up: 3–6 years	Literature search for full publications	*All abstracts* Significant ($p < 0.05$): 50% (177/354) Non-significant 47% (69/147) Equivocal 43% (144/335) *Controlled clinical trials* Significant ($p < 0.05$): 60% (84/140) Non-significant 48% (47/99) Equivocal 45% (39/87)	Significant results: $p < 0.05$. Equivocal results: no statements concerning the statistical significance of the main or the majority of otucomes
Vecchi *et al.* 2006[75]	Presented meeting abstracts Drug addiction	Clinical trials: CT = 100% Follow-up: > 5 years	Literature search for full publications	Positive 75% (120/161) Negative/null 47% (24/51) No statistical results 61% (198/325) No results: 39% (17/44)	Positive results: statistically significant results ($p < 0.05$) in favour of experimental arm Negative or null: significant results ($p < 0.05$) in the control arm or not significant ($p \geq 0.05$)
Zamakhshary *et al.* 2006[76]	Presented meeting abstracts Paediatric	Mixed: Basic = 25% Ob = 36% RCT = 1% Follow-up: < 2 years	Literature search for full publications	Significant ($p < 0.05$) 70% (105/151) Non-significant 41% (13/32)	Significant results: $p < 0.05$
Zaretsky and Imrie 2002[77]	Presented meeting abstracts Haemotology	Phase III trials (*n* = 57) Follow-up: 7 years	Literature search for full publications	The rates of publication of positive and negative results not significantly different ($p = 0.53$)	Only a short abstract available

Appendix 8

Main characteristics of manuscript cohort studies of publication bias: manuscripts submitted to journals

Study	Methods	Main findings
Olson *et al.* 2002[81]	Cohort study of 745 manuscripts of controlled trials submitted to *JAMA* from 02/1996 to 08/1999 *Outcome classification:* 1. Statistically significant (*p* < 0.05) for the primary outcome 2. Statistically non-significant 3. Unclear	*Proportion of studies with different results:* 51.4% (*n* = 383) with significant results 45.7% (*n* = 341) with non-significant results 2.8% (*n* = 21) with unclear results *Acceptance rate:* 20.4% (78/383) for significant results 15.0% (51/341) for non-significant results 19.0% (4/21) for unclear results. Logistic regression analysis: significant vs non-significant results OR = 1.30 (95% CI: 0.87 to 1.96)
Lee *et al.* 2006[78]	Cohort study of 1107 manuscripts of original research (including qualitative research, excluding single case reports) submitted to *BMJ, Lancet* and *Annals of Internal Medicine* during 01–03/2003 and during 11/2003–02/2004 *Outcome classification:* 1. Statistically significant (*p* < 0.05) for the primary outcome 2. Non-significant	*Proportion of different statistical results:* 86.8% (*n* = 718) with significant results 13.2% (*n* = 109) with non-significant results *Acceptance rate:* 4.9% (35/718) for significant results 6.4% (7/109) for non-significant results Multivariate analysis: OR = 0.83 (95% CI: 0.34 to 1.96)
Lynch *et al.* 2007[79]	Cohort study of 209 manuscripts of original research on hip or knee arthroplasty submitted to the *Journal of Bone and Joint Surgery (American Volume)* between 01/2004 and 06/2005 *Outcome classification:* 1. Positive or favourable or significant 2. Negative or non-supportive or no difference 3. Not analysable – unclear	*Proportion of studies with different results:* 70.8% (*n* = 148) with positive results 23.4% (*n* = 49) with negative results 5.7% (*n* = 12) with unclear results *Acceptance rate:* 30.4% (45/148) for positive results 36.7% (18/49) for negative results 8.3% (1/12) for unclear results Difference in publication rate between positive and negative outcomes was not statistically significant (*p* = 0.41)
Okike *et al.* 2008[80]	Cohort study of 855 manuscripts as scientific articles submitted to the *Journal of Bone and Joint Surgery (American Version)* between 01/2004 and 06/2005 *Outcome classification:* 1. Positive – favoured experimental item 2. Negative – favoured existing standard of care over the experimental item 3. Neutral – no difference 4. Not applicable	*Proportion of studies with different results:* 72.5% (*n* = 620) with positive results 12.3% (*n* = 105) with negative results 15.2% (*n* = 130) with neutral results *Acceptance rate:* 21.3% (132/620) for positive results 21.0% (22/105) for negative findings 24.6% (32/130) for neutral results Multivariate analysis: positive vs nonpositive OR = 0.92 (95% CI: 0.62 to 1.35)

Appendix 9

Outcome reporting bias –
characteristics of included studies

Study	Cohort or sample, methods	Findings	Notes
Bekkering et al. 2008[93]	Based on two systematic reviews of 767 observational studies (with 3284 results) of the association between diet and prostate or bladder cancer. The paper examined the proportion of results that had detail sufficient for use in meta-analyses estimating dose–response associations	61% of results were usable in dose–response meta-analyses. The most important reason for results not being usable was the absence of sufficient information on exposure levels in the different groups. Results that showed evidence of an association were more likely to be usable than results that found no such evidence	
Chan et al. 2004 (CMAJ)[7]	RCT protocols approved by the Canadian Institute of Health Research from 1990 to 1998 and subsequent journal publications. Reported and unreported outcomes were recorded from protocols and journal articles. If a published article provided insufficient data for meta-analysis, the outcome was defined as being incompletely reported (i.e. partial + qualitative + unreported). Used odds ratios to measure association between the completeness of outcome reporting and statistical significance Of 105 RCTs approved for funding, 48 published; with a total of 1402 outcomes measured: 1233 efficacy outcomes in 48 trials and 169 harm outcomes in 26 trials	The median number of participants per trial was 299. A median of 31% of efficacy outcomes and 59% of harm outcomes were incompletely reported. Statistically significant outcomes had higher odds than non-significant outcomes of being fully reported: OR=2.7 (95% CI: 1.5 to 5.0) for efficacy outcomes and OR=7.7 (95% CI: 0.5 to 111) for harm outcomes Primary outcomes differed between protocols and publications for 40% of the trials	Of the 48 published RCTs, only 30 for efficacy and 4 for harm outcomes were included for analyses. Only 22 trialists provided information about the statistical significance of unreported outcomes. Trials were also excluded if they had fully reported outcomes, or had only incompletely reported outcomes, or all outcomes being significant or non-significant. It is unclear whether the exclusion of trials might systematically affect the estimated associations between outcome reporting and study results
Chan et al. 2004 (JAMA)[6]	RCT protocols approved by the REC in Denmark in 1994–5. Reported and unreported outcomes were recorded from protocols, journal articles and a survey of trialists. If a published article provided insufficient data for meta-analysis, the outcome was defined as being incompletely reported. Odds ratios were used to measure association between the completeness of outcome reporting and statistical significance Identified 102 RCT protocols, 122 corresponding journal articles and 3736 outcomes	Overall 50% of efficacy and 65% of harm outcomes per trial were incompletely reported. Statistically significant outcomes had a higher odds of being fully reported compared with non-significant outcomes for both efficacy (OR=2.4; 95% CI: 1.4 to 4.0) and harm (OR=4.7; 95% CI: 1.8 to 12.0) data. In comparing published articles with protocols, 62% of trials had at least one primary outcome that was changed, introduced, or omitted. 86% of survey responders (42/49) denied the existence of unreported outcomes despite clear evidence to the contrary Exploratory meta-regression analysis found no significant association between the extent of bias and the source of funding, sample size, or number of study centres	Journal articles published mainly in 1998–9, before the 2001 CONSORT statement. Survey response rate was low, which may lead to an underestimation of bias. 49/99 trials measuring efficacy and 54/72 trials measuring harm outcomes were excluded from the analysis of reporting bias due to entire rows or columns being empty in the 2×2 table. 22% of efficacy and 35% of harm outcomes were ineligible for analysis due to unknown statistical significance
Chan et al. 2005[5]	RCTs indexed in PubMed appeared in Dec 2000. Trialists were surveyed to obtain information on unreported outcomes	Median percentage of incompletely reported outcomes per trial: 42% for efficacy outcomes and 50% for harm outcomes. 33% (169/505) of trials had at least one unreported efficacy outcome and 28% (85/308) of trials had unreported harms data. Pooled ORs for outcome reporting bias: 2.0 (95% CI: 1.6 to 2.7) for efficacy outcomes, and 1.9 (95% CI: 1.1 to 3.5) for harm outcomes	Low response rate may lead to underestimation of bias. Trial protocols were not reviewed so that fewer unreported outcomes were identified compared with previous two studies by the same investigators

Study	Cohort or sample, methods	Findings	Notes
	519 trials identified with 10,557 outcomes. Survey responders (69%) provided information on unreported outcomes (but unreliable)	Reasons given by authors for not reporting efficacy and harm outcomes, respectively: space constraints (47%, 25%); not clinically important (37%, 75%); not statistically significant (24%, 50%); not yet submitted (22%, 6%); not yet analysed (17%, 6%)	(44% of trials found to have one or more unreported outcomes in this study, compared with 76% and 98% of trials in the earlier studies based on trial protocols)
Furukawa et al. 2007[94]	A random sample of 156 Cochrane systematic reviews with 10 or more RCTs. Examined percentage of identified RCTs that contributed in meta-analysis; and the association between the percentage of RCTs contributing and pooled estimates (odds ratio and standardised mean difference)	A median of 46% (IQR 20–75%) of identified RCTs in each meta-analysis contributed to the pooled estimates. Regression analysis of percentage contributing RCTs and effect size: $\beta = -0.16$ (95% CI: -0.29 to -0.01) for OR and $\beta = -0.18$ (95% CI: -0.35 to -0.01) for SMD. When outcomes favoured the control, regression coefficient was $\beta = 0.16$ (95% CI: -0.16 to 0.45) for OR and $\beta = -0.08$ (95% CI: -0.37 to 0.22) for SMD	Concluded that the under-reporting appears to be biased. Discussed possible explanations for incomplete inclusion of outcomes. Did not assess potential confounders or explanatory variables
Ghersi et al. 2006[91] (abstract)	A comparison of 103 published RCTs and their protocols considered by Central Sydney Area Health service ERC from Jan 1992 to Dec 1996	17% of primary outcomes in the protocol were not reported as primary outcomes in the publication. 15% of primary outcomes in the publication not declared as primary outcomes in the protocol. Trials where all of the comparisons were statistically significant were more likely to report fully all of their comparisons ($p = 0.06$)	Only abstract available. Not clear about the association between statistical significant and the selective reporting of primary outcomes. Authors contacted for full publication
Hahn et al. 2002[36]	A pilot study comparing local REC approved protocols and results presented in subsequent publications. 41 (73%) responses received from trialists of 56 LREC approved projects (5 years ago). 18 published (but publications available only for 15. For the remaining three, two researchers did not agree to provide their article nor any references, and the copy of one presented as a conference poster was no longer available	Of the 15 published studies, six stated which outcome variables were of primary interest and four showed consistencies in the report. Eight mentioned an analysis plan. However, seven of these eight studies did not follow their prescribed analysis plan: the analysis of outcome variables or associations between certain variables was found to be missing from the report	A pilot study: before this study, 'the original protocol for a study and its subsequent report have never been compared in order to summarize the consistency between them using a structured framework for an assessment of within-study selection'
Kavvoura et al. 2007[98]	389 abstracts and 50 randomly selected full papers of epidemiological studies. Examined percentage of reporting statistically significant and non-significant results in the abstract and the association between the RRs and the type of contrast used	In the abstracts: 88% reported ≥ 1 significant RR and only 43% reported ≥ 1 non-significant RR. Full text of the 50 articles: a median of 9 (IQR 5–16) significant and 6 (IQR 3–16) non-significant RRs were presented. Paradoxically, the smallest presented RRs were based on the contrasts of extreme quintiles	The preponderance of significant findings was less prominent in the full texts of the articles. Selection of results by multiple analyses methods to measure the same outcome. The use of extreme contrasts may indicate zero or very small effect size
Kyzas et al. 2005[97]	Meta-analysis of studies on the association between the tumour suppressor protein (TP53) and mortality outcome of patients with head and neck squamous cell cancer. Study categories: (1) published and indexed with 'mortality' or 'survival'; (2) published but without 'mortality' or 'survival' index in MEDLINE and EMBASE; (3) retrieved: data retrieved from authors for those that suggested mortality collected but reported no usable data	n = number of trials (number of patients). Published/indexed: $n = 18$ (1364); RR = 1.27 (95% CI: 1.06 to 1.53). Published/not indexed: $n = 13$ (1028); RR = 1.13 (95% CI: 0.81 to 1.59). Retrieved: $n = 11$ (996); RR = 0.97 (95% CI: 0.72 to 1.29). All studies: $n = 42$ (3388); RR = 1.16 (95% CI: 0.99 to 1.35). The association was stronger by using the definitions preferred by each publication (RR = 1.38, 1.13–1.67) than when definitions were standardised (RR = 1.27, 1.06 to 1.53)	The definitions of outcomes were also selected to exaggerate the association. Discussed implications and approaches to dealing with the selective reporting bias

Study	Cohort or sample, methods	Findings	Notes
McCormack *et al.* 2004[92]	A meta-analysis based on aggregate published data was compared with an updated IPD meta-analysis of trials of hernia surgery	*Hernia recurrence outcomes:* numbers of contributing RCTs were similar and the results had no significant difference *Persisting pain outcome:* many more RCTs were included in IPD update (e.g. 20 vs 3), and results were qualitatively divergent	This case study indicates that some outcomes (e.g. persisting pain, an outcome rarely reported) may be more vulnerable than others (e.g. hernia recurrence) to the selective reporting. IPD meta-analysis may be an approach to dealing with selective reporting
Melander *et al.* 2003[43]	42 placebo controlled trials of 5 SSRIs submitted to the Swedish drug regulatory authority for marketing approval. Published versions were identified by searching literature databases and inquire to the sponsoring companies	All 21 trials with significant results but only 17 of the 21 trials with non-significant results were published. Many publications ignored the results of ITT analyses and reported the more favourable per protocol analyses only	Different from other studies: only one outcome (response rate) was considered. The selection based on different analysis methods (ITT vs per protocol). Also provided data on duplicate bias
Scharf and Colevas 2006[96]	A comparison of adverse events (AEs) reported in 22 published articles and corresponding protocols and data from Clinical Data Update System (CDUS). CDUS monitored phase II trials that were active between 03/1998 and 10/2003	27% of high-grade AEs in articles could not be matched to agent-attributable AEs in the CDUS. 28% of CDUS high-grade AEs could not be matched to AEs in the corresponding article. In 14 of 22 articles, the number of high-grade AEs in CDUS differed from the number in the articles by ≥20%. 58% of low-grade AEs in CDUS were reported in articles	Mismatch in AE reporting but not clear whether it was biased in terms of results. Anderson (2006)[523] pointed out some caveats about this study. Low-grade and recurrent AEs were under-reported
Williamson and Gamble 2005[95]	A comparison of results of Cochrane systematic reviews (CSRs) and the results of sensitivity analyses (by imputation) when within-study outcome selection bias was suspected. Included one motivating example and four CSRs from Sutton *et al.* (2000)[524]	The motivating example: five of the nine eligible trials included for mortality outcome. Imputed pooled estimate decreased treatment effect considerably For the four selected CSRs, within-study selection was evident or suspected in several trials, but the impact on the conclusions of the meta-analyses was minimal	A good discussion of relevant issues and the imputation method Is funnel plot appropriate to detect reporting bias? 'Assumed that the results were not selectively reported as continuous rather than binary results' – disagree or lack of evidence?
Williamson *et al.* 2006[99]	To estimate the percentage and impact of within-study selective reporting in an unselected cohort of 300 Cochrane systematic reviews		MRC-funded ongoing project

ERC, Ethics Review Committee; LREC, Local Research Ethics Committee; SMD, standardised mean deviation.

Appendix 10

Time lag bias –
included empirical studies

Cohorts of studies collated from various sources

Study	Methods	Results	Notes
Simes 1987[101]	38 trials located by a search of literature and trial registry, on advanced ovarian cancer and multiple myeloma Positive studies – those that showed significant survival difference	**Time to publication from study closure** |	Positive	Non-positive | |---|---|---| | 1–2 | 2 | 11 | | 3–5 | 4 | 9 | | 6–10 | 0 | 5 | Not yet published: 0 7	Three trials (unpublished) did not have survival information, and assumed to be not statistically significant
Stern and Simes 1997[24]	Retrospective cohort of 748 studies submitted to the Royal Prince Alfred Hospital REC from 1979 to 1988 Results of quantitative studies (n=218) were classified as: significant (p<0.05), non-significant trend (p=0.05–0.10) and non-significant or null (p≥0.10) Results of qualitative studies (n=103) classified as: striking, important/definite, or unimportant/negative	*Quantitative studies (n=218)* Median time from approval by REC to publication in journals 4.82 (3.87–5.72) years for studies with significant results (p<0.05), 7.99 (6.91 to ∝) for studies with null results (p>0.10) Adjusted hazard ratio (vs null results p>0.10):2.34 (95% CI:1.47 to 3.73) for significant results (p<0.05), and 0.43 (95% CI:0.15 to 1.24) for intermediate results (p=0.05–0.10) *Qualitative studies (n=103)* No clear evidence on publication bias	Analyses were based on 321 completed studies from 520 studies for which questionnaires were completed Results for clinical trials were similar to those for all quantitative studies
Ioannidis 1998[23]	Retrospective cohort of 66 completed phase 2/3 trials conducted between 1986 and 1996 by the AIDS clinical trials group and by Terry Beirn Community Programs for Clinical Research on AIDS Positive results defined as those with p<0.05 favouring the experimental therapy; and negative results defined as those with p>0.05 or favouring the control	Median time from start of enrolment to publication was 4.3 years for positive trials and 6.4 years for negative trials (p<0.001; HR=3.7 (95% CI: 1.8 to 7.7) Median time from completion of follow-up to publication was 1.7 years for positive trials and 3.0 for negative trials (p<0.001) Positive trials were submitted for publication more rapidly after completion than were negative trials (median 1.0 vs 1.6 years; p=0.001) and were published more rapidly after submission (median 0.8 vs 1.1 years; p=0.04)	The total number of eligible trials was 109
Misakian and Bero 1998[30]	61 completed studies identified by a survey of 89 organisations funding for research on passive smoking and investigators Results classified as significant (p<0.01), mixed (multiple primary outcomes with at least one significant outcome), or non-significant (p>0.05)	Median time from funding start date to publication was 3 years for significant studies, 6 years for mixed results, and 5 years for non-significant results Hazard ratio: 1.0 for non-significant results; 0.12 (95% CI: 0.02 to 0.97) for mixed; and 1.19 (95% CI: 0.95 to 3.84) for significant studies	Unclear whether the method for results classification was prespecified or not. If significant and mixed results were combined, the publication rate was lower than that for non-significant studies (73% vs 86%)

Study	Methods	Results	Notes
Min and Dickersin 2005[104]	242 observational studies that completed enrolment, initiated at Johns Hopkins University. Study results and full publication were verified by interview with the investigators Results were classified as statistically significant or not	Cox regression analysis found that time to full publication was associated with statistically significant results for the primary outcome specified (HR = 1.75, 95% CI: 1.14 to 2.93)	Only available as an abstract. Likely to be a subsample from the previous study by Min and Dickersin (1992)[20]
Dickersin 2002[82]	A cohort of 133 comparative studies submitted to and accepted for publication in *JAMA* Results classified as positive when $p < 0.05$, or negative when $p > 0.05$, for the main outcome	Median time between submission and publication was 7.8 months for studies with positive results vs 7.6 months for reports with negative results ($p = 0.44$)	Only accepted studies were included
Cronin and Sheldon 2004[27]	A cohort of 70 studies sponsored by NHS R&D programme from 1995 to 1998; with a survey of project leaders Results were classified as showed effect ($p < 0.05$ or important) or not	Time between completion of study and publication in a journal, univariate survival analysis hazard ratio 0.53 (95% CI: 0.25 to 1.1), $p = 0.10$	Contacted the author for details
Soares *et al.* 2005[103]	56 published phase III trials conducted by Radiation Therapy Oncology Group (RTOG) Results were classified according to the original trialists' preferences between experimental and standard treatment	Trial results were not associated with the average time to publication (from the date of trial initiation, or closure to accrual)	Published phase III trials only. An abstract without full publication
Hall *et al.* 2007[37]	53 published studies with sufficient data on time to publication, from 84 published studies from 190 research protocols submitted to the Capital District Health Authority REB in Halifax, Canada, for the period 1995–6 Results classified as statistically significant ($p < 0.05$) or not	Median time from the end of recruitment or follow-up as described in the methods section until publication was 2.71 (range 0.38 to 8.42) years. No difference in the time to publication for trials reporting significant results vs non-significant results (2.67 vs 3.00 years, $p = 0.87$)	Of the 84 published studies, 71 reported statistically significant results. The early phase studies were less likely to be published than late stage studies
Liebeskind *et al.* 2006[102]	Cases with available data from 159 acute stroke clinical trials fully published from 1955 to 1999 Results were classified as beneficial, or non-beneficial (harmful or neutral), according to authors' final judgement	*Mean time (beneficial vs no beneficial studies)* From enrolment initiation to publication ($n = 65$): 4.1 vs 4.2 years, $p = 0.70$ From enrolment completion to publication ($n = 55$): 2.0 vs 2.3 years, $p = 0.21$ From enrolment completion to submission ($n = 27$): 1.8 vs 2.2, $p = 0.19$ Between submission and acceptance ($n = 51$): 0.37 vs 0.37 years, $p = 0.92$ From acceptance to publication ($n = 61$): 0.38 vs 0.40 years, $p = 0.93$	Studies identified by searching electronic databases and Cochrane library. Subgroup analysis available for pharmaceutical sponsored trials. Also identified 19 trials in abstract form only and four unpublished trials

144

Cohorts of abstracts presented at meetings

Study	Methods	Results	Notes
Callaham et al. 1998[105]	493 research abstracts submitted for presentation at the 1991 meeting of the Society of Academic Emergency Medicine. Subsequent full publication verified by searching MEDLINE or contacting authors Results were classified as positive (beneficial or statistically significant) or negative	The mean time from the meeting to publication was no different between studies with positive results and those with negative results (1.6 vs 1.3 years, $p = 0.20$)	Callaham et al. compared their result with that of Stern and Simes (1997)[24]
Cheng et al. 1998[49]	178 abstracts of RCTs in cystic fibrosis (CF), presented at three CF conferences over a 30-year period. Searched Cochrane trial registers for full publications Results were classified as positive or negative, according to authors' conclusion about whether the test treatment was superior to the control	Log-rank tests did not show a significant difference in time to publication between results classified as positive or negative ($p = 0.54$). Cox regression model also did not demonstrate any evidence of an association between time to publication and results and sample size	No consistent factors were found
Eloubeidi et al. 2001[60]	461 research abstracts submitted to a GI endoscopic research meeting in 1994. Full publication was tracked by a literature search Results were classified as statistically significant ($p < 0.05$) or not	Multivariate Cox proportional hazards analysis: time to full publication for significant vs non-significant results HR = 1.92 (95% CI: 1.28 to 2.87) ($p = 0.0015$)	
Evers 2000[61]	151 abstracts of RCTs presented at the European Society of Human Reproduction and Embryology. Electronic databases and major journals were searched for subsequent full publication Results were classified as significant ($p < 0.05$) or not	Log-rank testing showed no significant differences in publication rate between studies with significant results and those with non-significant results ($\chi^2 = 3.06$, $p = 0.08$)	
Klassen et al. 2002[69]	447 abstracts of phase 3 RCTs presented at the Society for Paediatric Research Meeting (1992–5). Subsequent publication was ascertained by a literature search Results were classified according to authors' conclusion whether the result was in favour of the test treatment	After 5 years, the rate of publication was 66% for positive results and 44% for non-positive results	Time to publication from presentation was shown separately for positive and negative studies (but without statistical testing details)
Krzyzanowska et al. 2003[70]	510 abstracts from large phase 3 RCTs presented at an oncology conference. Subsequent full publication was tracked by a literature search and survey of original investigators Results were classified as significant ($p < 0.05$) or not for the primary end point	Median time from abstract presentation to full publication 2.2 years for significant results and 3.0 years for non-significant results Hazard ratio for significant vs non-significant results 1.4 (95% CI: 1.1 to 1.7)	From abstract presentation to publication

Temporal trends of reported effect size

Study	Methods	Results	Notes
Rothwell and Robertson 1997[106]	26 meta-analyses (with 241 trials in total) from a literature search were used to test the hypothesis that treatment effect might be related to year of publication	Early trials (year 1 and 2) overestimated the treatment effect compared with a meta-analysis of the subsequent trials in 20 of the 26 meta-analyses. The average difference in relative odds was 35% (95% CI: 15% to 55%)	The temporal trends were independent of trial size
Song and Gilbody 1998[107]	38 meta-analyses published in *BMJ* or *JAMA* during 1992–6. Rank correlation analysis of year of publication and the treatment effect	Four of the 38 meta-analyses showed a significant correlation ($p < 0.10$) between the year of publication and the treatment effect	More cases ($n = 10$) showed a significant correlation between sample size and treatment effect
Gehr et al. 2006[108]	Meta-analyses of lipid-lowering drugs pravastatin (RCTs = 64) and atorvastatin (RCTs = 35), and anti-glaucoma drugs timolol (RCTs = 75) and latanoprost (RCTs = 32) Regression analysis of reported effect size against year of publication	Pravastatin on LDL-C: effect size significantly reduced (-3.22% in every 5 years, $p < 0.0001$) Atorvastatin on LDL-C: no significant change in effect size ($+0.31\%$, $p = 0.86$) Timolol ($-0.56\,\text{mmHg}$, $p < 0.0001$) and latanoprost ($-1.78\,\text{mmHg}$, $p = 0.007$) decreased over time	Effect sizes in RCTs decreases over time in three of the four cases, caused mainly by baseline differences. The phenomenon was termed 'fading of reported effectiveness' by authors
Vaitkus and Brar 2007[109]	A case study: meta-analysis of *N*-acetylcysteine in the prevention of contrast-induced nephropathy, included 27 RCTs published only as manuscripts ($n = 12$) or as abstracts ($n = 2$) or both ($n = 13$)	Trials published earlier reported more favourable results than the trials published later	The study also reported bias related to journal impact factor and a comparison of full publication and abstracts
Jennions and Moller 2002[525]	44 published meta-analyses in ecology and evolution that provided sufficient data on primary studies included. Examined the relationship between estimated effect sizes and years of publication	At the original meta-analyses level, there was a significant negative relationship between year of publication and effect size ($r = -0.133$, $p < 0.01$; $n = 44$), the association remained after controlling for sample size ($r = -0.105$, $p < 0.01$; $n = 39$)	There was also a negative relationship between the sample size and effect size. Several possible explanations were mentioned
Leimu and Koricheva 2004[111]	Two meta-analyses of studies testing two plant defence theories in ecology Correlation analysis and cumulative meta-analysis	Correlation analyses revealed no significant association between the magnitude of effect size and publication year in either of the two meta-analyses	Studies of ecology and recommended the use of cumulative meta-analysis
Mieog and Ghersi 2005[526]	To explore the impact over time of the inclusion of data from abstracts on a Cochrane review of preoperative chemotherapy for early breast cancer	Inclusion of abstracts does not alter conclusions over time for overall survival and rate of mastectomy. For loco-regional treatment, inclusion of a recent abstract substantially amplifies the beneficial effect of preoperative chemotherapy	Focused on the use of abstracts in meta-analysis. Results obtained from a poster on internet. Note: different conclusion from Marinovich et al. (2005)[529] 'data maturity and systematic reviews of new health technologies'

LDL-C, low-density lipoprotein C; REB, Research Ethics Board.

Appendix 11

Grey literature bias – included empirical studies

Study	Methods	Main findings	Notes
Studies of multiple (10 or more) meta-analyses			
Fergusson et al. 2000[121]	10 meta-analyses of perioperative transfusion, identified from a project by the International Study of Perio-operative Transfusion Grey literature included 15 abstracts, one letter, three conference proceedings and one unpublished report	Reported treatment effect was (on average) statistically non-significantly greater in published trials compared with grey literature trials	Only five of the identified grey literature items were still not fully published at the time of analysis
McAuley et al. 2000[119]	41 meta-analyses that used binary outcomes, included 365 published and 102 trials available only as grey literature Grey literature included: abstracts (61%), unpublished trials (17%) or in press (3%), book chapters (6%), theses (2%), company reports (3%)	In 14 of the 41 analyses removal of the grey literature changed the estimate of treatment effect by ≥10% (treatment effect was greater in nine cases but smaller in five cases) On average published trials vs grey literature yielded larger treatment effect (ROR 1.15; 95% CI: 1.04 to 1.28)	A few places were unclear about the removal of abstracts from the meta-analyses
Egger et al. 2003[3]	60 meta-analyses of health-care interventions that included 630 published and 153 unpublished trials 153 unpublished studies including abstracts (45%), book sections (14%), theses (3%) and other forms (37%)	Reported treatment effect based on grey literature ranged from 97% more to 209% less beneficial than those from the corresponding published trials. Pooled effect estimates from the grey literature were on average greater than those from published trials (ROR: 1.07; 95% CI: 0.98 to 1.15)	Included only meta-analyses that conducted comprehensive literature search, used binary outcome and had five or more trials. Grey literature bias may be greater in meta-analyses that included fewer trials or without a comprehensive literature search
Burdett et al. 2003[120]	11 of the 13 individual patient meta-analyses of cancer, conducted at MRC Clinical Trial Unit (London), that included both published trials (n=75) and grey literature trials (n=45) Grey literature trials included unpublished data (53%), abstracts (38%), book chapters (7%) and non-English-language publications (2%)	Estimated treatment effect using data from fully published trials tended to be greater than that with grey literature trials, although the direction of bias was not always predictable Hazard ratio of hazard ratios (HRHR) for 11 meta-analyses of fully published data was 0.93 (95% CI: 0.90 to 0.97) compared with HRHR for 11 meta-analyses including all data of 0.96 (95% CI: 0.93 to 0.99)	The authors considered that the observed bias was 'less pronounced than reported by McAuley', possibly because of clinical trials' long history in the cancer field 11 cases were IPD meta-analyses, and 53% of grey literature was unpublished data. Grey literature in this study also included trials published in non-English-language journals (2% only)
Hopewell 2004[122]	17 meta-analyses in cancer; included 264 RCTs, identified from CDSR Grey literature included conference abstracts, letters, books, government and pharmaceutical reports, theses, file drawer data and personal correspondence	Reported treatment effect was (on average) statistically non-significantly smaller in grey literature trials than that in published trials. ROR = 1.05 (95% CI: 0.83 to 1.33)	Data obtained from Hopewell (2007)[113] (Cochrane Methodology Review)
Turner et al. 2008[45]	74 clinical trials of 12 antidepressants submitted and approved by FDA between 1984 and 2004. Of the 74 trials from FDA reviews, 23 were unpublished Unpublished trials from FDA databases	The effect size based on the journal articles (0.41; 95% CI: 0.36 to 0.45) was on average 32% (ranged from 11% to 69%) greater than the effect size based on the FDA reviews (0.31; 95% CI: 0.27 to 0.35) (sign test p<0.001)	Also included in the review of cohort studies. Negative findings (according to FDA's decision) were less likely to be published, and if published, the negative result was often conveyed as a positive outcome

Study	Methods	Main findings	Notes
Case studies			
Whittington *et al.* 2004[136]	A review of five SSRIs (fluoxetine, paroxetine, sertraline, citalopram, venlafaxine) in childhood depression, included five published and six unpublished trials (plus some unpublished data for published studies) Unpublished studies obtained from a review by the Committee on Safety of Medicines (UK)	Two published trials and unpublished data suggest that fluoxetine has a favourable risk-benefit profile. Published data from one trial of paroxetine and two trials of sertraline suggest equivocal or weak positive risk-benefit profiles; but in both cases, addition of unpublished data indicates that risks outweigh benefits. Data from unpublished trials of citalopram and venlafaxine show unfavourable risk-benefit profiles	The review included five published trials. Combined evidence from published and unpublished data suggested that paroxetine, sertraline, venlafaxine and citalopram are not efficacious and pose a possible increased risk of suicidal ideation, serious adverse events or both
Wallace *et al.* 2006[137]	A meta-analysis of SSRIs in paediatric depression, included six published trials and one unpublished trial for efficacy meta-analysis; and seven published and four unpublished trials for safety analysis Unpublished studies obtained from the website of the Committee on Safety of Medicines (UK)	The single unpublished trial (with a negative result) did not substantially influence interpretation of the overall efficacy rates. Similarly, omission of unpublished results did not influence interpretation of safety outcomes	The review concluded that 'unpublished studies did not substantially alter the risk-to-benefit determination', which is different from Whittington *et al.*'s[136] conclusions
Devine 1999[129]	Two meta-analyses of psychoeducational care. One meta-analysis of surgical patients included 80 published and 102 unpublished studies. Another meta-analysis of cancer patients included 43 published and 35 unpublished studies Grey literature: theses or dissertations	In the two meta-analyses, published studies yielded larger average estimates of effect than dissertations	Available only in the abstract form
Jeng *et al.* 1995[131]	An IPD meta-analysis of paternal cell immunisation for recurrent miscarriage, included four published and four unpublished RCTs Unpublished trials identified from submissions to the American Society for Reproductive Immunology	Treatment effect was greater by pooling data from published trials (RR 1.29; 95% CI: 1.03 to 1.60) than that from unpublished trials (RR 1.01; 95% CI: 0.74 to 1.28)	
Detsky *et al.* 1987[128]	A meta-analysis of perioperative parenteral nutrition for reducing complications and fatalities from major surgery Grey literature: abstracts	Results of studies presented as abstracts reported greater treatment effect than those of published trials. Pooled mean difference in fatality: 0.046 (*p* = 0.21) using data from published trials and 0.079 (*p* = 0.03) from abstracts	Note the direction of bias
Man-Son-Hing *et al.* 1998[133]	A meta-analysis of quinine for nocturnal leg cramps, included four published and four unpublished RCTs Grey literature: unpublished data from FDA, German or British drug regulatory bodies, and drug companies	Pooled reduction in number of cramps per 4-week period: 8.83 (95% CI: 4.16 to 13.49) from pooling published trials and 3.60 (95% CI: 2.15 to 5.05) from all trials	

Study	Methods	Main findings	Notes
Bhandari *et al.* 2000[527]	A meta-analysis of reamed vs non-reamed intramedullary nailing of lower extremity long bone fractures, included four published and five unpublished trials	Whether a study was published or unpublished did not significantly alter the relative risk of non-union between the two interventions	
Horn and Limburg 2001[528]	A meta-analysis of calcium antagonists for ischaemic stroke, used 18 published and four unpublished trials Unpublished studies identified by contacting investigators and companies	Published trials showed no difference between the treatment and placebo (RR 1.02; 95% CI: 0.96 to 1.08), whereas unpublished trials showed a significant unfavourable effect of treatment (RR 1.14; 95% CI: 1.00 to 1.30)	
MacLean *et al.* 2003[135]	A review of 15 published trials and 11 unpublished data from FDA, that reported relative risk of dyspepsia from NSAIDs Unpublished FDA data	Pooled RR of dyspepsia: 1.07 (95% CI: 0.70 to 1.63) using FDA data and 1.21 (95% CI: 0.81 to 1.81) using published trials. Meta-regression analyses found that estimates varied significantly by NSAID dose (*p* = 0.037) but were not related to whether the study was published or not (*p* = 0.73)	An example of possible confounding factors for the association between publication status and the treatment effect This study was presented by MacLean *et al.*[130] in 1999, but with seemingly different conclusions
Macleod *et al.* 2004[132]	A meta-analysis of animal experimental studies of nicotinamide for stroke, included 11 fully published studies and three in abstract only Grey literature: abstract only	Abstracts reported a significantly lower estimate of effect size than fully published studies (*p* < 0.001)	Animal experimental research
Marinovich *et al.* 2004[529]	A case study of four RCTs (three fully published) and 46 corresponding abstracts on sirolimus-eluting stents Grey literature: abstract only	Earliest abstracts for the three RCTs were found to have presented incomplete patient numbers (3/3), mis-stated primary and/or secondary end points (2/3), mis-stated baseline data (1/3) and missed presenting follow-up results (3/3)	Available in the abstract form only
McLeod and Weisz 2004[134]	A case study of randomised trials of psychotherapy in children and adolescents included 134 published trials and 121 dissertations Grey literature: dissertation	In weighted least squares analysis the pooled effect size for published studies (mean 0.50, SD = 1.64) was twice that for dissertations (mean 0.23, SD = 1.26), and the difference was statistically significant (*p* < 0.01)	The comparison of the methodological quality composite score favoured dissertations over published studies. Dissertations included smaller sample sizes

Study	Methods	Main findings	Notes
Martin *et al.* 2005[530]	Three meta-analyses of antipsychotics for schizophrenia: olanzapine vs typical antipsychotics (one published and four unpublished trials); quetiapine vs classical antipsychotics (three published and three unpublished trials); and risperidone vs typical antipsychotics (15 published and eight unpublished trials) Type of grey literature unclear	The pooled effect size using all trials (including unpublished data) was similar to that obtained when using only published trials in all three cases	It was unclear about why the three meta-analyses were selected The authors of this article suggested that caution was required when including grey literature in meta-analysis
Maguire *et al.* 2006[531]	A review of non-randomised studies in refractory epilepsy: 42 published vs 54 abstracts for add-on topiramate; 36 published vs 44 abstracts for add-on levetiracetam; 26 published vs 21 abstracts for add-on gabapentin therapy Grey literature: abstract only	Median reported responder and seizure freedom rates were higher in topiramate studies published in full vs those from abstracts (62% vs 55% and 16% vs 8%, respectively)	Available as an abstract; considered non-randomised studies No mention of findings for levetiracetam and gabapentin
Vaitkus and Brar 2007[109]	A meta-analysis of clinical trials of N-acetylcysteine in the prevention of contrast-induced nephropathy; included 12 fully published manuscripts only, 13 as abstracts followed by full publication, and two as abstracts only Grey literature: abstract	Published manuscripts consistently showed a more beneficial treatment-effect estimate than unpublished abstracts over time. 17 published meta-analyses in this area included on average 76% of contemporaneously available manuscripts and 13% of available abstracts	Significant publication bias identified among trials studying the ability of N-acetylcysteine to prevent contrast-induced nephropathy
Moller *et al.* 2005[532]	A review of basic animal research on the relationship between asymmetry and sexual selection, included 105 published and 20 unpublished data sets Unpublished data were obtained by a survey of investigators	Unpublished and published studies did not differ significantly in effect size adjusted for sample size, using random-effects meta-analysis	Basic research

CDSR, Cochrane Database of Systematic Reviews.

Appendix 12

Language bias –
included empirical studies

Study	Methods	Main findings	Notes
Egger et al. 1997[148]	40 pairs of RCTs (in the field of internal medicine), matched for first author and time of publication, with one report published in German and the other in English	Only 35% of German-language articles, compared with 65% of English language articles, reported significant differences (p < 0.05) in the main end point (McNemar's test p = 0.002)	RCTs within the same pair were likely to consider different research questions as the matching was only by author and time of publication
Heres et al. 2004[150]	21 pairs of RCTs on neuroscience, matched for the key authors and time of publication, with one in German and the other in English	33.3% of German articles and 57.1% of English-language articles reported significant findings (Wilcoxon's test p = 0.14)	Key authors defined as the first, second or last author Similar methods and findings to Egger et al. 1997[148]

Studies of multiple meta-analyses in which results of English trials were compared with results of non-English-language trials

Study	Methods	Main findings	Notes
Gregoire et al. 1995[147]	28 meta-analyses with language restrictions, identified from eight medical journals. For seven of the meta-analyses, 11 studies published in the excluded languages were identified by repeating the original search strategies without language restrictions	The statistical significance was changed in one meta-analysis of selective gut decontamination from an OR of 0.70 (95% CI: 0.45 to 1.09) in the original analysis to an OR of 0.67 (95% CI: 0.47 to 0.95) after including studies published in German and Swiss	However, the difference between the original meta-analyses and updated meta-analyses was non-significant, even when considering the meta-analysis of selective gut decontamination
Moher et al. 2000[151]	19 meta-analyses that included 206 English and 33 non-English-language trials Non-English languages: one in Dutch, one in Danish, 12 in French, 12 in German, four in Italian, two in Spanish and one in Chinese	There was no significant difference in the estimated treatment effect for language-restricted analyses compared with language inclusive analyses (ROR = 0.96; 95% CI: 0.78 to 1.18) Language inclusive meta-analyses had narrower confidence intervals (average width = 0.79; 95% CI: 0.51 to 1.07) compared with language-restricted meta-analyses (average width = 0.92; 95% CI: 0.53 to 1.32)	The analysis is based on a small number of meta-analyses. Limitations like sampling frame, clinical topics and interventions could have affected the results. More cases were included in Moher et al. 2003[4]
Moher et al. 2003[4,152]	42 meta-analyses that included 529 English and 133 non-English-language trials Non-English languages: 57 in German, 52 in French, 15 in Italian, six in Spanish, three in Danish, two in Dutch, one each in Japanese and Portuguese	Language-restricted meta-analyses vs language inclusive meta-analyses (random-effects): ROR = 1.11 (95% CI: 0.92 to 1.34) for all cases; ROR = 1.02 (0.83 to 1.26) for 34 meta-analyses of conventional interventions; and ROR = 1.63 (1.03 to 2.60) for eight meta-analyses of complementary and alternative medicine	There were only minor differences in the quality of reports of RCTs published in English vs in other languages
Egger et al. 2003[3,153,333]	50 meta-analyses (each had five or more trials and combined binary outcomes) that included 485 English and 115 non-English-language trials Non-English languages: 36.5% in German, 25.2% in French, 10.4% in Italian, 7.0% in Japanese, 6.1% in Spanish, 5.2% in Portuguese, 7.0% in four other European languages, and 2.6% in Chinese	Treatment effect estimates were on average 16% more beneficial in non-English-language trials (ROR = 0.84; 95% CI: 0.74 to 0.97). The change in estimated treatment effect after including or excluding non-English-language trials was less than 5% in 29 (58%) meta-analyses. In the remaining 21 meta-analyses, five (10%) showed more benefit and 16 (32%) showed less benefit after exclusion of non-English-anguage trials None of the meta-analyses changed statistical significance at the 5% level	Compared with English-language trials, non-English-language trials included fewer participants but were more likely to show statistically significant results, and tended to be of lower methodological quality

Appendix 13
Citation bias – included empirical studies

Study	Methods	Main findings	Notes
Chapman *et al.* 2009[162]	An examination of citations of 42 studies on tobacco smoking among schizophrenia subjects	A 10% increase in reported prevalence of smoking was associated with a 61% (95% CI: 30% to 98%) increase in citation rate. After adjusting for journal impact factor, a 10% increase in prevalence of smoking was associated with a 28% (1% to 62%) increase in citation rate	Prevalence of smoking in a special population with schizophrenia
Callaham *et al.* 2002[161]	Citations of 204 published articles originally submitted to a 1991 emergency medicine meeting were identified by searching Science Citation Index. Calculated number of times an article was cited per year and mean impact factor	Mean citations per year = 2.04 (95% CI: 1.6 to 2.4) in 440 different journals. Factors associated with citation: JIF, newsworthiness score, quality score. Positive outcome bias not evident	Examined studies from one meeting only
Kjaergard and Gluud 2002[163]	530 RCTs of hepato-biliary diseases. Trial results and quality were assessed. Number of citations for each trial was obtained from Science Citation Index. Linear regression analysis conducted	Positive association between a statistically significant result and the citation frequency (β = 1.21; 95% CI: 1.10 to 1.33). Disease area and allocation concealment were also significantly predictors of the citation frequency	Several confounders like accessibility and popularity of a journal, which might be associated with citation rates, were not included
Nieminen *et al.* 2007[164]	368 research papers published in 1996 in four psychiatric journals. Regression analysis to relate citation frequency to statistical significance (*p* < 0.05), adjusting for confounders	Median number of citations: 33 for significant results (*n* = 287) and 16 for non-significant results (*n* = 81). Citation rate ratio for papers reporting '*p* < 0.05' on primary outcome was 1.63 (95% CI: 1.32 to 2.02)	'Self-citations' were excluded
Schmidt *et al.* 2005[165]	A review of 42 reviews that included trials on clinical effects of physical interventions on house-dust mite antigens. Compared the proportion of trial references with significant results between cited references and all trials available	Of the 38 positive reviews and in terms of selection of references, 10 reviews were neutral, 27 reviews had a positive selection and one a negative selection. The four reviews that did not recommend physical interventions all had a negative selection of references	The study identified severe bias in selection of references in narrative reviews

Appendix 14

Reasons given by investigators for studies not being published

Study	Reasons for non-publication			
Cooper 1997[35] Cohort of studies submitted for review by a human subjects committee Mixed design	**Why the study was not prepared for a journal publication (*n* = 159)** Publication not an aim: 48% Class project only: 30% Assistant lost interest: 26% No significant results: 22% Results were not interesting: 20% Design or operational problems: 12% Researchers did not recall: 6% Others lost interest: 2%			
Decullier 2005[28] Cohort of research protocols Mixed designs	**Reasons given by investigators for not publishing (*n* = 102)** Negative results: 27 (26%) Writing or submission in progress: 23 (23%) Published in other forms: 23 (23%) Paper rejected: 5 (5%) Other reasons: 17 (17%) Not available: 7 (7%)			
Dickersin 1992[20] Cohort of research protocols Mixed design	**Main reasons**	**Total**	**School of Medicine**	**School of Public Health**
	Total unpublished studies	*124 (100%)*	*65*	*59*
	Manuscript rejected by journal	6 (5%)	2	4
	Total not submitted	*118 (95%)*	*63*	*55*
	Results not interesting	37 (30%)	26	11
	Design or operational problems	40 (32%)	17	23
	Publication not an aim	16 (13%)	8	8
	Other reasons	25 (20%)	12	13
Dickersin 1993[21] Cohort of research protocols Clinical trials	**Total unpublished trials: 100% (*n* = 14)** Not interesting or no time: 42.8% Co-investigator/operational problems: 37.5% Data analysis not completed: 14.3% Rejected by journal: 0% No reason given: 7.1%			

Study	Reasons for non-publication				
Easterbrook 1991[22]	**Total**	**Significant**	**Non-significant**	**Null**	
		(*n* = 78)	(*n* = 23)	(*n* = 12)	(*n* = 43)
Cohort of research protocols	Submitted or published elsewhere	35 (45%)	20	4	11
Mixed design	Not submitted/published at all	43 (55%)	3	8	32
	Null results	26 (33%)			26
	Methodology or logistic problem	21 (27%)	3	5	13
	Sponsor has control of data	19 (24%)	11	2	6
	Analysis incomplete	19 (24%)	10	2	7
	Manuscript rejected	16 (21%)	7	1	8
	Publication not aim of study	13 (17%)	6	4	3
	Too busy or lost interest	11 (14%)	3	5	3
	Unimportant results	10 (13%)	2	1	7
	Co-investigator left	5 (6%)	0	1	4

Camacho 2005[232]	**Factors affecting the publication of phase I clinical trials**			
Survey of authors of abstracts presented at annual meeting of the American Society of Clinical Oncology in 1997	**Reason**	**Novel agent**	**Non-novel**	**Total**
		(*n* = 36)	(*n* = 29)	(*n* = 65)
	Lack of time	12	11	23 (35%)
Clinical trials	Manuscript in preparation	10	5	15 (23%)
	Relocation of authors	11	3	14 (22%)
	Incomplete study	6	7	13 (20%)
	Results considered not interesting	7	4	11 (17%)
	Rejection from peer-reviewed journal	3	2	5 (8%)
	Manuscript submitted	1	5	6 (9%)
	Not in the sponsor's interest	2	1	3 (5%)
	Conflict of interest	1	0	1 (2%)
	Other	1	1	2 (3%)

Novel – agents not approved by the Food and Drug Administration at the time of submission

Non-novel – at least one agent approved

De Bellefeuille 1992[50]

Cohort of meeting abstracts

Mixed design

Reasons for non-publication (based on *n* = 41 respondents)

Lack of time/other resources: 13 (32%)

Insufficient priority: 9 (22%)

Incomplete study with intent to publish eventually: 5 (12%)

Article not accepted for publication: 4 (10%)

Modification of data after submission of abstract: 1 (2%)

Other: 12 (29%)

Study	Reasons for non-publication
Hartling 2004[233] Survey of authors of abstracts presented at the Society for Paediatric Research meetings from 1992 to 1995 Clinical trials	**Total number of unpublished studies *n* = 47** **Total number of unsubmitted studies *n* = 39** **Important reasons given by authors for non-publication:** Not enough time (*n* = 39): 56.4% Too much trouble with co-authors (*n* = 38): 28.9% Thought that journal was unlikely to accept (*n* = 38): 26.3% Results were not statistically significant (*n* = 38): 23.7% Results were not important enough (*n* = 38): 18.4% Others published with similar findings (*n* = 38): 15.8% Study quality poor (*n* = 37): 13.5% Not worth the trouble (*n* = 37): 10.8% Results did not support the hypothesis (*n* = 38): 5.3%
Hashkes 2003[67] Cohort of meeting abstracts Mixed design	**Reasons for non-submission of abstract for publication (*n* = 97)** Case report: 8 (8%) Previously reported: 5 (5%) Non-positive results: 2 (2%) Methodological problems: 2 (2%) Desire to expand paper: 42 (43%) Low priority or lack of time: 47 (48%) Fear of rejection: 13 (13%) Author moved or passed away: 4 (4%) No decision on journal: 1 (1%)
Hopewell 2001[234] Cohort of abstracts at meetings on systematic reviews Methodological research	**Reasons for non-publication of abstracts (*n* = 22)** Low priority or too busy: 9 (24%) Not deemed appropriate: 7 (19%) Findings became rapidly outdated: 2 (5%) Rejected by journal as not deemed relevant to the general readership: 1 Subject area was too specific with limited interest to a wider audience: 1 Internal Cochrane issue: 1 Concerns over unity of approach: 1 Note: Authors of 15 non-published abstracts did not given a reason. These 15 unpublished abstracts were not included
Krzyzanowska 2003[70] Survey of authors of abstracts presented at the annual meeting of the American Society of Clinical Oncology 1989–98 Clinical trials	**Reasons for lack of publication (based on 40 responses)** Lack of time, funds, or other resources: 14 (35%) Study incomplete, with eventual intent to publish: 6 (15%) Article submitted, but not accepted for publication: 5 (13%) Manuscript in preparation: 5 (13%) Manuscript under review: 4 (10%) Insufficient priority to warrant publication: 4 (10%) Other: 5 (13%) Not provided: 6 (15%)

Study	Reasons for non-publication
Sanossian 2006[72] Survey of authors of research abstracts presented at the annual International Stroke Conference in 2000 Mixed design	**Reasons for non-publication (*n* = 74)** No time: 28 (38%) Low priority: 11 (15%) Co-author responsibility or lack of participation: 10 (14%) Study ongoing: 8 (11%) Methodological limitations: 6 (8%) Different version published: 3 (4%) Other similar articles published: 2 (3%) Does not recall: 1 (1%) No reason given: 5 (7%)
Scherer *et al.* 1994[54]	**Number of unpublished abstracts of RCTs (*n* = 32)** Incomplete studies: 16% Manuscript rejected: 19% No time to prepare: 28% Problem of study design: 9%
Sprague 2003[236] Cohort of meeting abstracts and survey of authors Mixed design	**Reasons for failure to submit a manuscript to a journal**

Reason (*n* = 71)[a]	No. of responses
No time to prepare for publication	33 (46%)
Study is still ongoing	22 (31%)
Responsibility for manuscript belongs to a co-author	14 (20%)
Difficulty with co-authors (lack of participation)	12 (17%)
Pursuit of publication given a low priority	9 (13%)
Low likelihood of acceptance for publication because of methodological limitations of study (e.g. weak study design or small sample size)	9 (13%)
Other papers with similar findings already published	3 (4%)
Plan to submit paper for publication	3 (3 (4%)
Results not important enough	1 (1%)
Statistical analysis was not positive	1 (1%)
Low likelihood of acceptance by journal because of insufficient interest to readers	1 (1%)
Different version of data published	1 (1%)

Study	Reasons for non-publication
Vuckovic-Dekic 2001[237] Survey of Serbian authors of abstracts presented at Congress of the Balkan Union of Oncology 1996–8 Mixed design	**Reasons for not submitting studies (*n* = 21)** Not enough time: 10 (48%) Thought journals unlikely to accept: 2 (10%) Results not important enough: 1 (5%) Other papers with similar findings: 1 (5%) Too much trouble with co-authors: 1 (5%) Other reasons: 6 (29%)

Study	Reasons for non-publication
Weber 1998[230] Cohort of meeting abstracts Mixed designs	**Reasons for failure to submit to a journal (*n* = 179)** Not enough time: 74 (41%) Thought journals unlikely to accept: 35 (20%) Results not important enough: 21 (12%) Trouble with co-authors: 16 (9%) Not worth the trouble: 13 (7%) Other papers with similar findings: 11 (6%) Statistical analysis not positive: 7 (4%) Other reasons: 40 (22%)
Blumenthal et *al.* 1997[231] Survey of life sciences faculty members at 50 universities in the USA	**Reasons given for delay to publication (*n* = 412)** Patent application submission: 46% Protection of scientific lead: 31% Patent negotiation: 26% Resolution of intellectual property ownership: 17% Slow dissemination of undesired results: 28%
Dickersin 1987[228] Survey of authors of published trials Clinical trials	**Total unsubmitted trials: 100% (*n* = 102)** Analysis in progress: 14.7% Results negative: 34.3% Lack of interest: 15.7% Sample size or poor methodology: 4.9% Controversy: 2.9% Other or unknown: 27.5%
Machan et *al.* 2006[234] (conference abstract) An email survey of members of European Federation of Medical Informatics and International Medical Informatics Association Evaluation studies	**Unpublished evaluation studies (*n* = 104)** Generalisability limited: 26% Study not yet finished: 18% No time for writing: 11% Results seemed not of interest to others: 10% Methods inadequate/sampling insufficient: 9% Organisations prohibited publication: 9% Rejected by journal: 6% Results too negative: 5% No interest in academic output: 5% Evaluation of first prototype only: 4%
Misakian 1998[30]	**Reasons for unpublished results (*n* = 59) of passive smoking** Ongoing data collection or analysis: 56% Lack of time: 44% Competing priorities: 19% Statistically non-significant results: 3% Manuscript rejected: 7%
Rotton et *al.* 1995[229]	**Proportion of reasons given by 468 authors for not publishing** Failure to replicate: 5% Manuscript rejected: 33% Non-hypothesised results: 5% Inexplicable results: 22% Non-significance: 60%

a   The number of responses is greater than the number of respondents because respondents were allowed to choose more than one response

There may be two or more reasons for each unpublished study.

Appendix 15
Study findings and the acceptance of submitted manuscripts

Study	Methods	Main findings
Olson *et al.* 2002[81]	Cohort study of 745 manuscripts of controlled trials submitted to *JAMA* from 02/1996 to 08/1999 *Outcome classification:* 1. Statistically significant ($p < 0.05$) for the primary outcome 2. Statistically non-significant 3. Unclear	*Proportion of studies with different results:* 51.4% ($n = 383$) with significant results 45.7% ($n = 341$) with non-significant results 2.8% ($n = 21$) with unclear results *Acceptance rate:* 20.4% (78/383) for significant results 15.0% (51/341) for non-significant results 19.0% (4/21) for unclear results Logistic regression analysis: significant vs non-significant results OR = 1.30 (95% CI: 0.87 to 1.96)
Lee *et al.* 2006[78]	Cohort study of 1107 manuscripts of original research (including qualitative research, excluding single case reports) submitted to *BMJ*, *Lancet* and *Annals of Internal Medicine* between 01/2003 and 03/2003 and between 11/2003 and 02/2004 *Outcome classification:* 1. Statistically significant ($p < 0.05$) for the primary outcome 2. Non-significant	*Proportion of different statistical results:* 86.8% ($n = 718$) with significant results 13.2% ($n = 109$) with non-significant results *Acceptance rate:* 4.9% (35/718) for significant results 6.4% (7/109) for non-significant results Multivariate analysis: OR = 0.83 (95% CI: 0.34 to 1.96)
Lynch *et al.* 2007[79]	Cohort study of 209 manuscripts of original research on hip or knee arthroplasty submitted to the *Journal of Bone and Joint Surgery (American Volume)* between 01/2004 and 06/2005 *Outcome classification:* 1. Positive or favourable or significant 2. Negative or non-supportive or no difference 3. Not analysable – unclear	*Proportion of studies with different results:* 70.8% ($n = 148$) with positive results 23.4% ($n = 49$) with negative results 5.7% ($n = 12$) with unclear results *Acceptance rate:* 30.4% (45/148) for positive results 36.7% (18/49) for negative results 8.3% (1/12) for unclear results Difference in publication rate between positive and negative outcomes was not statistically significant ($p = 0.41$)
Okike *et al.* 2008[80]	Cohort study of 855 manuscripts as scientific articles submitted to the *Journal of Bone and Joint Surgery (American Version)* between 01/2004 and 06/2005 *Outcome classification:* 1. Positive – favoured experimental item 2. Negative – favoured existing standard of care over the experimental item 3. Neutral – no difference 4. Not applicable	*Proportion of studies with different results:* 72.5% ($n = 620$) with positive results 12.3% ($n = 105$) with negative results 15.2% ($n = 130$) with neutral results *Acceptance rate:* 21.3% (132/620) for positive results 21.0% (22/105) for negative findings 24.6% (32/130) for neutral results Multivariate analysis: positive vs non-positive OR = 0.92 (95% CI: 0.62 to 1.35)

Appendix 16

Case studies indicating pharmaceutical companies or industry research sponsorship as a source of publication and related biases

Cases	Brief descriptions
Nathan and Weatherall 1999[305] and 2002[306] Publication suppression: Deferiprone for the prevention of iron toxicity in patients with thalassaemia	A company-sponsored trial in 1989 found that the drug might be harmful. The company took legal action against the investigator, Dr Nancy Olivieri, in order to stop the disclosure of the negative finding
Rennie 1997[276,309] Publication suppression: Bioequivalence of brand name and generic forms of thyroxine sodium	A company-sponsored study in 1987 by Dong *et al.* showed bioequivalence of generic and brand name levothyroxine. The publication of the trial was suppressed for 7 years by the pharmaceutical company due to deleterious effect of results on price of company's product
Skolnick 1998[308] Publication suppression: *HTA* report on cholesterol-lowering statin drugs	A pharmaceutical company tried unsuccessfully to suppress the publication of findings from a health technology assessment on cholesterol-lowering statin drugs by the Canadian Coordinating Office of Health Technology Assessment in 1997
Millstone *et al.* 1994[304] Publication suppression: Increased somatic cells in cow's milk and bovine somatotrophin (BST)	Millstone *et al.* reported that their meta-analysis with unsupportive results of BST was blocked by a pharmaceutical company using legal rights over raw data
Shuchman 1999[307] Publication suppression: Ontario Ministry of Health: omeprazole and draft prescribing guidelines	Shuchman reported that a company threatened legal action over draft prescribing guidelines that concluded that all proton pump inhibitors had equivalent effect on peptic ulcers and gastro-oesophageal reflux disease. However, the company responded by saying that the company 'is not pursuing any legal action against any physician'[534]
McCarthy 2000[310–316] Publication suppression: Remune (HIV-1 immunogen) for HIV infection	A company-sponsored trial found no difference in efficacy between the vaccine and placebo. The manufacturer of Remune attempted to block the paper's publication because authors refused to include a post-hoc subgroup analysis
Lauritsen 1987[317] Non-publication: Prostaglandin for gastric ulcer	A company-sponsored trial compared prostaglandin analogue with ranitidine for gastric ulcer, and stopped in 1985. Ranitidine was better than prostaglandin in all centres except one. One of the trial centres in Denmark had asked for a copy of the report in March 1986 but had still not received the full report by April 1987
Symmonds *et al.* 2004[318] and Panahloo 2004[336] Non-publication: Neuraminidase inhibitors (oseltamivir) for asthmatic children suffering from influenza	Two trials showed no significant difference in time to freedom from illness between children taking the drug and placebo. Data were submitted to European Marketing Authorisation, but not published
Wilmshurst 1986[321] and 1987[322] Non-publication: Amrinone	A company discontinued trials that showed negative results and failed to report adverse events of amrinone

Cases	Brief descriptions
van Heteren 2001[319] Non-publication: Deep venous thrombosis after using third generation oral contraceptive pills	Results of a study on the risk of deep venous thrombosis after using third generation contraceptive pills were submitted to the European Medicines Evaluation Agency in 1999, but remained unpublished. The company stated that 'the study was not submitted for publication because it was felt that the study did not offer any new scientific information'
van Veldhuisen and Poole-Wilson 2001[320] Non-publication: 'Negative' drug trials in patients with chronic heart failure	van Veldhuisen and Poole-Wilson (2001) discussed three unpublished trials that were terminated prematurely because of increased mortality or adverse effects. These trials were presented at conferences but not fully published
Gottlieb 2001[324,326,327,535] Selective publication: Celecoxib for arthritis	A trial published in *JAMA* in 2000 concluded that celecoxib was associated with a lower incidence of symptomatic ulcers and ulcer complications compared with ibuprofen and diclofenac. However, the publication was based on the 6-month data. Unpublished 12-month data (submitted to FDA) was much less favourable for celecoxib
Topol 2004[214,215,217,323] Selective publication: Rofecoxib for arthritis	A trial of rofecoxib vs naproxen in patients with rheumatoid arthritis did not include three cases of myocardial infarctions (MIs) in the rofecoxib arm. The authors of the trial explained that the three MIs were observed after the cut-off date for reporting cardiovascular events[216,536]
Steinman 2006[331] Selective publication: Internal industry documents and the promotion of gabapentin for off-label uses	Steinman *et al.* reviewed internal industry documents about the promotion of gabapentin for off-label uses. They found that the company's 'management expressed concern that negative results could harm promotional efforts, and several documents indicate the intention to publish and publicis results only if they reflected favourably on gabapentin'
Garland 2004[220] Selective publication: Paroxetine and venlafaxine for depression in children and adolescents	Garland reported that none of the large negative trials of paroxetine and venlafaxine in children and adolescents were published. The GlaxoSmithKline internal document revealed that the company experts advised staff to withhold data about SSRI use in children.[222] GSK faces US lawsuit over concealment of trial results in 2004[223]
Reines 2004[330,337] Selective publication: Rofecoxib for Alzheimer's disease or cognitive impairment	The two published trials of rofecoxib for Alzheimer's disease only mentioned on-treatment mortality in the text without any statistical analyses, and concluded that rofecoxib is well tolerated. However, the company's unpublished intention-to-treat analyses and the independent analyses based on data provided by the sponsor in the New Jersey Vioxx litigation found a statistically significant increase in total mortality (HR 2.99; 95% CI: 1.55 to 5.56; and HR 2.13; 95% CI: 1.55 to 5.77 respectively)[219]
Whittington *et al.* 2004[136] Selective publication: Selective serotonin reuptake inhibitors (SSRIs) in children and adolescents	Whittington *et al.* compared results of published trials and unpublished data. They concluded that published data presented a favourable risk-benefit profile, whereas unpublished data indicated that risks could outweigh benefits of these drugs (except fluoxetine) in children and adolescents
Applegate *et al.* 1997[328] Selective publication: Isradipine	Several investigators of a multicentre trial of isradipine dropped out when the paper was in preparation, because 'the sponsor of the study was attempting to wield undue influence on the nature of the final paper'

Cases	Brief descriptions
Metcalfe *et al.* 2008[335] Selective (or delayed) publication: Trastuzumab (Herceptin®) for early breast cancer	An industry-sponsored three-arm trial (NCCTG-N9831) directly compared sequential, concurrent Herceptin, and usual care control. According to a conference abstract in 2005, interim results indicated concurrent Herceptin was more effective than sequential therapy (HR 0.64; 95% CI: 0.46 to 0.91). However, a journal paper in 2005 reported only data on concurrent therapy and the control, without including sequential-group data. Because of these missing data, 'sequential trastuzumab seems more effective than it probably is'.[335] The principal investigator for the trial responded that the publication of concurrent therapy data was according to an analysis plan prespecified while data on sequential therapy were not sufficiently mature[537]
Lenzer 2002[329] Delayed publication: Alteplase (a thrombolytic agent) for acute ischaemic stroke	A trial found that alteplase did not improve stroke recovery and increased mortality. The negative result was not published for 6 years after the trial's completion
Lenzer 2002[329] Delayed publication: Release of results from a trial on ezetimibe – a cholesterol-lowering drug	Negative results of a trial on ezetimibe were released by the company only after a US Congressional inquiry was set up to look into why the results had not been published 2 years after the study was completed
Alasbali *et al.* 2009[332] Discrepancy between results and abstract conclusions Topical prostaglandins	Alasbali *et al.* examined the discrepancy between the statistical significance of the publication's main outcome measure and its abstract conclusions. The published abstract conclusion was not consistent with the results of the main outcome measure in 18 of 29 industry-funded studies compared with zero of 10 non-industry-funded studies on the efficacy of topical prostaglandin analogues

Appendix 17
List of 347 reviews assessed

Treatment reviews (*n* = 100)

Abdulla J, Kober L, Christensen E, Torp-Pedersen C. Effect of beta-blocker therapy on functional status in patients with heart failure – a meta-analysis. *Eur J Heart Fail* 2006;**8**:522–31.

Agosti R, Duke RK, Chrubasik JE, Chrubasik S. Effectiveness of Petasites hybridus preparations in the prophylaxis of migraine: a systematic review. *Phytomedicine* 2006;**13**:743–6.

Auperin A, Le Pechoux C, Pignon JP, Koning C, Jeremic B, Clamon G, *et al*. Concomitant radio-chemotherapy based on platin compounds in patients with locally advanced non-small cell lung cancer (NSCLC): a meta-analysis of individual data from 1764 patients. *Ann Oncol* 2006;**17**:473–83.

Ayalon L, Gum AM, Feliciano L, Arean PA. Effectiveness of nonpharmacological interventions for the management of neuropsychiatric symptoms in patients with dementia: a systematic review. *Arch Intern Med* 2006;**166**:2182–8.

Bainbridge D, Cheng DC, Martin JE, Novick R. NSAID-analgesia, pain control and morbidity in cardiothoracic surgery. *Can J Anaesth* 2006;**53**:46–59.

Beelmann A, Losel F. Child social skills training in developmental crime prevention: effects on antisocial behavior and social competence. *Psicothema* 2006;**18**:603–10.

Ben Amar M. Cannabinoids in medicine: A review of their therapeutic potential. *J Ethnopharmacol* 2006;**105**:1–25.

Bergman R, Parkes M. Systematic review: the use of mesalazine in inflammatory bowel disease. *Aliment Pharmacol Ther* 2006;**23**:841–55.

Bouza C, Lopez T, Magro A, Navalpotro L, Amate JM. Efficacy and safety of balloon kyphoplasty in the treatment of vertebral compression fractures: a systematic review. *Eur Spine J* 2006;**15**:1050–67.

Bria E, Ciccarese M, Giannarelli D, Cuppone F, Nistico C, Nuzzo C, *et al*. Early switch with aromatase inhibitors as adjuvant hormonal therapy for postmenopausal breast cancer: pooled-analysis of 8794 patients. *Cancer Treat Rev* 2006;**32**:325–32.

Brok J, Gluud LL, Gluud C. Ribavirin monotherapy for chronic hepatitis C infection: a Cochrane Hepato-

Biliary Group systematic review and meta-analysis of randomized trials. *Am J Gastroenterol* 2006;**101**:842–7.

Bujko K, Kepka L, Michalski W, Nowacki MP. Does rectal cancer shrinkage induced by preoperative radio(chemo) therapy increase the likelihood of anterior resection? A systematic review of randomised trials. *Radiother Oncol* 2006;**80**:4–12.

Busquets JM, Hwang PH. Endoscopic resection of sinonasal inverted papilloma: a meta-analysis. *Otolaryngol Head Neck Surg* 2006;**134**:476–82.

Chavez-Tapia NC, Barrientos-Gutierrez T, Tellez-Avila FI, Sanchez-Avila F, Montano-Reyes MA, Uribe M. Insulin sensitizers in treatment of nonalcoholic fatty liver disease: Systematic review. *World J Gastroenterol* 2006;**12**:7826–31.

Clarke J, van Tulder M, Blomberg S, de Vet H, van der Heijden G, Bronfort G. Traction for low back pain with or without sciatica: an updated systematic review within the framework of the Cochrane collaboration. *Spine* 2006;**31**:1591–9.

Collins CE, Warren J, Neve M, McCoy P, Stokes BJ. Measuring effectiveness of dietetic interventions in child obesity: a systematic review of randomized trials. *Arch Pediatr Adolesc Med* 2006;**160**:906–22.

Cook SA, Rosser R, Salmon P. Is cosmetic surgery an effective psychotherapeutic intervention? A systematic review of the evidence. *J Plast Reconstr Aesthet Surg* 2006;**59**:1133–51.

Dahlof CG, Pascual J, Dodick DW, Dowson AJ. Efficacy, speed of action and tolerability of almotriptan in the acute treatment of migraine: pooled individual patient data from four randomized, double-blind, placebo-controlled clinical trials. *Cephalalgia* 2006;**26**:400–8.

Davey P, Brown E, Fenelon L, Finch R, Gould I, Holmes A, *et al*. Systematic review of antimicrobial drug prescribing in hospitals. *Emerg Infect Dis* 2006;**12**:211–6.

De Schryver EL, Algra A, van Gijn J. Dipyridamole for preventing stroke and other vascular events in patients with vascular disease. *Cochrane Database Syst Rev* 2006(2):CD001820.

Demyttenaere S, Feldman LS, Fried GM. Effect of pneumoperitoneum on renal perfusion and function: a systematic review. *Surg Endosc* 2007;**21**:152–60.

169

Dibra A, Kastrati A, Alfonso F, Seyfarth M, Perez-Vizcayno MJ, Mehilli J, *et al*. Effectiveness of drug-eluting stents in patients with bare-metal in-stent restenosis: meta-analysis of randomized trials. *J Am Coll Cardiol* 2007;**49**:616–23.

Dos Santos-Neto LL, de Vilhena Toledo MA, Medeiros-Souza P, de Souza GA. The use of herbal medicine in Alzheimer's disease-a systematic review. *Evid Based Complement Alternat Med* 2006;**3**:441–5.

Eisenberg E, McNicol ED, Carr DB. Efficacy of mu-opioid agonists in the treatment of evoked neuropathic pain: Systematic review of randomized controlled trials. *Eur J Pain* 2006;**10**:667–76.

Elahi MM, Khan JS. Living with off-pump coronary artery surgery: evolution, development, and clinical potential for coronary heart disease patients. *Heart Surg Forum* 2006;**9**:E630–7.

Elder JS, Diaz M, Caldamone AA, Cendron M, Greenfield S, Hurwitz R, *et al*. Endoscopic therapy for vesicoureteral reflux: a meta-analysis. I. Reflux resolution and urinary tract infection. *J Urol* 2006;**175**:716–22.

Elkins MR, Jones A, van der Schans C. Positive expiratory pressure physiotherapy for airway clearance in people with cystic fibrosis. *Cochrane Database Syst Rev* 2006(2):CD003147.

Faber E, Kuiper JI, Burdorf A, Miedema HS, Verhaar JA. Treatment of impingement syndrome: a systematic review of the effects on functional limitations and return to work. *J Occup Rehabil* 2006;**16**:7–25.

Fabrizi F, Dixit V, Martin P. Meta-analysis: anti-viral therapy of hepatitis B virus-associated glomerulonephritis. *Aliment Pharmacol Ther* 2006;**24**:781–8.

Falagas ME, Manta KG, Ntziora F, Vardakas KZ. Linezolid for the treatment of patients with endocarditis: a systematic review of the published evidence. *J Antimicrob Chemother* 2006;**58**:273–80.

Finer NN, Barrington KJ. Nitric oxide for respiratory failure in infants born at or near term. *Cochrane Database Syst Rev* 2006(4):CD000399.

Francis J, Johnson B, Niehaus M. Quality of life in patients with implantable cardioverter defibrillators. *Indian Pacing Electrophysiol J* 2006;**6**:173–81.

Fung AT, Reid SE, Jones MP, Healey PR, McCluskey PJ, Craig JC. Meta-analysis of randomised controlled trials comparing latanoprost with brimonidine in the treatment of open-angle glaucoma, ocular hypertension or normal-tension glaucoma. *Br J Ophthalmol* 2007;**91**:62–8.

Gafter-Gvili A, Paul M, Fraser A, Leibovici L. Effect of quinolone prophylaxis in afebrile neutropenic patients on microbial resistance: systematic review and meta-analysis. *J Antimicrob Chemother* 2007;**59**:5–22.

Gagner M, Boza C. Laparoscopic duodenal switch for morbid obesity. *Expert Rev Med Devices* 2006;**3**:105–12.

Gajdos P, Chevret S, Toyka K. Intravenous immunoglobulin for myasthenia gravis. *Cochrane Database Syst Rev* 2006(2):CD002277.

Getz M, Hutzler Y, Vermeer A. Effects of aquatic interventions in children with neuromotor impairments: a systematic review of the literature. *Clin Rehabil* 2006;**20**:927–36.

Glasmacher A, Hahn C, Hoffmann F, Naumann R, Goldschmidt H, von Lilienfeld-Toal M, *et al*. A systematic review of phase-II trials of thalidomide monotherapy in patients with relapsed or refractory multiple myeloma. *Br J Haematol* 2006;**132**:584–93.

Goonetilleke KS, Siriwardena AK. Systematic review of peri-operative nutritional supplementation in patients undergoing pancreaticoduodenectomy. *JPancreas* 2006;**7**:5–13.

Gupta VK. Botulinum toxin – a treatment for migraine? A systematic review. *Pain Med* 2006;**7**:386–94.

Hadley G, Derry S, Moore RA. Imiquimod for actinic keratosis: systematic review and meta-analysis. *J Invest Dermatol* 2006;**126**:1251–5.

Helin RD, Angeles ST, Bhat R. Octreotide therapy for chylothorax in infants and children: A brief review. *Pediatr Crit Care Med* 2006;**7**:576–9.

Hickey BE, Francis D, Lehman MH. Sequencing of chemotherapy and radiation therapy for early breast cancer. *Cochrane Database Syst Rev* 2006(4):CD005212.

Ho KM, Sheridan DJ. Meta-analysis of frusemide to prevent or treat acute renal failure. *BMJ* 2006;**333**:420.

Huang HY, Caballero B, Chang S, Alberg AJ, Semba RD, Schneyer CR, *et al*. The efficacy and safety of multivitamin and mineral supplement use to prevent cancer and chronic disease in adults: a systematic review for a National Institutes of Health state-of-the-science conference. *Ann Intern Med* 2006;**145**:372–85.

Ingram C, Courneya KS, Kingston D. The effects of exercise on body weight and composition in breast cancer survivors: an integrative systematic review. *Oncol Nurs Forum* 2006;**33**:937–47; quiz 948–50.

Issa AM, Mojica WA, Morton SC, Traina S, Newberry SJ, Hilton LG, *et al*. The efficacy of omega-3 fatty acids on cognitive function in aging and dementia: a systematic review. *Dement Geriatr Cogn Disord* 2006;**21**:88–96.

Jimbo M, Nease DE Jr, Ruffin MTt, Rana GK. Information technology and cancer prevention. *CA Cancer J Clin* 2006;**56**:26–36; quiz 48–9.

Johnson CE, Danhauer JL, Reith AC, Latiolais LN. A systematic review of the nonacoustic benefits of bone-anchored hearing AIDS. *Ear Hear* 2006;**27**:703–13.

Kalanda GC, Hill J, Verhoeff FH, Brabin BJ. Comparative efficacy of chloroquine and sulphadoxine–pyrimethamine in pregnant women and children: a meta-analysis. *Trop Med Int Health* 2006;**11**:569–77.

Keeley EC, Boura JA, Grines CL. Comparison of primary and facilitated percutaneous coronary interventions for ST-elevation myocardial infarction: quantitative review of randomised trials. *Lancet* 2006;**367**:579–88.

Kelley GA, Kelley KS. Aerobic exercise and HDL2-C: a meta-analysis of randomized controlled trials. *Atherosclerosis* 2006;**184**:207–15.

Khanna A, Walker GR, Livingstone AS, Arheart KL, Rocha-Lima C, Koniaris LG. Is adjuvant 5-FU-based chemoradiotherapy for resectable pancreatic adenocarcinoma beneficial? A meta-analysis of an unanswered question. *J Gastrointest Surg* 2006;**10**:689–97.

Kingma JJ, de Knikker R, Wittink HM, Takken T. Eccentric overload training in patients with chronic Achilles tendinopathy: a systematic review. *Br J Sports Med* 2007;**41**:e3.

Kirby D, Obasi A, Laris BA. The effectiveness of sex education and HIV education interventions in schools in developing countries. *World Health Organ Tech Rep Ser* 2006;**938**:103–50; discussion 317–41.

Kleiner-Fisman G, Herzog J, Fisman DN, Tamma F, Lyons KE, Pahwa R, *et al.* Subthalamic nucleus deep brain stimulation: summary and meta-analysis of outcomes. *Mov Disord* 2006;**21**(Suppl. 14):S290–304.

Kyrgiou M, Salanti G, Pavlidis N, Paraskevaidis E, Ioannidis JP. Survival benefits with diverse chemotherapy regimens for ovarian cancer: meta-analysis of multiple treatments. *J Natl Cancer Inst* 2006;**98**:1655–63.

Lander JA, Weltman BJ, So SS. EMLA and amethocaine for reduction of children's pain associated with needle insertion. *Cochrane Database Syst Rev* 2006;**3**:CD004236.

Law M, Rudnicka AR. Statin safety: a systematic review. *Am J Cardiol* 2006;**97**(8A):52C–60C.

Le Corvoisier P, Hittinger L, Chanson P, Montagne O, Macquin-Mavier I, Maison P. Cardiac effects of growth hormone treatment in chronic heart failure: A meta-analysis. *J Clin Endocrinol Metab* 2007;**92**:180–5.

Lee SJ, Schover LR, Partridge AH, Patrizio P, Wallace WH, Hagerty K, *et al.* American Society of Clinical Oncology recommendations on fertility preservation in cancer patients. *J Clin Oncol* 2006;**24**:2917–31.

Liberatore Rdel R Jr, Damiani D. Insulin pump therapy in type 1 diabetes mellitus. *J Pediatr (Rio J)* 2006;**82**:249–54.

Lodi G, Sardella A, Bez C, Demarosi F, Carrassi A. Interventions for treating oral leukoplakia. *Cochrane Database Syst Rev* 2006(4):CD001829.

Lyman GH, Glaspy J. Are there clinical benefits with early erythropoietic intervention for chemotherapy-induced anemia? A systematic review. *Cancer* 2006;**106**:223–33.

McCart MR, Priester PE, Davies WH, Azen R. Differential effectiveness of behavioral parent-training and cognitive-behavioral therapy for antisocial youth: a meta-analysis. *J Abnorm Child Psychol* 2006;**34**:527–43.

McPhail MJ, Abu-Hilal M, Johnson CD. A meta-analysis comparing suprapubic and transurethral catheterization for bladder drainage after abdominal surgery. *Br J Surg* 2006;**93**:1038–44.

Meijering S, Corstjens AM, Tulleken JE, Meertens JH, Zijlstra JG, Ligtenberg JJ. Towards a feasible algorithm for tight glycaemic control in critically ill patients: a systematic review of the literature. *Crit Care* 2006;**10**:R19.

Milne AC, Avenell A, Potter J. Meta-analysis: protein and energy supplementation in older people. *Ann Intern Med* 2006;**144**:37–48.

Murphy MH, Nevill AM, Murtagh EM, Holder RL. The effect of walking on fitness, fatness and resting blood pressure: a meta-analysis of randomised, controlled trials. *Prev Med* 2007;**44**:377–85.

Niebauer K, Dewilde S, Fox-Rushby J, Revicki DA. Impact of omalizumab on quality-of-life outcomes in patients with moderate-to-severe allergic asthma. *Ann Allergy Asthma Immunol* 2006;**96**:316–26.

Noble J, Ellis PM, Mackay JA, Evans WK. Second-line or subsequent systemic therapy for recurrent or progressive non-small cell lung cancer: a systematic review and practice guideline. *J Thorac Oncol* 2006;**1**:1042–58.

Oktay K, Cil AP, Bang H. Efficiency of oocyte cryopreservation: a meta-analysis. *Fertil Steril* 2006;**86**:70–80.

Oliver D, Connelly JB, Victor CR, Shaw FE, Whitehead A, Genc Y, *et al.* Strategies to prevent falls and fractures in hospitals and care homes and effect of cognitive impairment: systematic review and meta-analyses. *BMJ* 2007;**334**:82.

Papakostas GI, Fava M. A meta-analysis of clinical trials comparing milnacipran, a serotonin–norepinephrine

reuptake inhibitor, with a selective serotonin reuptake inhibitor for the treatment of major depressive disorder. *Eur Neuropsychopharmacol* 2007;**17**:32–6.

Paul M, Yahav D, Fraser A, Leibovici L. Empirical antibiotic monotherapy for febrile neutropenia: systematic review and meta-analysis of randomized controlled trials. *J Antimicrob Chemother* 2006;**57**:176–89.

Petignat P, du Bois A, Bruchim I, Fink D, Provencher DM. Should intraperitoneal chemotherapy be considered as standard first-line treatment in advanced stage ovarian cancer? *Crit Rev Oncol Hematol* 2007;**62**:137–47.

Petrie J, Bunn F, Byrne G. Parenting programmes for preventing tobacco, alcohol or drugs misuse in children <18: a systematic review. *Health Educ Res* 2007;**22**:177–91.

Pittler MH, Karagulle MZ, Karagulle M, Ernst E. Spa therapy and balneotherapy for treating low back pain: meta-analysis of randomized trials. *Rheumatology (Oxford)* 2006;**45**:880–4.

Poulsen S, Errboe M, Lescay Mevil Y, Glenny AM. Potassium containing toothpastes for dentine hypersensitivity. *Cochrane Database Syst Rev* 2006;**3**:CD001476.

Rahimi R, Nikfar S, Abdollahi M. Meta-analysis finds use of inhaled corticosteroids during pregnancy safe: a systematic meta-analysis review. *Hum Exp Toxicol* 2006;**25**:447–52.

Ram FS. Use of theophylline in chronic obstructive pulmonary disease: examining the evidence. *Curr Opin Pulm Med* 2006;**12**:132–9.

Reilly KA, Barker KL, Shamley D. A systematic review of lateral wedge orthotics – how useful are they in the management of medial compartment osteoarthritis? *Knee* 2006;**13**:177–83.

Roder V, Mueller DR, Mueser KT, Brenner HD. Integrated psychological therapy (IPT) for schizophrenia: is it effective? *Schizophr Bull* 2006;**32**(Suppl. 1):S81–93.

Rojas MX, Granados C. Oral antibiotics versus parenteral antibiotics for severe pneumonia in children. *Cochrane Database Syst Rev* 2006(2):CD004979.

Ross S, Soroka D, Karahalios A, Glazener CM, Hay-Smith EJ, Drutz HP. Incontinence-specific quality of life measures used in trials of treatments for female urinary incontinence: a systematic review. *Int Urogynecol J Pelvic Floor Dysfunct* 2006;**17**:272–85.

Sjosten N, Kivela SL. The effects of physical exercise on depressive symptoms among the aged: a systematic review. *Int J Geriatr Psychiatry* 2006;**21**:410–8.

Tatrow K, Montgomery GH. Cognitive behavioral therapy techniques for distress and pain in breast cancer patients: a meta-analysis. *J Behav Med* 2006;**29**:17–27.

Temel Y, Kessels A, Tan S, Topdag A, Boon P, Visser-Vandewalle V. Behavioural changes after bilateral subthalamic stimulation in advanced Parkinson disease: a systematic review. *Parkinsonism Relat Disord* 2006;**12**:265–72.

Triantos CK, Goulis J, Patch D, Papatheodoridis GV, Leandro G, Samonakis D, *et al.* An evaluation of emergency sclerotherapy of varices in randomized trials: looking the needle in the eye. *Endoscopy* 2006;**38**:797–807.

Turner A, Rabiu M. Patching for corneal abrasion. *Cochrane Database Syst Rev* 2006(2):CD004764.

Uchida T, Bakhai A, Almonacid A, Shibata T, Cox B, Kuntz RE. A meta-analysis of randomized controlled trials of intracoronary gamma- and beta-radiation therapy for in-stent restenosis. *Heart Vessels* 2006;**21**:368–74.

van Nooten J, Oh H, Pierce B, Koning FJ, Jadad AR. Spiritual care as eHealth: a systematic review. *J Pastoral Care Counsel* 2006;**60**:387–94.

Van Vliet HA, Grimes DA, Helmerhorst FM, Schulz KF. Biphasic versus triphasic oral contraceptives for contraception. *Cochrane Database Syst Rev* 2006;**3**:CD003283.

Vuillermin P, South M, Robertson C. Parent-initiated oral corticosteroid therapy for intermittent wheezing illnesses in children. *Cochrane Database Syst Rev* 2006;**3**:CD005311.

Weeks A, Alfirevic Z. Oral misoprostol administration for labor induction. *Clin Obstet Gynecol* 2006;**49**:658–71.

Whelan AM, Jurgens TM, Bowles SK. Natural health products in the prevention and treatment of osteoporosis: systematic review of randomized controlled trials. *Ann Pharmacother* 2006;**40**:836–49.

Wilson AD, Childs S. Effects of interventions aimed at changing the length of primary care physicians' consultation. *Cochrane Database Syst Rev* 2006(1):CD003540.

Wind J, Polle SW, Fung Kon Jin PH, Dejong CH, von Meyenfeldt MF, Ubbink DT, *et al.* Systematic review of enhanced recovery programmes in colonic surgery. *Br J Surg* 2006;**93**:800–9.

Yan BM, Myers RP. Neurolytic celiac plexus block for pain control in unresectable pancreatic cancer. *Am J Gastroenterol* 2007;**102**:430–8.

Zhang J, Ding EL, Song Y. Adverse effects of cyclooxygenase 2 inhibitors on renal and arrhythmia

events: meta-analysis of randomized trials. *JAMA* 2006;**296**:1619–32.

Risk factor reviews (*n* = 100)

Aalto TJ, Malmivaara A, Kovacs F, Herno A, Alen M, Salmi L, *et al.* Preoperative predictors for postoperative clinical outcome in lumbar spinal stenosis: systematic review. *Spine* 2006;**31**:E648–63.

Abhinav K, Al-Chalabi A, Hortobagyi T, Leigh PN. Electrical injury and amyotrophic lateral sclerosis: a systematic review of the literature. *J Neurol Neurosurg Psychiatry* 2007;**78**:450–3.

Akobeng AK, Ramanan AV, Buchan I, Heller RF. Effect of breast feeding on risk of coeliac disease: a systematic review and meta-analysis of observational studies. *Arch Dis Child* 2006;**91**:39–43.

Alder N, Fenty J, Warren F, Sutton AJ, Rushton L, Jones DR, *et al.* Meta-analysis of mortality and cancer incidence among workers in the synthetic rubber-producing industry. *Am J Epidemiol* 2006;**164**:405–20.

Alexander DD, Mink PJ, Mandel JH, Kelsh MA. A meta-analysis of occupational trichloroethylene exposure and multiple myeloma or leukaemia. *Occup Med (Lond)* 2006;**56**:485–93.

Altman MR, Lydon-Rochelle MT. Prolonged second stage of labor and risk of adverse maternal and perinatal outcomes: a systematic review. *Birth* 2006;**33**:315–22.

Ambroise D, Wild P, Moulin JJ. Update of a meta-analysis on lung cancer and welding. *Scand J Work Environ Health* 2006;**32**:22–31.

Anderson MA, Levsen J, Dusio ME, Bryant PJ, Brown SM, Burr CM, *et al.* Evidence-based factors in readmission of patients with heart failure. *J Nurs Care Qual* 2006;**21**:160–7.

Asia PCSC. Coronary risk prediction for those with and without diabetes. *Eur J Cardiovasc Prev Rehabil* 2006;**13**:30–6.

Atsma F, Bartelink ML, Grobbee DE, van der Schouw YT. Postmenopausal status and early menopause as independent risk factors for cardiovascular disease: a meta-analysis. *Menopause* 2006;**13**:265–79.

Azarpazhooh A, Leake JL. Systematic review of the association between respiratory diseases and oral health. *J Periodontol* 2006;**77**:1465–82.

Batty GD, Deary IJ, Gottfredson LS. Premorbid (early life) IQ and later mortality risk: systematic review. *Ann Epidemiol* 2007;**17**:278–88.

Beam JW, Buckley B. Community-acquired methicillin-resistant *Staphylococcus aureus*: prevalence and risk factors. *J Athl Train* 2006;**41**:337–40.

Brimble KS, Walker M, Margetts PJ, Kundhal KK, Rabbat CG. Meta-analysis: peritoneal membrane transport, mortality, and technique failure in peritoneal dialysis. *J Am Soc Nephrol* 2006;**17**:2591–8.

Butterworth AS, Higgins JP, Pharoah P. Relative and absolute risk of colorectal cancer for individuals with a family history: a meta-analysis. *Eur J Cancer* 2006;**42**:216–27.

Carter OB. The weighty issue of Australian television food advertising and childhood obesity. *Health Promot J Austr* 2006;**17**:5–11.

Cheng JY, Ng EM, Ko JS, Chen RY. Physical activity and erectile dysfunction: meta-analysis of population-based studies. *Int J Impot Res* 2007;**19**:245–52.

Cornish J, Tan E, Teare J, Teoh TG, Rai R, Clark SK, *et al.* A meta-analysis on the influence of inflammatory bowel disease on pregnancy. *Gut* 2007;**56**:830–7.

Curtis KM, Mohllajee AP, Martins SL, Peterson HB. Combined oral contraceptive use among women with hypertension: a systematic review. *Contraception* 2006;**73**:179–88.

da Silva Dal Pizzol T, Knop FP, Mengue SS. Prenatal exposure to misoprostol and congenital anomalies: systematic review and meta-analysis. *Reprod Toxicol* 2006;**22**:666–71.

Dauchet L, Amouyel P, Hercberg S, Dallongeville J. Fruit and vegetable consumption and risk of coronary heart disease: a meta-analysis of cohort studies. *J Nutr* 2006;**136**:2588–93.

den Boer JJ, Oostendorp RA, Beems T, Munneke M, Oerlemans M, Evers AW. A systematic review of bio-psychosocial risk factors for an unfavourable outcome after lumbar disc surgery. *Eur Spine J* 2006;**15**:527–36.

Ding EL, Song Y, Malik VS, Liu S. Sex differences of endogenous sex hormones and risk of type 2 diabetes: a systematic review and meta-analysis. *JAMA* 2006;**295**:1288–99.

Dionne CE, Dunn KM, Croft PR. Does back pain prevalence really decrease with increasing age? A systematic review. *Age Ageing* 2006;**35**:229–34.

Fahmy NM, Mahmud S, Aprikian AG. Delay in the surgical treatment of bladder cancer and survival: systematic review of the literature. *Eur Urol* 2006;**50**:1176–82.

Falagas ME, Kopterides P. Risk factors for the isolation of multi-drug-resistant *Acinetobacter baumannii* and *Pseudomonas aeruginosa*: a systematic review of the literature. *J Hosp Infect* 2006;**64**:7–15.

Falleti MG, Maruff P, Burman P, Harris A. The effects of growth hormone (GH) deficiency and GH replacement on cognitive performance in adults: a meta-analysis of the current literature. *Psychoneuroendocrinology* 2006;**31**:681–91.

Flores-Mateo G, Navas-Acien A, Pastor-Barriuso R, Guallar E. Selenium and coronary heart disease: a meta-analysis. *Am J Clin Nutr* 2006;**84**:762–73.

Franks HM, Roesch SC. Appraisals and coping in people living with cancer: a meta-analysis. *Psychooncology* 2006;**15**:1027–37.

Furber AS, Maheswaran R, Newell JN, Carroll C. Is smoking tobacco an independent risk factor for HIV infection and progression to AIDS? A systemic review. *Sex Transm Infect* 2007;**83**:41–6.

Galassi A, Reynolds K, He J. Metabolic syndrome and risk of cardiovascular disease: a meta-analysis. *Am J Med* 2006;**119**:812–9.

Gheeraert PJ, De Buyzere ML, Taeymans YM, Gillebert TC, Henriques JP, De Backer G, *et al.* Risk factors for primary ventricular fibrillation during acute myocardial infarction: a systematic review and meta-analysis. *Eur Heart J* 2006;**27**:2499–510.

Gomes B, Higginson IJ. Factors influencing death at home in terminally ill patients with cancer: systematic review. *BMJ* 2006;**332**:515–21.

Hay JL, McCaul KD, Magnan RE. Does worry about breast cancer predict screening behaviors? A meta-analysis of the prospective evidence. *Prev Med* 2006;**42**:401–8.

He FJ, Nowson CA, MacGregor GA. Fruit and vegetable consumption and stroke: meta-analysis of cohort studies. *Lancet* 2006;**367**:320–6.

Hernandez-Diaz S, Varas-Lorenzo C, Garcia Rodriguez LA. Non-steroidal antiinflammatory drugs and the risk of acute myocardial infarction. *Basic Clin Pharmacol Toxicol* 2006;**98**:266–74.

Hobbs CG, Sterne JA, Bailey M, Heyderman RS, Birchall MA, Thomas SJ. Human papillomavirus and head and neck cancer: a systematic review and meta-analysis. *Clin Otolaryngol* 2006;**31**:259–66.

Holscher T, Bentzen SM, Baumann M. Influence of connective tissue diseases on the expression of radiation side effects: a systematic review. *Radiother Oncol* 2006;**78**:123–30.

Huang JS, Lee TA, Lu MC. Prenatal programming of childhood overweight and obesity. *Matern Child Health J* 2007;**11**:461–73.

Hughes JR, Carpenter MJ. Does smoking reduction increase future cessation and decrease disease risk? A qualitative review. *Nicotine Tob Res* 2006;**8**:739–49.

Jackson CA, Sudlow CL. Is hypertension a more frequent risk factor for deep than for lobar supratentorial intracerebral haemorrhage? *J Neurol Neurosurg Psychiatry* 2006;**77**:1244–52.

Jewett M, Rendon R, Dranitsaris G, Drachenberg D, Tanguay S, Donnelly B, *et al.* Does prolonging the time to renal cancer surgery affect long-term cancer control: a systematic review of the literature. *Can J Urol* 2006;**13**(Suppl. 3):54–61.

Kamphuis CB, Giskes K, de Bruijn GJ, Wendel-Vos W, Brug J, van Lenthe FJ. Environmental determinants of fruit and vegetable consumption among adults: a systematic review. *Br J Nutr* 2006;**96**:620–35.

Kantovitz KR, Pascon FM, Rontani RM, Gaviao MB. Obesity and dental caries – A systematic review. *Oral Health Prev Dent* 2006;**4**:137–44.

Kasper JS, Giovannucci E. A meta-analysis of diabetes mellitus and the risk of prostate cancer. *Cancer Epidemiol Biomarkers Prev* 2006;**15**:2056–62.

Khambalia A, Joshi P, Brussoni M, Raina P, Morrongiello B, Macarthur C. Risk factors for unintentional injuries due to falls in children aged 0–6 years: a systematic review. *Inj Prev* 2006;**12**:378–81.

Knol MJ, Twisk JW, Beekman AT, Heine RJ, Snoek FJ, Pouwer F. Depression as a risk factor for the onset of type 2 diabetes mellitus. A meta-analysis. *Diabetologia* 2006;**49**:837–45.

Koppes LL, Dekker JM, Hendriks HF, Bouter LM, Heine RJ. Meta-analysis of the relationship between alcohol consumption and coronary heart disease and mortality in type 2 diabetic patients. *Diabetologia* 2006;**49**:648–52.

Kubo A, Corley DA. Body mass index and adenocarcinomas of the esophagus or gastric cardia: a systematic review and meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2006;**15**:872–8.

Langan SM, Williams HC. What causes worsening of eczema? A systematic review. *Br J Dermatol* 2006;**155**:504–14.

Larsson SC, Orsini N, Wolk A. Milk, milk products and lactose intake and ovarian cancer risk: a meta-analysis of epidemiological studies. *Int J Cancer* 2006;**118**:431–41.

Leandro G, Mangia A, Hui J, Fabris P, Rubbia-Brandt L, Colloredo G, *et al.* Relationship between steatosis, inflammation, and fibrosis in chronic hepatitis C: a meta-analysis of individual patient data. *Gastroenterology* 2006;**130**:1636–42.

Lee PN, Forey BA. Environmental tobacco smoke exposure and risk of stroke in nonsmokers: A review with meta-analysis. *J Stroke Cerebrovasc Dis* 2006;**15**:190–201.

Leeners B, Richter-Appelt H, Imthurn B, Rath W. Influence of childhood sexual abuse on pregnancy, delivery, and the early postpartum period in adult women. *J Psychosom Res* 2006;**61**:139–51.

Littlejohn C. Does socio-economic status influence the acceptability of, attendance for, and outcome of, screening and brief interventions for alcohol misuse: a review. *Alcohol Alcohol* 2006;**41**:540–5.

Ma H, Bernstein L, Pike MC, Ursin G. Reproductive factors and breast cancer risk according to joint estrogen and progesterone receptor status: a meta-analysis of epidemiological studies. *Breast Cancer Res* 2006;**8**:R43.

Maguire S, Mann M, John N, Ellaway B, Sibert JR, Kemp AM. Does cardiopulmonary resuscitation cause rib fractures in children? A systematic review. *Child Abuse Negl* 2006;**30**:739–51.

Mahon NE, Yarcheski A, Yarcheski TJ, Cannella BL, Hanks MM. A meta-analytic study of predictors for loneliness during adolescence. *Nurs Res* 2006;**55**:308–15.

Maki DG, Kluger DM, Crnich CJ. The risk of bloodstream infection in adults with different intravascular devices: a systematic review of 200 published prospective studies. *Mayo Clin Proc* 2006;**81**:1159–71.

Mehra R, Moore BA, Crothers K, Tetrault J, Fiellin DA. The association between marijuana smoking and lung cancer: a systematic review. *Arch Intern Med* 2006;**166**:1359–67.

Mills EJ, Nachega JB, Bangsberg DR, Singh S, Rachlis B, Wu P, *et al.* Adherence to HAART: a systematic review of developed and developing nation patient-reported barriers and facilitators. *PLoS Med* 2006;**3**:e438.

Mills EJ, Seely D, Rachlis B, Griffith L, Wu P, Wilson K, *et al.* Barriers to participation in clinical trials of cancer: a meta-analysis and systematic review of patient-reported factors. *Lancet Oncol* 2006;**7**:141–8.

Mizoue T, Tanaka K, Tsuji I, Wakai K, Nagata C, Otani T, *et al.* Alcohol drinking and colorectal cancer risk: an evaluation based on a systematic review of epidemiologic evidence among the Japanese population. *Jpn J Clin Oncol* 2006;**36**:582–97.

Murphy VE, Clifton VL, Gibson PG. Asthma exacerbations during pregnancy: incidence and association with adverse pregnancy outcomes. *Thorax* 2006;**61**:169–76.

Naito M, Yuasa H, Nomura Y, Nakayama T, Hamajima N, Hanada N. Oral health status and health-related quality of life: a systematic review. *J Oral Sci* 2006;**48**:1–7.

Nakagami T, Qiao Q, Tuomilehto J, Balkau B, Tajima N, Hu G, *et al.* Screen-detected diabetes, hypertension and hypercholesterolemia as predictors of cardiovascular mortality in five populations of Asian origin: the DECODA study. *Eur J Cardiovasc Prev Rehabil* 2006;**13**:555–61.

Neri M, Ugolini D, Bonassi S, Fucic A, Holland N, Knudsen LE, *et al.* Children's exposure to environmental pollutants and biomarkers of genetic damage. II. Results of a comprehensive literature search and meta-analysis. *Mutat Res* 2006;**612**:14–39.

Newman LA, Griffith KA, Jatoi I, Simon MS, Crowe JP, Colditz GA. Meta-analysis of survival in African American and white American patients with breast cancer: ethnicity compared with socioeconomic status. *J Clin Oncol* 2006;**24**:1342–9.

Omarova A, Phillips CJ. A meta-analysis of literature data relating to the relationships between cadmium intake and toxicity indicators in humans. *Environ Res* 2007;**103**:432–40.

Onega T, Baron J, MacKenzie T. Cancer after total joint arthroplasty: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2006;**15**:1532–7.

Ong KK, Loos RJ. Rapid infancy weight gain and subsequent obesity: systematic reviews and hopeful suggestions. *Acta Paediatr* 2006;**95**:904–8.

Papatheodoridis GV, Sougioultzis S, Archimandritis AJ. Effects of *Helicobacter pylori* and nonsteroidal anti-inflammatory drugs on peptic ulcer disease: a systematic review. *Clin Gastroenterol Hepatol* 2006;**4**:130–42.

Paradies Y. A systematic review of empirical research on self-reported racism and health. *Int J Epidemiol* 2006;**35**:888–901.

Paydarfar JA, Birkmeyer NJ. Complications in head and neck surgery: a meta-analysis of postlaryngectomy pharyngocutaneous fistula. *Arch Otolaryngol Head Neck Surg* 2006;**132**:67–72.

Reis FJ, Sousa TA, Oliveira MS, Dantas N, Silveira M, Braghiroly MI, *et al.* Is hepatitis C virus a cause of idiopathic dilated cardiomyopathy?: A systematic review of literature. *Braz J Infect Dis* 2006;**10**:199–202.

Rhodes RE, Smith NE. Personality correlates of physical activity: a review and meta-analysis. *Br J Sports Med* 2006;**40**:958–65.

Robertson L, Wu O, Langhorne P, Twaddle S, Clark P, Lowe GD, *et al.* Thrombophilia in pregnancy: a systematic review. *Br J Haematol* 2006;**132**:171–96.

Rocha AT, de Vasconcellos AG, da Luz Neto ER, Araujo DM, Alves ES, Lopes AA. Risk of venous thromboembolism and efficacy of thromboprophylaxis in hospitalized obese medical patients and in obese patients undergoing bariatric surgery. *Obes Surg* 2006;**16**:1645–55.

Rodrigues MC, Mello RR, Fonseca SC. Learning difficulties in schoolchildren born with very low birth weight. *J Pediatr (Rio J)* 2006;**82**:6–14.

Simpson SH, Eurich DT, Majumdar SR, Padwal RS, Tsuyuki RT, Varney J, *et al.* A meta-analysis of the association between adherence to drug therapy and mortality. *BMJ* 2006;**333**:15.

Singh G, Wu O, Langhorne P, Madhok R. Risk of acute myocardial infarction with nonselective non-steroidal anti-inflammatory drugs: a meta-analysis. *Arthritis Res Ther* 2006;**8**:R153.

Smeets RJ, Wade D, Hidding A, Van Leeuwen PJ, Vlaeyen JW, Knottnerus JA. The association of physical deconditioning and chronic low back pain: a hypothesis-oriented systematic review. *Disabil Rehabil* 2006;**28**:673–93.

Szilagyi A, Nathwani U, Vinokuroff C, Correa JA, Shrier I. The effect of lactose maldigestion on the relationship between dairy food intake and colorectal cancer: a systematic review. *Nutr Cancer* 2006;**55**:141–50.

Thornton J, Kelly SP, Harrison RA, Edwards R. Cigarette smoking and thyroid eye disease: a systematic review. *Eye* 2007;**21**:1135–45.

Tinazzi E, Ficarra V, Simeoni S, Artibani W, Lunardi C. Reactive arthritis following BCG immunotherapy for urinary bladder carcinoma: a systematic review. *Rheumatol Int* 2006;**26**:481–8.

Tokumaru O, Haruki K, Bacal K, Katagiri T, Yamamoto T, Sakurai Y. Incidence of cancer among female flight attendants: a meta-analysis. *J Travel Med* 2006;**13**:127–32.

Truong KD, Ma S. A systematic review of relations between neighborhoods and mental health. *J Ment Health Policy Econ* 2006;**9**:137–54.

Vamvakas EC. Pneumonia as a complication of blood product transfusion in the critically ill: transfusion-related immunomodulation (TRIM). *Crit Care Med* 2006;**34**(5 Suppl.):S151–9.

van der Horst K, Oenema A, Ferreira I, Wendel-Vos W, Giskes K, van Lenthe F, *et al.* A systematic review of environmental correlates of obesity-related dietary behaviors in youth. *Health Educ Res* 2007;**22**:203–26.

Van Maele-Fabry G, Libotte V, Willems J, Lison D. Review and meta-analysis of risk estimates for prostate cancer in pesticide manufacturing workers. *Cancer Causes Control* 2006;**17**:353–73.

van Velzen JM, van Bennekom CA, Polomski W, Slootman JR, van der Woude LH, Houdijk H. Physical capacity and walking ability after lower limb amputation: a systematic review. *Clin Rehabil* 2006;**20**:999–1016.

Wakai K, Inoue M, Mizoue T, Tanaka K, Tsuji I, Nagata C, *et al.* Tobacco smoking and lung cancer risk: an evaluation based on a systematic review of epidemiological evidence among the Japanese population. *Jpn J Clin Oncol* 2006;**36**:309–24.

Warner L, Stone KM, Macaluso M, Buehler JW, Austin HD. Condom use and risk of gonorrhea and Chlamydia: a systematic review of design and measurement factors assessed in epidemiologic studies. *Sex Transm Dis* 2006;**33**(1):36–51.

Whalley GA, Gamble GD, Doughty RN. The prognostic significance of restrictive diastolic filling associated with heart failure: a meta-analysis. *Int J Cardiol* 2007;**116**:70–7.

Wind J, Lagarde SM, Ten Kate FJ, Ubbink DT, Bemelman WA, van Lanschot JJ. A systematic review on the significance of extracapsular lymph node involvement in gastrointestinal malignancies. *Eur J Surg Oncol* 2007;**33**:401–8.

Wohl M, Gorwood P. Paternal ages below or above 35 years old are associated with a different risk of schizophrenia in the offspring. *Eur Psychiatry* 2007;**22**:22–6.

Zhang B, Wing YK. Sex differences in insomnia: a meta-analysis. *Sleep* 2006;**29**:85–93.

Zhang XF, Attia J, D'Este C, Ma XY. The relationship between higher blood pressure and ischaemic, haemorrhagic stroke among Chinese and Caucasians: meta-analysis. *Eur J Cardiovasc Prev Rehabil* 2006;**13**:429–37.

Zhao Y, Wang S, Aunan K, Seip HM, Hao J. Air pollution and lung cancer risks in China – a meta-analysis. *Sci Total Environ* 2006;**366**:500–13.

Zumkeller N, Brenner H, Zwahlen M, Rothenbacher D. *Helicobacter pylori* infection and colorectal cancer risk: a meta-analysis. *Helicobacter* 2006;**11**:75–80.

Diagnostic reviews (*n* = 50)

Alongi F, Ragusa P, Montemaggi P, Bona CM. Combining independent studies of diagnostic fluorodeoxyglucose positron-emission tomography and computed tomography in mediastinal lymph node staging for non-small cell lung cancer. *Tumori* 2006;**92**:327–33.

Bandera E, Botteri M, Minelli C, Sutton A, Abrams KR, Latronico N. Cerebral blood flow threshold of ischemic

penumbra and infarct core in acute ischemic stroke: a systematic review. *Stroke* 2006;**37**:1334–9.

Beattie WS, Abdelnaem E, Wijeysundera DN, Buckley DN. A meta-analytic comparison of preoperative stress echocardiography and nuclear scintigraphy imaging. *Anesth Analg* 2006;**102**:8–16.

Benatar M. A systematic review of diagnostic studies in myasthenia gravis. *Neuromuscul Disord* 2006;**16**:459–67.

Benjaminse A, Gokeler A, van der Schans CP. Clinical diagnosis of an anterior cruciate ligament rupture: a meta-analysis. *J Orthop Sports Phys Ther* 2006;**36**:267–88.

Biagini E, Shaw LJ, Poldermans D, Schinkel AF, Rizzello V, Elhendy A, *et al.* Accuracy of non-invasive techniques for diagnosis of coronary artery disease and prediction of cardiac events in patients with left bundle branch block: a meta-analysis. *Eur J Nucl Med Mol Imaging* 2006;**33**:1442–51.

Brealey S, Scally A, Hahn S, Thomas N, Godfrey C, Crane S. Accuracy of radiographers red dot or triage of accident and emergency radiographs in clinical practice: a systematic review. *Clin Radiol* 2006;**61**:604–15.

Carnes D, Ashby D, Underwood M. A systematic review of pain drawing literature: should pain drawings be used for psychologic screening? *Clin J Pain* 2006;**22**:449–57.

Chappuis F, Rijal S, Soto A, Menten J, Boelaert M. A meta-analysis of the diagnostic performance of the direct agglutination test and rK39 dipstick for visceral leishmaniasis. *BMJ* 2006;**333**:723.

Christou MA, Siontis GC, Katritsis DG, Ioannidis JP. Meta-analysis of fractional flow reserve versus quantitative coronary angiography and noninvasive imaging for evaluation of myocardial ischemia. *Am J Cardiol* 2007;**99**:450–6.

Davenport C, Cheng EY, Kwok YT, Lai AH, Wakabayashi T, Hyde C, *et al.* Assessing the diagnostic test accuracy of natriuretic peptides and ECG in the diagnosis of left ventricular systolic dysfunction: a systematic review and meta-analysis. *Br J Gen Pract* 2006;**56**:48–56.

de Graaf I, Prak A, Bierma-Zeinstra S, Thomas S, Peul W, Koes B. Diagnosis of lumbar spinal stenosis: a systematic review of the accuracy of diagnostic tests. *Spine* 2006;**31**:1168–76.

Detsky ME, McDonald DR, Baerlocher MO, Tomlinson GA, McCrory DC, Booth CM. Does this patient with headache have a migraine or need neuroimaging? *JAMA* 2006;**296**:1274–83.

Doria AS, Moineddin R, Kellenberger CJ, Epelman M, Beyene J, Schuh S, *et al.* US or CT for diagnosis of appendicitis in children and adults? A meta-analysis. *Radiology* 2006;**241**:83–94.

Gisbert JP, Abraira V. Accuracy of *Helicobacter pylori* diagnostic tests in patients with bleeding peptic ulcer: a systematic review and meta-analysis. *Am J Gastroenterol* 2006;**101**:848–63.

Gisbert JP, de la Morena F, Abraira V. Accuracy of monoclonal stool antigen test for the diagnosis of *H. pylori* infection: a systematic review and meta-analysis. *Am J Gastroenterol* 2006;**101**:1921–30.

Gupta SG, Wang LC, Penas PF, Gellenthin M, Lee SJ, Nghiem P. Sentinel lymph node biopsy for evaluation and treatment of patients with Merkel cell carcinoma: The Dana-Farber experience and meta-analysis of the literature. *Arch Dermatol* 2006;**142**:685–90.

Hogg K, Brown G, Dunning J, Wright J, Carley S, Foex B, *et al.* Diagnosis of pulmonary embolism with CT pulmonary angiography: a systematic review. *Emerg Med J* 2006;**23**:172–8.

Hollingworth W, Medina LS, Lenkinski RE, Shibata DK, Bernal B, Zurakowski D, *et al.* A systematic literature review of magnetic resonance spectroscopy for the characterization of brain tumors. *AJNR Am J Neuroradiol* 2006;**27**:1404–11.

Jones AE, Fiechtl JF, Brown MD, Ballew JJ, Kline JA. Procalcitonin test in the diagnosis of bacteremia: a meta-analysis. *Ann Emerg Med* 2007;**50**:34–41.

Karassa FB, Afeltra A, Ambrozic A, Chang DM, De Keyser F, Doria A, *et al.* Accuracy of anti-ribosomal P protein antibody testing for the diagnosis of neuropsychiatric systemic lupus erythematosus: an international meta-analysis. *Arthritis Rheum* 2006;**54**:312–24.

King J, Thatcher N, Pickering C, Hasleton P. Sensitivity and specificity of immunohistochemical antibodies used to distinguish between benign and malignant pleural disease: a systematic review of published reports. *Histopathology* 2006;**49**:561–8.

Koliopoulos G, Arbyn M, Martin-Hirsch P, Kyrgiou M, Prendiville W, Paraskevaidis E. Diagnostic accuracy of human papillomavirus testing in primary cervical screening: a systematic review and meta-analysis of non-randomized studies. *Gynecol Oncol* 2007;**104**:232–46.

Locker T, Goodacre S, Sampson F, Webster A, Sutton AJ. Meta-analysis of plethysmography and rheography in the diagnosis of deep vein thrombosis. *Emerg Med J* 2006;**23**:630–5.

Lou L, Lagravere MO, Compton S, Major PW, Flores-Mir C. Accuracy of measurements and reliability of landmark identification with computed tomography (CT) techniques in the maxillofacial area: a systematic review. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2007;**104**:402–11.

Major MP, Flores-Mir C, Major PW. Assessment of lateral cephalometric diagnosis of adenoid hypertrophy and posterior upper airway obstruction: a systematic review. *Am J Orthod Dentofacial Orthop* 2006;**130**:700–8.

Martin A, Portaels F, Palomino JC. Colorimetric redox-indicator methods for the rapid detection of multidrug resistance in *Mycobacterium tuberculosis*: a systematic review and meta-analysis. *J Antimicrob Chemother* 2007;**59**:175–83.

Martin JL, Williams KS, Sutton AJ, Abrams KR, Assassa RP. Systematic review and meta-analysis of methods of diagnostic assessment for urinary incontinence. *Neurourol Urodyn* 2006;**25**:674–83; discussion 684.

Myers ER, Bastian LA, Havrilesky LJ, Kulasingam SL, Terplan MS, Cline KE, *et al.* Management of adnexal mass. *Evid Rep Technol Assess (Full Rep)* 2006(130):1–145.

Nijkeuter M, Ginsberg JS, Huisman MV. Diagnosis of deep vein thrombosis and pulmonary embolism in pregnancy: a systematic review. *J Thromb Haemost* 2006;**4**:496–500.

Peacock F, Morris DL, Anwaruddin S, Christenson RH, Collinson PO, Goodacre SW, *et al.* Meta-analysis of ischemia-modified albumin to rule out acute coronary syndromes in the emergency department. *Am Heart J* 2006;**152**:253–62.

Purkayastha S, Chow A, Athanasiou T, Cambaroudis A, Panesar S, Kinross J, *et al.* Does serum procalcitonin have a role in evaluating the severity of acute pancreatitis? A question revisited. *World J Surg* 2006;**30**:1713–21.

Riquelme A, Calvo M, Salech F, Valderrama S, Pattillo A, Arellano M, *et al.* Value of adenosine deaminase (ADA) in ascitic fluid for the diagnosis of tuberculous peritonitis: a meta-analysis. *J Clin Gastroenterol* 2006;**40**:705–10.

Rodgers M, Nixon J, Hempel S, Aho T, Kelly J, Neal D, *et al.* Diagnostic tests and algorithms used in the investigation of haematuria: systematic reviews and economic evaluation. *Health Technol Assess* 2006;**10**:iii-iv, xi-259.

Rubinstein SM, Pool JJ, van Tulder MW, Riphagen, II, de Vet HC. A systematic review of the diagnostic accuracy of provocative tests of the neck for diagnosing cervical radiculopathy. *Eur Spine J* 2007;**16**:307–19.

Sgouros SN, Pereira SP. Systematic review: sphincter of Oddi dysfunction – non-invasive diagnostic methods and long-term outcome after endoscopic sphincterotomy. *Aliment Pharmacol Ther* 2006;**24**:237–46.

Stein PD, Beemath A, Kayali F, Skaf E, Sanchez J, Olson RE. Multidetector computed tomography for the diagnosis of coronary artery disease: a systematic review. *Am J Med* 2006;**119**:203–16.

Steingart KR, Henry M, Ng V, Hopewell PC, Ramsay A, Cunningham J, *et al.* Fluorescence versus conventional sputum smear microscopy for tuberculosis: a systematic review. *Lancet Infect Dis* 2006;**6**:570–81.

Steingart KR, Ng V, Henry M, Hopewell PC, Ramsay A, Cunningham J, *et al.* Sputum processing methods to improve the sensitivity of smear microscopy for tuberculosis: a systematic review. *Lancet Infect Dis* 2006;**6**:664–74.

Sun Z, Jiang W. Diagnostic value of multislice computed tomography angiography in coronary artery disease: a meta-analysis. *Eur J Radiol* 2006;**60**:279–86.

Tomlinson A, Khanal S, Ramaesh K, Diaper C, McFadyen A. Tear film osmolarity: determination of a referent for dry eye diagnosis. *Invest Ophthalmol Vis Sci* 2006;**47**:4309–15.

Tuon FF, Litvoc MN, Lopes MI. Adenosine deaminase and tuberculous pericarditis – a systematic review with meta-analysis. *Acta Trop* 2006;**99**:67–74.

Vakil N, Moayyedi P, Fennerty MB, Talley NJ. Limited value of alarm features in the diagnosis of upper gastrointestinal malignancy: systematic review and meta-analysis. *Gastroenterology* 2006;**131**:390–401; quiz 659–60.

van der Zaag-Loonen HJ, Dikkers R, de Bock GH, Oudkerk M. The clinical value of a negative multi-detector computed tomographic angiography in patients suspected of coronary artery disease: A meta-analysis. *Eur Radiol* 2006;**16**:2748–56.

Verma D, Kapadia A, Eisen GM, Adler DG. EUS vs MRCP for detection of choledocholithiasis. *Gastrointest Endosc* 2006;**64**:248–54.

Wardlaw JM, Chappell FM, Best JJ, Wartolowska K, Berry E. Non-invasive imaging compared with intra-arterial angiography in the diagnosis of symptomatic carotid stenosis: a meta-analysis. *Lancet* 2006;**367**:1503–12.

Whiting P, Harbord R, Main C, Deeks JJ, Filippini G, Egger M, *et al.* Accuracy of magnetic resonance imaging for the diagnosis of multiple sclerosis: systematic review. *BMJ* 2006;**332**:875–84.

Whiting P, Westwood M, Bojke L, Palmer S, Richardson G, Cooper J, *et al.* Clinical effectiveness and cost-effectiveness of tests for the diagnosis and investigation of urinary tract infection in children: a systematic review and economic model. *Health Technol Assess* 2006;**10**:iii-iv, xi–xiii, 1–154.

Will O, Purkayastha S, Chan C, Athanasiou T, Darzi AW, Gedroyc W, *et al.* Diagnostic precision of nanoparticle-enhanced MRI for lymph-node metastases: a meta-analysis. *Lancet Oncol* 2006;**7**:52–60.

Xing Y, Foy M, Cox DD, Kuerer HM, Hunt KK, Cormier JN. Meta-analysis of sentinel lymph node biopsy after preoperative chemotherapy in patients with breast cancer. *Br J Surg* 2006;**93**:539–46.

Genetic reviews (*n* = 50)

Akomolafe A, Lunetta KL, Erlich PM, Cupples LA, Baldwin CT, Huyck M, *et al*. Genetic association between endothelial nitric oxide synthase and Alzheimer disease. *Clin Genet* 2006;**70**:49–56.

Annese V, Valvano MR, Palmieri O, Latiano A, Bossa F, Andriulli A. Multidrug resistance 1 gene in inflammatory bowel disease: a meta-analysis. *World J Gastroenterol* 2006;**12**:3636–44.

Aoki T, Hirota T, Tamari M, Ichikawa K, Takeda K, Arinami T, *et al*. An association between asthma and TNF-308G/A polymorphism: meta-analysis. *J Hum Genet* 2006;**51**:677–85.

Arias A, Feinn R, Kranzler HR. Association of an Asn40Asp (A118G) polymorphism in the mu-opioid receptor gene with substance dependence: a meta-analysis. *Drug Alcohol Depend* 2006;**83**:262–8.

Baglietto L, Jenkins MA, Severi G, Giles GG, Bishop DT, Boyle P, *et al*. Measures of familial aggregation depend on definition of family history: meta-analysis for colorectal cancer. *J Clin Epidemiol* 2006;**59**:114–24.

Barroso I, Luan J, Sandhu MS, Franks PW, Crowley V, Schafer AJ, *et al.*. Meta-analysis of the Gly482Ser variant in PPARGC1A in type 2 diabetes and related phenotypes. *Diabetologia* 2006;**49**:501–5.

Blomqvist ME, Reynolds C, Katzov H, Feuk L, Andreasen N, Bogdanovic N, *et al*. Towards compendia of negative genetic association studies: an example for Alzheimer disease. *Hum Genet* 2006;**119**:29–37.

Boccia S, La Torre G, Gianfagna F, Mannocci A, Ricciardi G. Glutathione S-transferase T1 status and gastric cancer risk: a meta-analysis of the literature. *Mutagenesis* 2006;**21**:115–23.

Borlak J, Reamon-Buettner SM. N-acetyltransferase 2 (NAT2) gene polymorphisms in colon and lung cancer patients. *BMC Med Genet* 2006;**7**:58.

Burwick RM, Ramsay PP, Haines JL, Hauser SL, Oksenberg JR, Pericak-Vance MA, *et al*. APOE epsilon variation in multiple sclerosis susceptibility and disease severity: some answers. *Neurology* 2006;**66**:1373–83.

Camargo MC, Mera R, Correa P, Peek RM Jr, Fontham ET, Goodman KJ, *et al*. Interleukin-1beta and interleukin-1 receptor antagonist gene polymorphisms and gastric cancer: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2006;**15**:1674–87.

Casas JP, Cavalleri GL, Bautista LE, Smeeth L, Humphries SE, Hingorani AD. Endothelial nitric oxide synthase gene polymorphisms and cardiovascular disease: a HuGE review. *Am J Epidemiol* 2006;**164**:921–35.

Daher S, Sass N, Oliveira LG, Mattar R. Cytokine genotyping in preeclampsia. *Am J Reprod Immunol* 2006;**55**:130–5.

Delgado-Vega AM, Anaya JM. Meta-analysis of HLA-DRB1 polymorphism in Latin American patients with rheumatoid arthritis. *Autoimmun Rev* 2007;**6**:402–8.

Diep CB, Kleivi K, Ribeiro FR, Teixeira MR, Lindgjaerde OC, Lothe RA. The order of genetic events associated with colorectal cancer progression inferred from meta-analysis of copy number changes. *Genes Chromosomes Cancer* 2006;**45**:31–41.

Fang Y, Rivadeneira F, van Meurs JB, Pols HA, Ioannidis JP, Uitterlinden AG. Vitamin D receptor gene BsmI and TaqI polymorphisms and fracture risk: a meta-analysis. *Bone* 2006;**39**:938–45.

Fletcher O, Johnson N, Palles C, dos Santos Silva I, McCormack V, Whittaker J, *et al*. Inconsistent association between the STK15 F31I genetic polymorphism and breast cancer risk. *J Natl Cancer Inst* 2006;**98**:1014–8.

Gao J, Shan G, Sun B, Thompson PJ, Gao X. Association between polymorphism of tumour necrosis factor alpha-308 gene promoter and asthma: a meta-analysis. *Thorax* 2006;**61**:466–71.

Healy DG, Abou-Sleiman PM, Casas JP, Ahmadi KR, Lynch T, Gandhi S, *et al*. UCHL-1 is not a Parkinson's disease susceptibility gene. *Ann Neurol* 2006;**59**:627–33.

Huang X, Chen P, Kaufer DI, Troster AI, Poole C. Apolipoprotein E and dementia in Parkinson disease: a meta-analysis. *Arch Neurol* 2006;**63**:189–93.

Huang Y, Han S, Li Y, Mao Y, Xie Y. Different roles of MTHFR C677T and A1298C polymorphisms in colorectal adenoma and colorectal cancer: a meta-analysis. *J Hum Genet* 2007;**52**:73–85.

Jeong SH, Joo EJ, Ahn YM, Lee KY, Kim YS. Investigation of genetic association between human Frizzled homolog 3 gene (FZD3) and schizophrenia: results in a Korean population and evidence from meta-analysis. *Psychiatry Res* 2006;**143**:1–11.

Kendler KS, Baker JH. Genetic influences on measures of the environment: a systematic review. *Psychol Med* 2007;**37**:615–26.

Koschny R, Holland H, Koschny T, Vitzthum HE. Comparative genomic hybridization pattern of non-anaplastic and anaplastic oligodendrogliomas – a meta-analysis. *Pathol Res Pract* 2006;**202**:23–30.

179

Lee SA, Lee KM, Park SK, Choi JY, Kim B, Nam J, *et al.* Genetic polymorphism of XRCC3 Thr241Met and breast cancer risk: case-control study in Korean women and meta-analysis of 12 studies. *Breast Cancer Res Treat* 2007;**103**:71–6.

Lee YH, Rho YH, Choi SJ, Ji JD, Song GG, Nath SK, *et al.* The PTPN22 C1858T functional polymorphism and autoimmune diseases – a meta-analysis. *Rheumatology (Oxford)* 2007;**46**:49–56.

Lewis SJ, Lawlor DA, Davey Smith G, Araya R, Timpson N, Day IN, *et al.* The thermolabile variant of MTHFR is associated with depression in the British Women's Heart and Health Study and a meta-analysis. *Mol Psychiatry* 2006;**11**:352–60.

Li D, He L. Association study of the G-protein signaling 4 (RGS4) and proline dehydrogenase (PRODH) genes with schizophrenia: a meta-analysis. *Eur J Hum Genet* 2006;**14**:1130–5.

Li H, Tai BC. RNASEL gene polymorphisms and the risk of prostate cancer: a meta-analysis. *Clin Cancer Res* 2006;**12**:5713–9.

Marti A, Ochoa MC, Sanchez-Villegas A, Martinez JA, Martinez-Gonzalez MA, Hebebrand J, *et al.* Meta-analysis on the effect of the N363S polymorphism of the glucocorticoid receptor gene (GRL) on human obesity. *BMC Med Genet* 2006;**7**:50.

Medica I, Kastrin A, Peterlin B. Genetic polymorphisms in vasoactive genes and preeclampsia: a meta-analysis. *Eur J Obstet Gynecol Reprod Biol* 2007;**131**:115–26.

Nishimura F, Shibasaki M, Ichikawa K, Arinami T, Noguchi E. Failure to find an association between CD14–159C/T polymorphism and asthma: a family-based association test and meta-analysis. *Allergol Int* 2006;**55**:55–8.

Noso S, Ikegami H, Fujisawa T, Kawabata Y, Asano K, Hiromine Y, *et al.* Association of SUMO4, as a candidate gene for IDDM5, with susceptibility to type 1 diabetes in Asian populations. *Ann N Y Acad Sci* 2006;**1079**:41–6.

Philibert RA. A meta-analysis of the association of the HOPA12bp polymorphism and schizophrenia. *Psychiatr Genet* 2006;**16**:73–6.

Qian L, Zhao J, Shi Y, Zhao X, Feng G, Xu F, *et al.* Brain-derived neurotrophic factor and risk of schizophrenia: an association study and meta-analysis. *Biochem Biophys Res Commun* 2007;**353**:738–43.

Serretti A, Kato M, De Ronchi D, Kinoshita T. Meta-analysis of serotonin transporter gene promoter polymorphism (5-HTTLPR) association with selective serotonin reuptake inhibitor efficacy in depressed patients. *Mol Psychiatry* 2007;**12**:247–57.

Shirts BH, Wood J, Yolken RH, Nimgaonkar VL. Association study of IL10, IL1beta, and IL1RN and schizophrenia using tag SNPs from a comprehensive database: suggestive association with rs16944 at IL1beta. *Schizophr Res* 2006;**88**:235–44.

Talkowski ME, Seltman H, Bassett AS, Brzustowicz LM, Chen X, Chowdari KV, *et al.* Evaluation of a susceptibility gene for schizophrenia: genotype based meta-analysis of RGS4 polymorphisms from thirteen independent samples. *Biol Psychiatry* 2006;**60**:152–62.

Tello-Ruiz MK, Curley C, DelMonte T, Giallourakis C, Kirby A, Miller K, *et al.* Haplotype-based association analysis of 56 functional candidate genes in the IBD6 locus on chromosome 19. *Eur J Hum Genet* 2006;**14**:780–90.

Tenesa A, Campbell H, Barnetson R, Porteous M, Dunlop M, Farrington SM. Association of MUTYH and colorectal cancer. *Br J Cancer* 2006;**95**:239–42.

Thakkinstian A, Bowe S, McEvoy M, Smith W, Attia J. Association between apolipoprotein E polymorphisms and age-related macular degeneration: A HuGE review and meta-analysis. *Am J Epidemiol* 2006;**164**:813–22.

Tonjes A, Scholz M, Loeffler M, Stumvoll M. Association of Pro12Ala polymorphism in peroxisome proliferator-activated receptor gamma with Pre-diabetic phenotypes: meta-analysis of 57 studies on nondiabetic individuals. *Diabetes Care* 2006;**29**:2489–97.

Tripathy CB, Roy N. Meta-analysis of glutathione S-transferase M1 genotype and risk toward head and neck cancer. *Head Neck* 2006;**28**:217–24.

Tsantes AE, Nikolopoulos GK, Bagos PG, Vaiopoulos G, Travlou A. Lack of association between the platelet glycoprotein Ia C807T gene polymorphism and coronary artery disease: a meta-analysis. *Int J Cardiol* 2007;**118**:189–96.

Vollmert C, Hahn S, Lamina C, Huth C, Kolz M, Schopfer-Wendels A, *et al.* Calpain-10 variants and haplotypes are associated with polycystic ovary syndrome in Caucasians. *Am J Physiol Endocrinol Metab* 2007;**292**:E836–44.

Wang Z, Wei J, Zhang X, Guo Y, Xu Q, Liu S, *et al.* A review and re-evaluation of an association between the NOTCH4 locus and schizophrenia. *Am J Med Genet B Neuropsychiatr Genet* 2006;**141**:902–6.

Webb EL, Rudd MF, Houlston RS. Case-control, kin-cohort and meta-analyses provide no support for STK15 F31I as a low penetrance colorectal cancer allele. *Br J Cancer* 2006;**95**:1047–9.

Wells PS, Anderson JL, Scarvelis DK, Doucette SP, Gagnon F. Factor XIII Val34Leu variant is protective

against venous thromboembolism: a HuGE review and meta-analysis. *Am J Epidemiol* 2006;**164**:101–9.

Yang YC, Chang TY, Lee YJ, Su TH, Dang CW, Wu CC, *et al.* HLA-DRB1 alleles and cervical squamous cell carcinoma: experimental study and meta-analysis. *Hum Immunol* 2006;**67**:331–40.

Zintzaras E, Kitsios G, Stefanidis I. Endothelial NO synthase gene polymorphisms and hypertension: a meta-analysis. *Hypertension* 2006;**48**:700–10.

Reviews that explicitly tested publication bias (*n* = 47)

Bell CM, Urbach DR, Ray JG, Bayoumi A, Rosen AB, Greenberg D, *et al.* Bias in published cost effectiveness studies: systematic review. *BMJ* 2006;**332**:699–703.

Boodhwani M, Rubens FD, Sellke FW, Mesana TG, Ruel M. Mortality and myocardial infarction following surgical versus percutaneous revascularization of isolated left anterior descending artery disease: a meta-analysis. *Eur J Cardiothorac Surg* 2006;**29**:65–70.

Chen XC, Xu MT, Zhou W, Han CL, Chen WQ. A meta-analysis of relationship between beta-fibrinogen gene -148C/T polymorphism and susceptibility to cerebral infarction in Han Chinese. *Chin Med J (Engl)* 2007;**120**:1198–202.

Chodosh J, Morton SC, Mojica W, Maglione M, Suttorp MJ, Hilton L, *et al.* Meta-analysis: chronic disease self-management programs for older adults. *Ann Intern Med* 2005;**143**:427–38.

Clark EM, Tobias JH, Ness AR. Association between bone density and fractures in children: a systematic review and meta-analysis. *Pediatrics* 2006;**117**:e291–7.

Cruciani M, Lipsky BA, Mengoli C, de Lalla F. Are granulocyte colony-stimulating factors beneficial in treating diabetic foot infections?: A meta-analysis. *Diabetes Care* 2005;**28**:454–60.

Cruciani M, Mengoli C, Malena M, Bosco O, Serpelloni G, Grossi P. Antifungal prophylaxis in liver transplant patients: a systematic review and meta-analysis. *Liver Transpl* 2006;**12**:850–8.

Cruz DN, Perazella MA, Bellomo R, de Cal M, Polanco N, Corradi V, *et al.* Effectiveness of polymyxin B-immobilized fiber column in sepsis: a systematic review. *Crit Care* 2007;**11**:R47.

Dherani M, Pope D, Mascarenhas M, Smith KR, Weber M, Bruce N. Indoor air pollution from unprocessed solid fuel use and pneumonia risk in children aged under five years: a systematic review and meta-analysis. *Bull World Health Organ* 2008;**86**:390–8C.

Doulton TW, He FJ, MacGregor GA. Systematic review of combined angiotensin-converting enzyme inhibition and angiotensin receptor blockade in hypertension. *Hypertension* 2005;**45**:880–6.

Doust JA, Pietrzak E, Dobson A, Glasziou P. How well does B-type natriuretic peptide predict death and cardiac events in patients with heart failure: systematic review. *BMJ* 2005;**330**:625.

Garcia-Closas M, Malats N, Silverman D, Dosemeci M, Kogevinas M, Hein DW, *et al.* NAT2 slow acetylation, GSTM1 null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses. *Lancet* 2005;**366**:649–59.

Gautam M, Cheruvattath R, Balan V. Recurrence of autoimmune liver disease after liver transplantation: a systematic review. *Liver Transpl* 2006;**12**:1813–24.

Griffin S, Ellis S, Fitzgerald-Barron A, Rose J, Egger M. Nebulised steroid in the treatment of croup: a systematic review of randomised controlled trials. *Br J Gen Pract* 2000;**50**:135–41.

Hayden JA, van Tulder MW, Malmivaara AV, Koes BW. Meta-analysis: exercise therapy for nonspecific low back pain. *Ann Intern Med* 2005;**142**:765–75.

Hayden JA, van Tulder MW, Tomlinson G. Systematic review: strategies for using exercise therapy to improve outcomes in chronic low back pain. *Ann Intern Med* 2005;**142**:776–85.

Hulten E, Jackson JL, Douglas K, George S, Villines TC. The effect of early, intensive statin therapy on acute coronary syndrome: a meta-analysis of randomized controlled trials. *Arch Intern Med* 2006;**166**:1814–21.

Kasper JS, Giovannucci E. A meta-analysis of diabetes mellitus and the risk of prostate cancer. *Cancer Epidemiol Biomarkers Prev* 2006;**15**:2056–62.

Li D, Sham PC, Owen MJ, He L. Meta-analysis shows significant association between dopamine system genes and attention deficit hyperactivity disorder (ADHD). *Hum Mol Genet* 2006;**15**:2276–84.

Liu T, Zeng D, Zeng C, He X. Association between MYOC.mt1 promoter polymorphism and risk of primary open-angle glaucoma: a systematic review and meta-analysis. *Med Sci Monit* 2008;**14**:RA87–93.

Maggard MA, Shugarman LR, Suttorp M, Maglione M, Sugerman HJ, Livingston EH, *et al.* Meta-analysis: surgical treatment of obesity. *Ann Intern Med* 2005;**142**:547–59.

Meert AP, Martin B, Delmotte P, Berghmans T, Lafitte JJ, Mascaux C, *et al.* The role of EGF-R expression on patient survival in lung cancer: a systematic review with meta-analysis. *Eur Respir J* 2002;**20**:975–81.

Minelli C, Thompson JR, Tobin MD, Abrams KR. An integrated approach to the meta-analysis of genetic association studies using Mendelian randomization. *Am J Epidemiol* 2004;**160**:445–52.

Morris RK, Khan KS, Coomarasamy A, Robson SC, Kleijnen J. The value of predicting restriction of fetal growth and compromise of its wellbeing: Systematic quantitative overviews (meta-analysis) of test accuracy literature. *BMC Pregnancy Childbirth* 2007;**7**:3.

Muradin GS, Bosch JL, Stijnen T, Hunink MG. Balloon dilation and stent implantation for treatment of femoropopliteal arterial disease: meta-analysis. *Radiology* 2001;**221**:137–45.

Newman DJ, Mattock MB, Dawnay AB, Kerry S, McGuire A, Yaqoob M, *et al.* Systematic review on urine albumin testing for early detection of diabetic complications. *Health Technol Assess* 2005;**9**:iii–vi, xiii–163.

Nowak AK, Stockler MR, Chow PK, Findlay M. Use of tamoxifen in advanced-stage hepatocellular carcinoma. A systematic review. *Cancer* 2005;**103**:1408–14.

Ntais C, Polycarpou A, Ioannidis JP. Association of the CYP17 gene polymorphism with the risk of prostate cancer: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2003;**12**:120–6.

Orlando LA, Kulasingam SL, Matchar DB. Meta-analysis: the detection of pancreatic malignancy with positron emission tomography. *Aliment Pharmacol Ther* 2004;**20**:1063–70.

Owen CG, Shah A, Henshaw K, Smeeth L, Sheikh A. Topical treatments for seasonal allergic conjunctivitis: systematic review and meta-analysis of efficacy and effectiveness. *Br J Gen Pract* 2004;**54**:451–6.

Owen CG, Whincup PH, Gilg JA, Cook DG. Effect of breast feeding in infancy on blood pressure in later life: systematic review and meta-analysis. *BMJ* 2003;**327**:1189–95.

Ownby RL, Crocco E, Acevedo A, John V, Loewenstein D. Depression and risk for Alzheimer disease: systematic review, meta-analysis, and metaregression analysis. *Arch Gen Psychiatry* 2006;**63**:530–8.

Petticrew M, Bell R, Hunter D. Influence of psychological coping on survival and recurrence in people with cancer: systematic review. *BMJ* 2002;**325**:1066.

Qin LQ, Xu JY, Wang PY, Hoshi K. Soyfood intake in the prevention of breast cancer risk in women: a meta-analysis of observational epidemiological studies. *J Nutr Sci Vitaminol (Tokyo)* 2006;**52**:428–36.

Schernhammer ES, Colditz GA. Suicide rates among physicians: a quantitative and gender assessment (meta-analysis). *Am J Psychiatry* 2004;**161**:2295–302.

Selvin E, Marinopoulos S, Berkenblit G, Rami T, Brancati FL, Powe NR, *et al.* Meta-analysis: glycosylated hemoglobin and cardiovascular disease in diabetes mellitus. *Ann Intern Med* 2004;**141**:421–31.

Stone J, Sharpe M, Carson A, Lewis SC, Thomas B, Goldbeck R, *et al.* Are functional motor and sensory symptoms really more frequent on the left? A systematic review. *J Neurol Neurosurg Psychiatry* 2002;**73**:578–81.

Strippoli GF, Navaneethan SD, Johnson DW, Perkovic V, Pellegrini F, Nicolucci A, *et al.*. Effects of statins in patients with chronic kidney disease: meta-analysis and meta-regression of randomised controlled trials. *BMJ* 2008;**336**:645–51.

Tsai AC, Morton SC, Mangione CM, Keeler EB. A meta-analysis of interventions to improve care for chronic illnesses. *Am J Manag Care* 2005;**11**:478–88.

van Kempen EE, Kruize H, Boshuizen HC, Ameling CB, Staatsen BA, de Hollander AE. The association between noise exposure and blood pressure and ischemic heart disease: a meta-analysis. *Environ Health Perspect* 2002;**110**:307–17.

Van Maele-Fabry G, Willems JL. Occupation related pesticide exposure and cancer of the prostate: a meta-analysis. *Occup Environ Med* 2003;**60**:634–42.

Veglia F, Matullo G, Vineis P. Bulky DNA adducts and risk of cancer: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2003;**12**:157–60.

Wasnich RD, Miller PD. Antifracture efficacy of antiresorptive agents are related to changes in bone density. *J Clin Endocrinol Metab* 2000;**85**:231–6.

Wellman RJ, Sugarman DB, DiFranza JR, Winickoff JP. The extent to which tobacco marketing and tobacco use in films contribute to children's use of tobacco: a meta-analysis. *Arch Pediatr Adolesc Med* 2006;**160**:1285–96.

Winkley K, Ismail K, Landau S, Eisler I. Psychological interventions to improve glycaemic control in patients with type 1 diabetes: systematic review and meta-analysis of randomised controlled trials. *BMJ* 2006;**333**:65.

Zafarmand MH, van der Schouw YT, Grobbee DE, de Leeuw PW, Bots ML. The M235T polymorphism in the AGT gene and CHD risk: evidence of a Hardy-Weinberg equilibrium violation and publication bias in a meta-analysis. *PLoS ONE* 2008;**3**:e2533.

Zheng M, Bai J, Yuan B, Lin F, You J, Lu M, *et al.* Meta-analysis of prophylactic corticosteroid use in post-ERCP pancreatitis. *BMC Gastroenterology* 2008;**8**:6.

Appendix 18
Original study proposal

II.   DETAILS OF PROPOSED RESEARCH

Detailed outline of proposed research (see attached notes for guidance).

Title:

RM05/JH29: Dissemination bias in medical and health related research - an updated synthesis of empirical evidence and a critical assessment of available methods

Background

Synthesis of published research is becoming increasingly important in providing relevant and valid research evidence to clinical and health policy decision making. However, the validity of research synthesis based on published literature will be threatened if published studies comprise a biased selection of all studies that have been conducted.

A previous HTA monograph published in 2000 systematically reviewed studies that provided empirical evidence on publication and related biases, and studies that developed or tested methods for preventing, reducing, or detecting publication and related biases.[1] The review found evidence indicating that studies with significant or favourable results were more likely to be published, or were likely to be published earlier than those with non-significant results. There was limited and indirect evidence indicating the possibility of full publication bias, outcome reporting bias, duplicate publication bias, and language bias. The review identified little empirical evidence relating to the impact of publication and related biases on health policy, clinical decision making and the outcome of patient management. Considering that the spectrum of the accessibility of research results (dissemination profile) ranges from completely inaccessible to easily accessible, it was suggested that a single term "dissemination bias" could be used to denote all types of publication and related biases.[1]

In the previous HTA report, the available methods for dealing with dissemination biases were classified according to measures that could be taken before, during or after a literature review: to prevent publication bias before a literature review (eg, prospective registration of trials), to reduce or detect publication and related biases during a literature review (eg, locating grey literature or unpublished studies, and funnel plot related methods), and to minimise the impact of publication bias after a literature review (eg, confirmatory large scale trials, updating systematic reviews).[1] It was concluded that the ideal solution to publication bias is the prospective, universal registration of all studies at their inception. It was concluded, although debatable, that available statistical methods for detecting and adjusting publication bias should be mainly used for the purpose of sensitivity analysis.

Since the publication of the HTA review of publication and related biases, many new studies on publication and related biases have been published. For example, Egger et al (2003) provided further empirical evidence on publication bias, language bias, grey literature bias, and MEDLINE index bias,[2] and Moher et al (2003) evaluated language bias in meta-analyses of randomised controlled trials.[3] Recently, more convincing evidence on outcome reporting bias has been published.[4,5] The new empirical evidence may contradict or strengthen the empirical evidence included in the previous HTA report. Funnel plot and related statistical methods have been applied in new studies to collections of meta-analyses to estimate possible publication bias in systematic reviews.[6-8] There are also new published studies that investigated methods for dealing with publication bias (for example[9-11]).

Dubben and Beck-Bornholdt (2005) used the funnel plot approach, and found no evidence of publication bias in studies of publication bias.[12] They acknowledged that the analysis was handicapped by insufficient power (with only 26 included studies) and also by the diverse definitions of publication bias in the primary studies. However, Song et al pointed out that the study had other, more important, limitations so that dissemination bias of studies on publication bias could not be safely excluded.[13]

Purpose (aims, objectives)

1. To identify all relevant empirical studies published since 1998. Empirical studies are defined as those that provide empirical evidence on the existence, consequences, causes and/or risk factors of dissemination bias.

2. To identify all relevant methodological studies published since 1998. Methodological studies are those that have developed or investigated methods for preventing, reducing or detecting dissemination bias.

3. To categorise empirical and methodological studies identified according to a conceptual framework of dissemination profile, and to critically appraise studies that provided direct empirical evidence.

4. To synthesise findings from newly identified and previously included studies to enable us to assess whether each type of dissemination bias does exist, and if so the extent of the effect that it may have on results of systematic reviews and hence decision making.

5. To assess the possibility of dissemination bias of studies that provide empirical evidences on dissemination bias.

6. To assess the usefulness and limitations of available methods, and resources required to use these methods to combat each type of dissemination bias, through synthesis of the methodological studies.

7. To examine measures taken in a representative sample of published systematic reviews to prevent, reduce and detect different types of dissemination bias. We will include both narrative and quantitative (meta-analytic) systematic reviews that evaluated effect of healthcare interventions, systematic reviews that evaluated the accuracy of diagnostic tests, and systematic reviews of epidemiological studies that evaluated association of risk factors and health outcomes.

8. To bring together current evidence on the existence and scale of each type of dissemination bias, effects and costs of methods to combat these biases, and current use of these methods to create recommendations for reviewers, policy makers, health professionals and service users, and to disseminate these recommendations.

Investigation methods

The review contains three parts: (1) review of empirical and methodological studies; (2) an assessment of published systematic reviews; (3) synthesising all findings, providing and disseminating recommendations.

Part 1. Review of empirical and methodological studies

Methods used in the previous HTA report will be modified to identify and categorise relevant studies. We will adopt a new framework to categorise relevant studies, and important studies will be assessed using a more critical and structured approach. Details of the review methods are described below and in the Figure.

Criteria for inclusion and literature search strategies

A preliminary literature search indicated that there are a large number of potentially relevant studies in fields of medical and health related research, and searches in the area of social sciences produced few studies in the initial review, so we plan to focus on dissemination bias in health and related research. We will include studies that provide empirical evidence on the existence, consequences, causes, and/or risk factors of types of dissemination bias; and studies that develop or evaluate methods for preventing, reducing or detecting dissemination bias.

Literature searches for methodological studies are often difficult because of ill-defined boundaries and inappropriate indexing in commonly used bibliographic databases.[14] Our previous experience and initial searching suggests that the most productive and efficient methods include searching the Cochrane Methodology Register, references of retrieved articles and citation search of key studies.

We will search the **Cochrane Methodology Register** (CRM) and **MEDLINE** for relevant empirical and methodological studies published since 1998. We will compare studies identified from the MEDLINE and those identified from the Cochrane Methodology Register, to check the completeness and usefulness of the two bibliographic databases for methodological reviews. Key words used in the search of electronic databases will include: publication bias, dissemination bias, language bias, national bias, country bias, reporting bias, grey literature bias, conference/abstract bias, full publication bias, citation bias, time lag/delay bias, reference bias, selection bias, location bias, duplication or multiple publication bias, database bias, index bias, and file drawer. References (titles with or without abstracts) gathered by searching the CMR and MEDLINE will be assessed independently by two reviewers for inclusion or exclusion. Any disagreement will be discussed.

We will also search **EMBASE** (from 2005 to 2007, as EMBASE is searched for the Methodology Register, but we will ensure that we have included the most recent references), **Ahmed** (1998-2007), **Cinahl** (1998-2007), **PsychInfo** (1998-2007), **SIGLE** (1998-2007) and **Dissertation Abstracts** (1998-2007) for any additional relevant studies. Searching of EMBASE, Ahmed, Cinahl, PsychInfor, SIGLE, and Dissertation Abstracts will be conducted by one reviewer. **References of retrieved reviews and studies** will be examined by one reviewer to identify additional relevant studies, including any relevant studies published before 1998 but

missed in the previous HTA report. **Citations of the key studies** will also be searched. The literature search will not be restricted by publication language.

We have conducted a preliminary search of the Cochrane Methodology Register for relevant studies that published since 1998. According to titles (with or without abstracts) of identified references, there are a large number of possibly relevant studies (300-400, after excluding obvious duplicates). More than 200 references were empirical studies, including publication bias (n=26), publication of conference abstracts (n=66), outcome reporting bias (n=26), country or language bias (n=39), grey literature (n=13), time lag bias (n=15), causes of publication bias (n=16), citation bias (n=16). We found 58 studies of methods for dealing with publication bias, including 30 studies of statistical methods, 17 studies of literature search methods, and 11 studies of other methods (eg, trial registration or large scale confirmation trials).

Classification of identified studies

According to findings from our preliminary literature search, relevant studies are numerous in quantity and substantially diverse in quality. It is crucial to classify identified studies using a pre-specified structure to facilitate subsequent assessment and synthesis (Figure).

First , one reviewer will classify identified studies as (1) evidence studies or (2) methodological studies. Evidence studies are defined as studies that provide empirical evidence on the existence, extent, consequences, causes or risk factors of dissemination bias. Methodological studies are defined as those that develop or investigate methods for preventing, reducing or detecting dissemination bias. Some studies may be classified as both an evidence and a methods study.

Review of evidence studies

Evidence studies will be categorised into various types of dissemination bias, according to a framework of dissemination profile (that is, accessibility of research results): non-publication (never, or delayed); incomplete publication (e.g. biased outcome reporting, data dredged subgroup effects, biased full publication of conference abstracts); published but difficult to access (e.g. grey literature, language bias, database bias); other biased dissemination activities (e.g. citation bias, duplicate bias). It is possible that some studies may be included in more than one category.

Then evidence studies will be further separated into two groups: studies that provided direct evidence, and studies that provided indirect evidence. Direct evidence refers to data or observations that could be used to directly indicate dissemination bias, including admissions of bias on the part of those involved in the publication process, comparison of the results of published and unpublished studies, and the prospective or retrospective follow-up of dissemination profile of cohorts of studies. Indirect evidence refers to observations that could be explained indirectly by dissemination bias but other alternative explanations could not be excluded. For instance, a disproportionately high proportion of positive findings in the published literature might provide indirect evidence, as might larger effect sizes in smaller studies compared with larger studies.

We will apply a checklist of quality assessment to critically appraise studies that provided direct empirical evidence and studies that assessed association between sample size and effects in multiple meta-analyses, with regard to scientific rigorousness, the sample's representativeness, and appropriateness of data analyses and interpretation. Topic specific items will be considered if judged appropriate for different types of bias. More details about the proposed quality assessment are described below.

- *Scientific rigorousness*: We aim to detect potential threats to the validity of study results. For example, prospective studies are more valid than retrospective studies. Selection and inclusion of samples may be more or less biased, and whether assessments and judgements were independently duplicated.
- *Generalisability*: It is important to consider whether results of studies could be generalisable to different fields of research, settings, and designs. For example, dissemination bias in randomised trials may not be similar to that in epidemiological studies.
- *Appropriateness of data analysis*: We will assess whether the data analysis method is appropriate to address the objectives of the study.
- *Appropriateness of interpretation*: Limitations of the study should be considered when results of a study were interpreted.

Relevant studies included in the previous HTA report[1] will also be critically appraised using the same checklist.

Using a standardised appraisal form (Appendix 1), categorisation of all evidence studies and critical appraisal of selected studies will be independently carried out by two reviewers. Disagreements will be resolved by discussion.

Figure. Classification of identified relevant studies
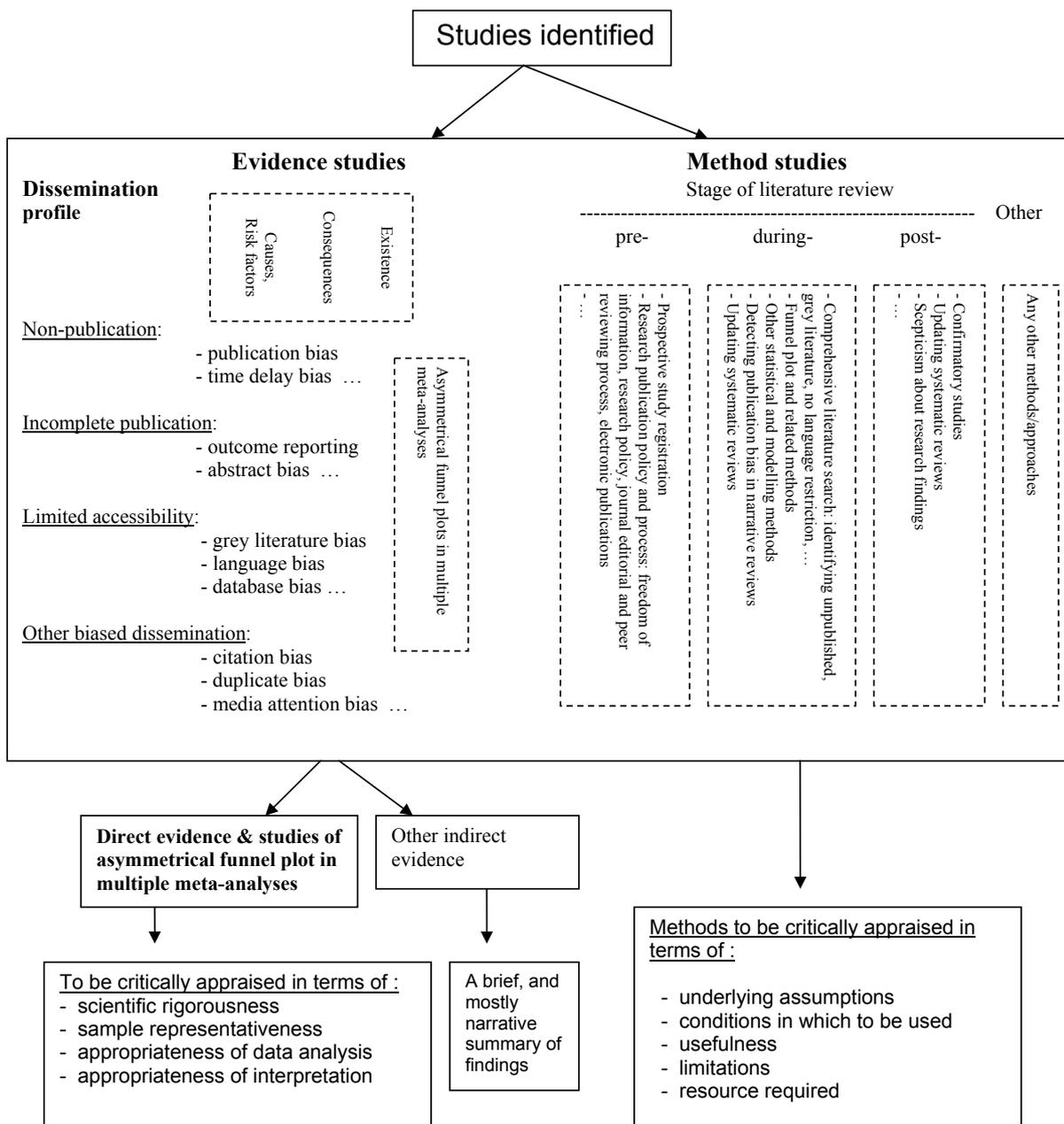
Assessment of dissemination bias in empirical evidence studies

In this review, we will assess whether dissemination bias is also a problem for studies that provide empirical evidence on dissemination bias. Results of included empirical evidence studies will be independently categorised by two reviewers as positive (significant dissemination bias), non-significant (no clear dissemination bias), or can't tell (see Appendix 1). Then we will examine the association of the results of empirical studies and studies' quality and dissemination profile (including time of publication, journal impact factor, number of citations after a given period of publication, and study's impact on guidelines/recommendations for systematic reviews).

Review of methodological studies

There may be multiple studies investigating the same method. Method studies will be categorised according to methods they investigated, to generate a list of available methods and corresponding studies identified. Then each method will be cross-classified from two aspects: (1) type of dissemination bias and (2) stage of literature review (see Figure). It is possible that the same method may be relevant to different types of dissemination bias or applicable to the different stage of a literature review. Using a standardised method classification sheet (Appendix 2), the review of method studies will be conducted by one reviewer and checked by a second reviewer. Disagreements will be resolved by discussion.

Based on findings of included studies, available methods will be critically appraised in terms of underlying assumptions, conditions under which the method could be used, usefulness, limitations, and resource required.

Presentation and summary of literature review findings

Data extracted from the included studies and results of critically appraisal will be presented in tables and described narratively. If appropriate, results from individual studies will be quantitatively pooled (for example, pooled odds ratio of full publication of conference abstracts with statistically significant results versus those with statistically non-significant results). Results of critical appraisal will be taken into consideration to interpret and explain findings from empirical and methodological studies. We will highlight whether findings from studies newly identified contradict or strengthen findings from studies included in the previous HTA report.

Part 2. Assessment of a sample of published systematic reviews

In the previous HTA report, 193 systematic reviews taken from the Database of Abstract of Reviews of Effectiveness (Centre for Reviews and Dissemination, University of York) were used to identify further evidence of dissemination bias and to illustrate the methods used in systematic reviews for dealing with publication bias. However, there are several shortcomings in the previous assessment. First, systematic reviews included in the DARE database might on average have better quality than those from the general bibliographic databases (such as MEDLINE) so that the representativeness of systematic reviews assessed in the previous HTA report may be questionable. Secondly, 91% of systematic reviews that evaluated the effectiveness of healthcare interventions and 9% that evaluated the accuracy of diagnostic technologies were not separately assessed. The problem of dissemination bias might be different between the two types of systematic reviews. Thirdly, systematic reviews of epidemiological studies of association between risk factors and health outcomes were not included in the previous HTA report.

To overcome these shortcomings, we plan to obtain a representative sample of systematic reviews from the general bibliographic database MEDLINE, including (1) systematic reviews of studies on effects of healthcare interventions, (2) systematic reviews of studies on accuracy of diagnostic tests, and (3) systematic reviews of epidemiological studies on association between risk factors and health outcomes.

> A preliminary search of MEDLINE using "systematic review" or "meta-analysis" (in titles or in abstracts) identified 2779 English-language references published in 2005. We examined the first 300 of the 2779 references and identified 109 systematic reviews that evaluated effects of healthcare interventions (including preventive interventions), 13 systematic reviews of studies of diagnostic tests, and 53 systematic reviews of epidemiological studies (including 18 systematic reviews of genetic studies). This preliminary exercise indicates that there are about 1009 systematic review of effects of health interventions, about 120 systematic reviews of diagnostic tests, and about 490 systematic reviews of epidemiological studies. The following approach is based on findings from this preliminary work.

Identifying and sampling systematic reviews

First, we will search MEDLINE for systematic reviews published in 2005. References identified from MEDLINE will be examined by one reviewer and categorised as systematic reviews of effects of healthcare interventions, systematic reviews of accuracy of diagnostic tests, or systematic reviews of epidemiological studies (genetic epidemiology or not). Then we will use computer-generated random numbers to obtain a random sample of 100 systematic reviews of effects of healthcare interventions, 50 systematic reviews of accuracy of diagnostic tests, and 100 systematic reviews of epidemiological studies (of which 50 will be reviews of genetic epidemiology studies) from all identified systematic reviews.

Extracting data from included systematic reviews

The data extraction from included systematic reviews will be independently conducted by two reviewers to collect the following information (see preliminary data extraction sheet in Appendix 3): type of review (effect, diagnostic, epidemiological), method of data synthesis (narrative or quantitative), whether the issue of publication bias was considered, whether unpublished studies or those published in non-English languages were searched for and included; methods used for dealing with publication bias; any evidence on the existence, extent and consequence of publication bias.

A checklist will also be applied independently by two reviewers to assess the overall quality of included systematic reviews (see Appendix 3). Any disagreements between the two reviewers will be resolved by discussion.

Analysing data from included systematic reviews

Data extracted from systematic reviews of effects of healthcare interventions, systematic reviews of accuracy of diagnostic test, and systematic reviews of epidemiological studies will be separately presented and compared. We will also examine the subgroup of genetic epidemiology reviews separately. We will compare findings from narrative systematic reviews and quantitative systematic reviews (meta-analyses). Systematic reviews of effects of healthcare interventions and systematic reviews of diagnostic accuracy published in 2005 will be compared with those included in the previous HTA report to examine whether the reporting and treatment of dissemination bias has improved over time.

Part 3. Synthesising findings from Part 1 and Part 2

Findings from Part 1 will illuminate the existence or otherwise, extent and potential impact on policy of different types of dissemination bias, and suggest a range of methods for dealing with such biases. Part 2 will provide findings about what actually happens in the practice of systematic reviews. Part 3 aims to compare findings from Part 1 and Part 2, and to identify gaps between empirical and methodological research on dissemination bias, and actual practice of systematic reviews (see the proposed summary table below). For example, considerable resource might be wasted in systematic reviews identifying, translating, and assessing studies published in languages other than English if evidence suggests that language bias is not a problem. Some statistical methods developed may have rarely been used in practice for various reasons (eg, too complicated or no additional advantages as compared with simple methods).

Bias category	Evidence of existence	Methods to combat this bias	Usefulness and limitations of the method	Resources required to use the method
Type of dissemination bias	Based on literature review	Based on literature review	Effect of the methods, as well as limitations, based on literature review and a sample of systematic reviews	Including time, staff, and costs. Based on literature review and experience

Main references

1. Song F, Eastwood AJ, Gilbody S, Duley L, Sutton AJ. Publication and related biases. *Health Technol Assess* 2000;**4**:1-115.

2. Egger M, Juni P, Bartlett C, Holenstein F, Sterne J. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. *Health Technol Assess.* 2003;**7**:1-76.

3. Moher D, Pham B, Lawson ML, Klassen TP. The inclusion of reports of randomised trials published in languages other than English in systematic reviews. *Health Technology Assessment* 2003;**7**:1-90.

4. Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *Jama* 2004;**291**:2457-65.

5. Chan AW,.Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. *BMJ* 2005;**330**:753.

6. Song F, Khan KS, Dinnes J, Sutton AJ. Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *Int J Epidemiol* 2002;**31**:88-95.

7. Sterne J, Jüni P, Schulz K, Altman D, Bartlett C, Egger M. Statistical methods for assessing the influence of study characteristics on treatment effects in 'meta-epidemiological' research. *Statistics in Medicine* 2002;**21**:1513-24.

8. Sutton AJ, Duval SJ, Tweedie RL, Abrams KR, Jones DR. Empirical assessment of effect of publication bias on meta-analyses. *BMJ* 2000;**320**:1574-7.

9. Terrin N, Schmid CH, Lau J, Olkin I. Adjusting for publication bias in the presence of heterogeneity. *Stat.Med* 2003;**22**:2113-26.

10. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005;**58**:882-93.

11. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Comparison of Two Methods to Detect Publication Bias in Meta-analysis. *JAMA: The Journal of the American Medical Association* 2006;**295**:676-80.

12. Dubben HH,.Beck-Bornholdt HP. Systematic review of publication bias in studies on publication bias. *BMJ* 2005.

13. Song F. Review of publication bias in studies on publication bias: studies on publication bias are probably susceptible to the bias they study. *BMJ* 2005;**331**:637-8.

14. Edwards S, Lilford RJ, Kiauka S. Different types of systematic review in health services research. In Black N, Brazier J, Fitzpatrick R, Reeves B, eds. *Health services research methods: a guide to best practice*, pp 255-9. London: BMJ Books, 1998.

Appendix 1. Data Extraction Sheet - Empirical Evidence Studies
Preliminary sheet, will be modified by pilot testing

Author (year): _____   Source: _____
Title: _____   Reviewer: _____

Study design & objectives:

Issues:
- ☐ Existence/identifying   ☐ Causes/risk factors
- ☐ Consequence   ☐ Other: _____

Categories:
- ☐ Non-publication   ☐ Incomplete publication
- ☐ Limited accessibility   ☐ Other: _____

Specific bias:
- ☐ Publication bias   ☐ Grey literature bias
- ☐ Language bias   ☐ Reporting bia
- ☐ Abstract bias   ☐ Time delay bias
- ☐ Database index bias   ☐ Citation bias
- ☐ Duplicate bias   ☐ Media attention bias
- ☐ Other: _____

Areas:
- ☐ general health   ☐ specific health (eg, obesity):
- ☐ other: _____

Study results: ☐ Significant/important bias   ☐ Non-significant   ☐ Can't tell
  Details:

Original authors' conclusions:

Evidence: ☐ Direct   ☐ Indirect
For indirect evidence, stop.   For studies with direct evidence, continue:

Scientific rigorousness (hints: prospective or retrospective, sample selection bias)

Sample representativeness (hints: research field, participants, outcomes, interventions; study designs)

Appropriateness of data analysis (hints: consider objectives, available data and methods of data analysis)

Appropriateness of interpretations (hints: limitations of the study should be taken into consideration)

Overall study quality: ☐ High (hint: without considerable concern on study validity)
                     ☐ Moderate   (hint: with some concern on study validity)
                     ☐ Low (hint: with considerable concern on study validity)
                     ☐ Can't tell

Any other comments:

Appendix 2. Data Extraction Sheet – Methodological Studies
  Preliminary sheet, will be modified by pilot testing

Author (year): _____   Source: _____   Reviewer: _____
Title: _____

Study design:

Study objectives:

    ☐ New method,   ☐ Established method, ☐ Evidence of usefulness/limitations
----
Methods:
    ☐ Study registration   ☐ Literature search
    ☐ Funnel plot   ☐ Statistical/modelling
    ☐ Updating reviews   ☐ Publication process
    ☐ Research ethics/policy   ☐ Confirmatory studies
    ☐ other: _____

Purpose:
    ☐ Preventing bias ☐ Reducing bias, ☐ Detecting bias, ☐ Adjusting bias,
    ☐ other: _____

Stage of literature review:
    ☐ Before literature review,   ☐ In literature review,   ☐ After literature review
----
What dissemination bias the method is relevant:
    ☐ Publication bias   ☐ Grey literature bias
    ☐ Language bias   ☐ Reporting bias
    ☐ Abstract bias   ☐ Time delay bias
    ☐ Database index bias   ☐ Citation bias
    ☐ Duplicate bias   ☐ Media attention bias
    ☐ Other: _____

Main findings and conclusions:

Resources required to use the method:

Reviewer's commentary (study's validity, scientific rigorousness, method's usefulness and limitations,
any empirical evidence provided):

Appendix 3. Data Extraction Sheet – Systematic reviews
Preliminary sheet, will be modified by pilot testing

Author (year): _____ **Source:** _____ Reviewer: _____

Objectives: _____ □ Effectiveness/adverse effects □. Diagnostic □. Genetic epidemiology
                                                                                                          □. Other epidemiology

Type of reviews: □. Narrative □. Meta-analysis

Designs of included studies: □ RCTs/CCTs (Study= ; patients=)
                                 □ Diagnostic accuracy studies (Study= ; patients=)
                                 □ Epidemiological studies (Study= ; patients=)
                                 □ Other_____ (Study= ; patients=)

How were differences between studies investigated?
    □ NA □ Narrative □ Statistical □ Meta-regression □ Sensitivity/subgrouup □ Other

Authors' conclusion:
    □ Significant/positive: At least one intervention recommended; or sig. difference found between interventions.
    □ Non-sig./not important: No intervention is recommended, or no sig. differences found among interventions.
    □ Unclear: No able to judge; neither positive nor negative; lack of evidence.

Sources searched to identify studies:
    □ Not stated
    □ MEDLINE □ EMBASE □ Psychlit □ Cochrane □ Bibliographies
    □ Handsearch □ Experts/authors □ Company □ Proceedings
    □ Other: _____

Non-English language studies:
    □ Unclear
    □ Searched Yes No If yes, search methods:
    □ Identified Yes No How many?
    □ Included Yes No If included, a). for main analysis b). for sensitivity analysis?

Unpublished studies:
    □ Unclear
    □ Searched Yes No If yes, search methods:
    □ Identified Yes No How many?
    □ Included Yes No If included, a). for main analysis b). for sensitivity analysis?

Issue of publication bias discussed? □ No □ Yes

Methods used for dealing with publication bias:
    □ Not used □ Identify unpublished studies
    □ Prospective register □ Fail-safe N
    □ Funnel plot □ Rank correlation
    □ Egger's method □ Large scale trials
    □ Modelling □ Other: _____
  Details:

Evidence on publication bias □ Not available □ Available, If available, details
    (such as, results of published trials versus unpublished trials; or shape of Funnel plot or related methods)

Systematic review's overall quality:

1. Well defined review question	□ Yes	□ Partially	□ No	□ Can't tell
2. Identification of all relevant studies	□ Yes	□ Partially	□ No	□ Can't tell
3. Appropriate assessment of study quality	□ Yes	□ Partially	□ No	□ Can't tell
4. Reliable and accurate data extraction	□ Yes	□ Partially	□ No	□ Can't tell
5. Appropriate investigation of heterogeneity	□ Yes	□ Partially	□ No	□ Can't tell
6. Appropriate data synthesis	□ Yes	□ Partially	□ No	□ Can't tell
7. Appropriate interpretation of results	□ Yes	□ Partially	□ No	□ Can't tell

Health Technology Assessment reports published to date

Volume 1, 1997

No. 1
Home parenteral nutrition: a systematic review.
By Richards DM, Deeks JJ, Sheldon TA, Shaffer JL.

No. 2
Diagnosis, management and screening of early localised prostate cancer.
A review by Selley S, Donovan J, Faulkner A, Coast J, Gillatt D.

No. 3
The diagnosis, management, treatment and costs of prostate cancer in England and Wales.
A review by Chamberlain J, Melia J, Moss S, Brown J.

No. 4
Screening for fragile X syndrome.
A review by Murray J, Cuckle H, Taylor G, Hewison J.

No. 5
A review of near patient testing in primary care.
By Hobbs FDR, Delaney BC, Fitzmaurice DA, Wilson S, Hyde CJ, Thorpe GH, *et al*.

No. 6
Systematic review of outpatient services for chronic pain control.
By McQuay HJ, Moore RA, Eccleston C, Morley S, de C Williams AC.

No. 7
Neonatal screening for inborn errors of metabolism: cost, yield and outcome.
A review by Pollitt RJ, Green A, McCabe CJ, Booth A, Cooper NJ, Leonard JV, *et al*.

No. 8
Preschool vision screening.
A review by Snowdon SK, Stewart-Brown SL.

No. 9
Implications of socio-cultural contexts for the ethics of clinical trials.
A review by Ashcroft RE, Chadwick DW, Clark SRL, Edwards RHT, Frith L, Hutton JL.

No. 10
A critical review of the role of neonatal hearing screening in the detection of congenital hearing impairment.
By Davis A, Bamford J, Wilson I, Ramkalawan T, Forshaw M, Wright S.

No. 11
Newborn screening for inborn errors of metabolism: a systematic review.
By Seymour CA, Thomason MJ, Chalmers RA, Addison GM, Bain MD, Cockburn F, *et al*.

No. 12
Routine preoperative testing: a systematic review of the evidence.
By Munro J, Booth A, Nicholl J.

No. 13
Systematic review of the effectiveness of laxatives in the elderly.
By Petticrew M, Watt I, Sheldon T.

No. 14
When and how to assess fast-changing technologies: a comparative study of medical applications of four generic technologies.
A review by Mowatt G, Bower DJ, Brebner JA, Cairns JA, Grant AM, McKee L.

Volume 2, 1998

No. 1
Antenatal screening for Down's syndrome.
A review by Wald NJ, Kennard A, Hackshaw A, McGuire A.

No. 2
Screening for ovarian cancer: a systematic review.
By Bell R, Petticrew M, Luengo S, Sheldon TA.

No. 3
Consensus development methods, and their use in clinical guideline development.
A review by Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CFB, Askham J, *et al*.

No. 4
A cost–utility analysis of interferon beta for multiple sclerosis.
By Parkin D, McNamee P, Jacoby A, Miller P, Thomas S, Bates D.

No. 5
Effectiveness and efficiency of methods of dialysis therapy for end-stage renal disease: systematic reviews.
By MacLeod A, Grant A, Donaldson C, Khan I, Campbell M, Daly C, *et al*.

No. 6
Effectiveness of hip prostheses in primary total hip replacement: a critical review of evidence and an economic model.
By Faulkner A, Kennedy LG, Baxter K, Donovan J, Wilkinson M, Bevan G.

No. 7
Antimicrobial prophylaxis in colorectal surgery: a systematic review of randomised controlled trials.
By Song F, Glenny AM.

No. 8
Bone marrow and peripheral blood stem cell transplantation for malignancy.
A review by Johnson PWM, Simnett SJ, Sweetenham JW, Morgan GJ, Stewart LA.

No. 9
Screening for speech and language delay: a systematic review of the literature.
By Law J, Boyle J, Harris F, Harkness A, Nye C.

No. 10
Resource allocation for chronic stable angina: a systematic review of effectiveness, costs and cost-effectiveness of alternative interventions.
By Sculpher MJ, Petticrew M, Kelland JL, Elliott RA, Holdright DR, Buxton MJ.

No. 11
Detection, adherence and control of hypertension for the prevention of stroke: a systematic review.
By Ebrahim S.

No. 12
Postoperative analgesia and vomiting, with special reference to day-case surgery: a systematic review.
By McQuay HJ, Moore RA.

No. 13
Choosing between randomised and nonrandomised studies: a systematic review.
By Britton A, McKee M, Black N, McPherson K, Sanderson C, Bain C.

No. 14
Evaluating patient-based outcome measures for use in clinical trials.
A review by Fitzpatrick R, Davey C, Buxton MJ, Jones DR.

No. 15
Ethical issues in the design and conduct of randomised controlled trials.
A review by Edwards SJL, Lilford RJ, Braunholtz DA, Jackson JC, Hewison J, Thornton J.

No. 16
Qualitative research methods in health technology assessment: a review of the literature.
By Murphy E, Dingwall R, Greatbatch D, Parker S, Watson P.

No. 17
The costs and benefits of paramedic skills in pre-hospital trauma care.
By Nicholl J, Hughes S, Dixon S, Turner J, Yates D.

No. 18
Systematic review of endoscopic ultrasound in gastro-oesophageal cancer.
By Harris KM, Kelly S, Berry E, Hutton J, Roderick P, Cullingworth J, *et al.*

No. 19
Systematic reviews of trials and other studies.
By Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F.

No. 20
Primary total hip replacement surgery: a systematic review of outcomes and modelling of cost-effectiveness associated with different prostheses.
A review by Fitzpatrick R, Shortall E, Sculpher M, Murray D, Morris R, Lodge M, *et al.*

Volume 3, 1999

No. 1
Informed decision making: an annotated bibliography and systematic review.
By Bekker H, Thornton JG, Airey CM, Connelly JB, Hewison J, Robinson MB, *et al.*

No. 2
Handling uncertainty when performing economic evaluation of healthcare interventions.
A review by Briggs AH, Gray AM.

No. 3
The role of expectancies in the placebo effect and their use in the delivery of health care: a systematic review.
By Crow R, Gage H, Hampson S, Hart J, Kimber A, Thomas H.

No. 4
A randomised controlled trial of different approaches to universal antenatal HIV testing: uptake and acceptability. Annex: Antenatal HIV testing – assessment of a routine voluntary approach.
By Simpson WM, Johnstone FD, Boyd FM, Goldberg DJ, Hart GJ, Gormley SM, *et al.*

No. 5
Methods for evaluating area-wide and organisation-based interventions in health and health care: a systematic review.
By Ukoumunne OC, Gulliford MC, Chinn S, Sterne JAC, Burney PGJ.

No. 6
Assessing the costs of healthcare technologies in clinical trials.
A review by Johnston K, Buxton MJ, Jones DR, Fitzpatrick R.

No. 7
Cooperatives and their primary care emergency centres: organisation and impact.
By Hallam L, Henthorne K.

No. 8
Screening for cystic fibrosis.
A review by Murray J, Cuckle H, Taylor G, Littlewood J, Hewison J.

No. 9
A review of the use of health status measures in economic evaluation.
By Brazier J, Deverill M, Green C, Harper R, Booth A.

No. 10
Methods for the analysis of quality-of-life and survival data in health technology assessment.
A review by Billingham LJ, Abrams KR, Jones DR.

No. 11
Antenatal and neonatal haemoglobinopathy screening in the UK: review and economic analysis.
By Zeuner D, Ades AE, Karnon J, Brown J, Dezateux C, Anionwu EN.

No. 12
Assessing the quality of reports of randomised trials: implications for the conduct of meta-analyses.
A review by Moher D, Cook DJ, Jadad AR, Tugwell P, Moher M, Jones A, *et al.*

No. 13
'Early warning systems' for identifying new healthcare technologies.
By Robert G, Stevens A, Gabbay J.

No. 14
A systematic review of the role of human papillomavirus testing within a cervical screening programme.
By Cuzick J, Sasieni P, Davies P, Adams J, Normand C, Frater A, *et al.*

No. 15
Near patient testing in diabetes clinics: appraising the costs and outcomes.
By Grieve R, Beech R, Vincent J, Mazurkiewicz J.

No. 16
Positron emission tomography: establishing priorities for health technology assessment.
A review by Robert G, Milne R.

No. 17 (Pt 1)
The debridement of chronic wounds: a systematic review.
By Bradley M, Cullum N, Sheldon T.

No. 17 (Pt 2)
Systematic reviews of wound care management: (2) Dressings and topical agents used in the healing of chronic wounds.
By Bradley M, Cullum N, Nelson EA, Petticrew M, Sheldon T, Torgerson D.

No. 18
A systematic literature review of spiral and electron beam computed tomography: with particular reference to clinical applications in hepatic lesions, pulmonary embolus and coronary artery disease.
By Berry E, Kelly S, Hutton J, Harris KM, Roderick P, Boyce JC, *et al.*

No. 19
What role for statins? A review and economic model.
By Ebrahim S, Davey Smith G, McCabe C, Payne N, Pickin M, Sheldon TA, *et al.*

No. 20
Factors that limit the quality, number and progress of randomised controlled trials.
A review by Prescott RJ, Counsell CE, Gillespie WJ, Grant AM, Russell IT, Kiauka S, *et al.*

No. 21
Antimicrobial prophylaxis in total hip replacement: a systematic review.
By Glenny AM, Song F.

No. 22
Health promoting schools and health promotion in schools: two systematic reviews.
By Lister-Sharp D, Chapman S, Stewart-Brown S, Sowden A.

No. 23
Economic evaluation of a primary care-based education programme for patients with osteoarthritis of the knee.
A review by Lord J, Victor C, Littlejohns P, Ross FM, Axford JS.

Volume 4, 2000

No. 1
The estimation of marginal time preference in a UK-wide sample (TEMPUS) project.
A review by Cairns JA, van der Pol MM.

No. 2
Geriatric rehabilitation following fractures in older people: a systematic review.
By Cameron I, Crotty M, Currie C, Finnegan T, Gillespie L, Gillespie W, *et al*.

No. 3
Screening for sickle cell disease and thalassaemia: a systematic review with supplementary research.
By Davies SC, Cronin E, Gill M, Greengross P, Hickman M, Normand C.

No. 4
Community provision of hearing aids and related audiology services.
A review by Reeves DJ, Alborz A, Hickson FS, Bamford JM.

No. 5
False-negative results in screening programmes: systematic review of impact and implications.
By Petticrew MP, Sowden AJ, Lister-Sharp D, Wright K.

No. 6
Costs and benefits of community postnatal support workers: a randomised controlled trial.
By Morrell CJ, Spiby H, Stewart P, Walters S, Morgan A.

No. 7
Implantable contraceptives (subdermal implants and hormonally impregnated intrauterine systems) versus other forms of reversible contraceptives: two systematic reviews to assess relative effectiveness, acceptability, tolerability and cost-effectiveness.
By French RS, Cowan FM, Mansour DJA, Morris S, Procter T, Hughes D, *et al*.

No. 8
An introduction to statistical methods for health technology assessment.
A review by White SJ, Ashby D, Brown PJ.

No. 9
Disease-modifying drugs for multiple sclerosis: a rapid and systematic review.
By Clegg A, Bryant J, Milne R.

No. 10
Publication and related biases.
A review by Song F, Eastwood AJ, Gilbody S, Duley L, Sutton AJ.

No. 11
Cost and outcome implications of the organisation of vascular services.
By Michaels J, Brazier J, Palfreyman S, Shackley P, Slack R.

No. 12
Monitoring blood glucose control in diabetes mellitus: a systematic review.
By Coster S, Gulliford MC, Seed PT, Powrie JK, Swaminathan R.

No. 13
The effectiveness of domiciliary health visiting: a systematic review of international studies and a selective review of the British literature.
By Elkan R, Kendrick D, Hewitt M, Robinson JJA, Tolley K, Blair M, *et al*.

No. 14
The determinants of screening uptake and interventions for increasing uptake: a systematic review.
By Jepson R, Clegg A, Forbes C, Lewis R, Sowden A, Kleijnen J.

No. 15
The effectiveness and cost-effectiveness of prophylactic removal of wisdom teeth.
A rapid review by Song F, O'Meara S, Wilson P, Golder S, Kleijnen J.

No. 16
Ultrasound screening in pregnancy: a systematic review of the clinical effectiveness, cost-effectiveness and women's views.
By Bricker L, Garcia J, Henderson J, Mugford M, Neilson J, Roberts T, *et al*.

No. 17
A rapid and systematic review of the effectiveness and cost-effectiveness of the taxanes used in the treatment of advanced breast and ovarian cancer.
By Lister-Sharp D, McDonagh MS, Khan KS, Kleijnen J.

No. 18
Liquid-based cytology in cervical screening: a rapid and systematic review.
By Payne N, Chilcott J, McGoogan E.

No. 19
Randomised controlled trial of non-directive counselling, cognitive–behaviour therapy and usual general practitioner care in the management of depression as well as mixed anxiety and depression in primary care.
By King M, Sibbald B, Ward E, Bower P, Lloyd M, Gabbay M, *et al*.

No. 20
Routine referral for radiography of patients presenting with low back pain: is patients' outcome influenced by GPs' referral for plain radiography?
By Kerry S, Hilton S, Patel S, Dundas D, Rink E, Lord J.

No. 21
Systematic reviews of wound care management: (3) antimicrobial agents for chronic wounds; (4) diabetic foot ulceration.
By O'Meara S, Cullum N, Majid M, Sheldon T.

No. 22
Using routine data to complement and enhance the results of randomised controlled trials.
By Lewsey JD, Leyland AH, Murray GD, Boddy FA.

No. 23
Coronary artery stents in the treatment of ischaemic heart disease: a rapid and systematic review.
By Meads C, Cummins C, Jolly K, Stevens A, Burls A, Hyde C.

No. 24
Outcome measures for adult critical care: a systematic review.
By Hayes JA, Black NA, Jenkinson C, Young JD, Rowan KM, Daly K, *et al*.

No. 25
A systematic review to evaluate the effectiveness of interventions to promote the initiation of breastfeeding.
By Fairbank L, O'Meara S, Renfrew MJ, Woolridge M, Sowden AJ, Lister-Sharp D.

No. 26
Implantable cardioverter defibrillators: arrhythmias. A rapid and systematic review.
By Parkes J, Bryant J, Milne R.

No. 27
Treatments for fatigue in multiple sclerosis: a rapid and systematic review.
By Brañas P, Jordan R, Fry-Smith A, Burls A, Hyde C.

No. 28
Early asthma prophylaxis, natural history, skeletal development and economy (EASE): a pilot randomised controlled trial.
By Baxter-Jones ADG, Helms PJ, Russell G, Grant A, Ross S, Cairns JA, *et al*.

No. 29
Screening for hypercholesterolaemia versus case finding for familial hypercholesterolaemia: a systematic review and cost-effectiveness analysis.
By Marks D, Wonderling D, Thorogood M, Lambert H, Humphries SE, Neil HAW.

No. 30
A rapid and systematic review of the clinical effectiveness and cost-effectiveness of glycoprotein IIb/IIIa antagonists in the medical management of unstable angina.
By McDonagh MS, Bachmann LM, Golder S, Kleijnen J, ter Riet G.

No. 21
Systematic reviews of the effectiveness of day care for people with severe mental disorders: (1) Acute day hospital versus admission; (2) Vocational rehabilitation; (3) Day hospital versus outpatient care.
By Marshall M, Crowther R, Almaraz- Serrano A, Creed F, Sledge W, Kluiter H, *et al*.

No. 22
The measurement and monitoring of surgical adverse events.
By Bruce J, Russell EM, Mollison J, Krukowski ZH.

No. 23
Action research: a systematic review and guidance for assessment.
By Waterman H, Tillen D, Dickson R, de Koning K.

No. 24
A rapid and systematic review of the clinical effectiveness and cost-effectiveness of gemcitabine for the treatment of pancreatic cancer.
By Ward S, Morris E, Bansback N, Calvert N, Crellin A, Forman D, *et al*.

No. 25
A rapid and systematic review of the evidence for the clinical effectiveness and cost-effectiveness of irinotecan, oxaliplatin and raltitrexed for the treatment of advanced colorectal cancer.
By Lloyd Jones M, Hummel S, Bansback N, Orr B, Seymour M.

No. 26
Comparison of the effectiveness of inhaler devices in asthma and chronic obstructive airways disease: a systematic review of the literature.
By Brocklebank D, Ram F, Wright J, Barry P, Cates C, Davies L, *et al*.

No. 27
The cost-effectiveness of magnetic resonance imaging for investigation of the knee joint.
By Bryan S, Weatherburn G, Bungay H, Hatrick C, Salas C, Parry D, *et al*.

No. 28
A rapid and systematic review of the clinical effectiveness and cost-effectiveness of topotecan for ovarian cancer.
By Forbes C, Shirran L, Bagnall A-M, Duffy S, ter Riet G.

No. 29
Superseded by a report published in a later volume.

No. 30
The role of radiography in primary care patients with low back pain of at least 6 weeks duration: a randomised (unblinded) controlled trial.
By Kendrick D, Fielding K, Bentley E, Miller P, Kerslake R, Pringle M.

No. 31
Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients.
By McColl E, Jacoby A, Thomas L, Soutter J, Bamford C, Steen N, *et al*.

No. 32
A rapid and systematic review of the clinical effectiveness and cost-effectiveness of paclitaxel, docetaxel, gemcitabine and vinorelbine in non-small-cell lung cancer.
By Clegg A, Scott DA, Sidhu M, Hewitson P, Waugh N.

No. 33
Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives.
By Brookes ST, Whitley E, Peters TJ, Mulheran PA, Egger M, Davey Smith G.

No. 34
Depot antipsychotic medication in the treatment of patients with schizophrenia: (1) Meta-review; (2) Patient and nurse attitudes.
By David AS, Adams C.

No. 35
A systematic review of controlled trials of the effectiveness and cost-effectiveness of brief psychological treatments for depression.
By Churchill R, Hunot V, Corney R, Knapp M, McGuire H, Tylee A, *et al*.

No. 36
Cost analysis of child health surveillance.
By Sanderson D, Wright D, Acton C, Duree D.

Volume 6, 2002

No. 1
A study of the methods used to select review criteria for clinical audit.
By Hearnshaw H, Harker R, Cheater F, Baker R, Grimshaw G.

No. 2
Fludarabine as second-line therapy for B cell chronic lymphocytic leukaemia: a technology assessment.
By Hyde C, Wake B, Bryan S, Barton P, Fry-Smith A, Davenport C, *et al*.

No. 3
Rituximab as third-line treatment for refractory or recurrent Stage III or IV follicular non-Hodgkin's lymphoma: a systematic review and economic evaluation.
By Wake B, Hyde C, Bryan S, Barton P, Song F, Fry-Smith A, *et al*.

No. 4
A systematic review of discharge arrangements for older people.
By Parker SG, Peet SM, McPherson A, Cannaby AM, Baker R, Wilson A, *et al*.

No. 5
The clinical effectiveness and cost-effectiveness of inhaler devices used in the routine management of chronic asthma in older children: a systematic review and economic evaluation.
By Peters J, Stevenson M, Beverley C, Lim J, Smith S.

No. 6
The clinical effectiveness and cost-effectiveness of sibutramine in the management of obesity: a technology assessment.
By O'Meara S, Riemsma R, Shirran L, Mather L, ter Riet G.

No. 7
The cost-effectiveness of magnetic resonance angiography for carotid artery stenosis and peripheral vascular disease: a systematic review.
By Berry E, Kelly S, Westwood ME, Davies LM, Gough MJ, Bamford JM, *et al*.

No. 8
Promoting physical activity in South Asian Muslim women through 'exercise on prescription'.
By Carroll B, Ali N, Azam N.

No. 9
Zanamivir for the treatment of influenza in adults: a systematic review and economic evaluation.
By Burls A, Clark W, Stewart T, Preston C, Bryan S, Jefferson T, *et al*.

No. 10
A review of the natural history and epidemiology of multiple sclerosis: implications for resource allocation and health economic models.
By Richards RG, Sampson FC, Beard SM, Tappenden P.

No. 11
Screening for gestational diabetes: a systematic review and economic evaluation.
By Scott DA, Loveman E, McIntyre L, Waugh N.

No. 12
The clinical effectiveness and cost-effectiveness of surgery for people with morbid obesity: a systematic review and economic evaluation.
By Clegg AJ, Colquitt J, Sidhu MK, Royle P, Loveman E, Walker A.

No. 13
The clinical effectiveness of trastuzumab for breast cancer: a systematic review.
By Lewis R, Bagnall A-M, Forbes C, Shirran E, Duffy S, Kleijnen J, *et al*.

No. 14
The clinical effectiveness and cost-effectiveness of vinorelbine for breast cancer: a systematic review and economic evaluation.
By Lewis R, Bagnall A-M, King S, Woolacott N, Forbes C, Shirran L, *et al*.

No. 7
The clinical effectiveness and cost-effectiveness of routine dental checks: a systematic review and economic evaluation.
By Davenport C, Elley K, Salas C, Taylor-Weetman CL, Fry-Smith A, Bryan S, *et al*.

No. 8
A multicentre randomised controlled trial assessing the costs and benefits of using structured information and analysis of women's preferences in the management of menorrhagia.
By Kennedy ADM, Sculpher MJ, Coulter A, Dwyer N, Rees M, Horsley S, *et al*.

No. 9
Clinical effectiveness and cost–utility of photodynamic therapy for wet age-related macular degeneration: a systematic review and economic evaluation.
By Meads C, Salas C, Roberts T, Moore D, Fry-Smith A, Hyde C.

No. 10
Evaluation of molecular tests for prenatal diagnosis of chromosome abnormalities.
By Grimshaw GM, Szczepura A, Hultén M, MacDonald F, Nevin NC, Sutton F, *et al*.

No. 11
First and second trimester antenatal screening for Down's syndrome: the results of the Serum, Urine and Ultrasound Screening Study (SURUSS).
By Wald NJ, Rodeck C, Hackshaw AK, Walters J, Chitty L, Mackinson AM.

No. 12
The effectiveness and cost-effectiveness of ultrasound locating devices for central venous access: a systematic review and economic evaluation.
By Calvert N, Hind D, McWilliams RG, Thomas SM, Beverley C, Davidson A.

No. 13
A systematic review of atypical antipsychotics in schizophrenia.
By Bagnall A-M, Jones L, Lewis R, Ginnelly L, Glanville J, Torgerson D, *et al*.

No. 14
Prostate Testing for Cancer and Treatment (ProtecT) feasibility study.
By Donovan J, Hamdy F, Neal D, Peters T, Oliver S, Brindle L, *et al*.

No. 15
Early thrombolysis for the treatment of acute myocardial infarction: a systematic review and economic evaluation.
By Boland A, Dundar Y, Bagust A, Haycox A, Hill R, Mujica Mota R, *et al*.

No. 16
Screening for fragile X syndrome: a literature review and modelling.
By Song FJ, Barton P, Sleightholme V, Yao GL, Fry-Smith A.

No. 17
Systematic review of endoscopic sinus surgery for nasal polyps.
By Dalziel K, Stein K, Round A, Garside R, Royle P.

No. 18
Towards efficient guidelines: how to monitor guideline use in primary care.
By Hutchinson A, McIntosh A, Cox S, Gilbert C.

No. 19
Effectiveness and cost-effectiveness of acute hospital-based spinal cord injuries services: systematic review.
By Bagnall A-M, Jones L, Richardson G, Duffy S, Riemsma R.

No. 20
Prioritisation of health technology assessment. The PATHS model: methods and case studies.
By Townsend J, Buxton M, Harper G.

No. 21
Systematic review of the clinical effectiveness and cost-effectiveness of tension-free vaginal tape for treatment of urinary stress incontinence.
By Cody J, Wyness L, Wallace S, Glazener C, Kilonzo M, Stearns S, *et al*.

No. 22
The clinical and cost-effectiveness of patient education models for diabetes: a systematic review and economic evaluation.
By Loveman E, Cave C, Green C, Royle P, Dunn N, Waugh N.

No. 23
The role of modelling in prioritising and planning clinical trials.
By Chilcott J, Brennan A, Booth A, Karnon J, Tappenden P.

No. 24
Cost–benefit evaluation of routine influenza immunisation in people 65–74 years of age.
By Allsup S, Gosney M, Haycox A, Regan M.

No. 25
The clinical and cost-effectiveness of pulsatile machine perfusion versus cold storage of kidneys for transplantation retrieved from heart-beating and non-heart-beating donors.
By Wight J, Chilcott J, Holmes M, Brewer N.

No. 26
Can randomised trials rely on existing electronic data? A feasibility study to explore the value of routine data in health technology assessment.
By Williams JG, Cheung WY, Cohen DR, Hutchings HA, Longo MF, Russell IT.

No. 27
Evaluating non-randomised intervention studies.
By Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovitch C, Song F, *et al*.

No. 28
A randomised controlled trial to assess the impact of a package comprising a patient-orientated, evidence-based self-help guidebook and patient-centred consultations on disease management and satisfaction in inflammatory bowel disease.
By Kennedy A, Nelson E, Reeves D, Richardson G, Roberts C, Robinson A, *et al*.

No. 29
The effectiveness of diagnostic tests for the assessment of shoulder pain due to soft tissue disorders: a systematic review.
By Dinnes J, Loveman E, McIntyre L, Waugh N.

No. 30
The value of digital imaging in diabetic retinopathy.
By Sharp PF, Olson J, Strachan F, Hipwell J, Ludbrook A, O'Donnell M, *et al*.

No. 31
Lowering blood pressure to prevent myocardial infarction and stroke: a new preventive strategy.
By Law M, Wald N, Morris J.

No. 32
Clinical and cost-effectiveness of capecitabine and tegafur with uracil for the treatment of metastatic colorectal cancer: systematic review and economic evaluation.
By Ward S, Kaltenthaler E, Cowan J, Brewer N.

No. 33
Clinical and cost-effectiveness of new and emerging technologies for early localised prostate cancer: a systematic review.
By Hummel S, Paisley S, Morgan A, Currie E, Brewer N.

No. 34
Literature searching for clinical and cost-effectiveness studies used in health technology assessment reports carried out for the National Institute for Clinical Excellence appraisal system.
By Royle P, Waugh N.

No. 19
A rapid and systematic review and economic evaluation of the clinical and cost-effectiveness of newer drugs for treatment of mania associated with bipolar affective disorder.
By Bridle C, Palmer S, Bagnall A-M, Darba J, Duffy S, Sculpher M, *et al.*

No. 20
Liquid-based cytology in cervical screening: an updated rapid and systematic review and economic analysis.
By Karnon J, Peters J, Platt J, Chilcott J, McGoogan E, Brewer N.

No. 21
Systematic review of the long-term effects and economic consequences of treatments for obesity and implications for health improvement.
By Avenell A, Broom J, Brown TJ, Poobalan A, Aucott L, Stearns SC, *et al.*

No. 22
Autoantibody testing in children with newly diagnosed type 1 diabetes mellitus.
By Dretzke J, Cummins C, Sandercock J, Fry-Smith A, Barrett T, Burls A.

No. 23
Clinical effectiveness and cost-effectiveness of prehospital intravenous fluids in trauma patients.
By Dretzke J, Sandercock J, Bayliss S, Burls A.

No. 24
Newer hypnotic drugs for the short-term management of insomnia: a systematic review and economic evaluation.
By Dündar Y, Boland A, Strobl J, Dodd S, Haycox A, Bagust A, *et al.*

No. 25
Development and validation of methods for assessing the quality of diagnostic accuracy studies.
By Whiting P, Rutjes AWS, Dinnes J, Reitsma JB, Bossuyt PMM, Kleijnen J.

No. 26
EVALUATE hysterectomy trial: a multicentre randomised trial comparing abdominal, vaginal and laparoscopic methods of hysterectomy.
By Garry R, Fountain J, Brown J, Manca A, Mason S, Sculpher M, *et al.*

No. 27
Methods for expected value of information analysis in complex health economic models: developments on the health economics of interferon-β and glatiramer acetate for multiple sclerosis.
By Tappenden P, Chilcott JB, Eggington S, Oakley J, McCabe C.

No. 28
Effectiveness and cost-effectiveness of imatinib for first-line treatment of chronic myeloid leukaemia in chronic phase: a systematic review and economic analysis.
By Dalziel K, Round A, Stein K, Garside R, Price A.

No. 29
VenUS I: a randomised controlled trial of two types of bandage for treating venous leg ulcers.
By Iglesias C, Nelson EA, Cullum NA, Torgerson DJ, on behalf of the VenUS Team.

No. 30
Systematic review of the effectiveness and cost-effectiveness, and economic evaluation, of myocardial perfusion scintigraphy for the diagnosis and management of angina and myocardial infarction.
By Mowatt G, Vale L, Brazzelli M, Hernandez R, Murray A, Scott N, *et al.*

No. 31
A pilot study on the use of decision theory and value of information analysis as part of the NHS Health Technology Assessment programme.
By Claxton K, Ginnelly L, Sculpher M, Philips Z, Palmer S.

No. 32
The Social Support and Family Health Study: a randomised controlled trial and economic evaluation of two alternative forms of postnatal support for mothers living in disadvantaged inner-city areas.
By Wiggins M, Oakley A, Roberts I, Turner H, Rajan L, Austerberry H, *et al.*

No. 33
Psychosocial aspects of genetic screening of pregnant women and newborns: a systematic review.
By Green JM, Hewison J, Bekker HL, Bryant, Cuckle HS.

No. 34
Evaluation of abnormal uterine bleeding: comparison of three outpatient procedures within cohorts defined by age and menopausal status.
By Critchley HOD, Warner P, Lee AJ, Brechin S, Guise J, Graham B.

No. 35
Coronary artery stents: a rapid systematic review and economic evaluation.
By Hill R, Bagust A, Bakhai A, Dickson R, Dündar Y, Haycox A, *et al.*

No. 36
Review of guidelines for good practice in decision-analytic modelling in health technology assessment.
By Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, *et al.*

No. 37
Rituximab (MabThera®) for aggressive non-Hodgkin's lymphoma: systematic review and economic evaluation.
By Knight C, Hind D, Brewer N, Abbott V.

No. 38
Clinical effectiveness and cost-effectiveness of clopidogrel and modified-release dipyridamole in the secondary prevention of occlusive vascular events: a systematic review and economic evaluation.
By Jones L, Griffin S, Palmer S, Main C, Orton V, Sculpher M, *et al.*

No. 39
Pegylated interferon α-2a and -2b in combination with ribavirin in the treatment of chronic hepatitis C: a systematic review and economic evaluation.
By Shepherd J, Brodin H, Cave C, Waugh N, Price A, Gabbay J.

No. 40
Clopidogrel used in combination with aspirin compared with aspirin alone in the treatment of non-ST-segment-elevation acute coronary syndromes: a systematic review and economic evaluation.
By Main C, Palmer S, Griffin S, Jones L, Orton V, Sculpher M, *et al.*

No. 41
Provision, uptake and cost of cardiac rehabilitation programmes: improving services to under-represented groups.
By Beswick AD, Rees K, Griebsch I, Taylor FC, Burke M, West RR, *et al.*

No. 42
Involving South Asian patients in clinical trials.
By Hussain-Gambles M, Leese B, Atkin K, Brown J, Mason S, Tovey P.

No. 43
Clinical and cost-effectiveness of continuous subcutaneous insulin infusion for diabetes.
By Colquitt JL, Green C, Sidhu MK, Hartwell D, Waugh N.

No. 44
Identification and assessment of ongoing trials in health technology assessment reviews.
By Song FJ, Fry-Smith A, Davenport C, Bayliss S, Adi Y, Wilson JS, *et al.*

No. 45
Systematic review and economic evaluation of a long-acting insulin analogue, insulin glargine
By Warren E, Weatherley-Jones E, Chilcott J, Beverley C.

203

No. 46
Supplementation of a home-based exercise programme with a class-based programme for people with osteoarthritis of the knees: a randomised controlled trial and health economic analysis.

By McCarthy CJ, Mills PM, Pullen R, Richardson G, Hawkins N, Roberts CR, *et al*.

No. 47
Clinical and cost-effectiveness of once-daily versus more frequent use of same potency topical corticosteroids for atopic eczema: a systematic review and economic evaluation.

By Green C, Colquitt JL, Kirby J, Davidson P, Payne E.

No. 48
Acupuncture of chronic headache disorders in primary care: randomised controlled trial and economic analysis.

By Vickers AJ, Rees RW, Zollman CE, McCarney R, Smith CM, Ellis N, *et al*.

No. 49
Generalisability in economic evaluation studies in healthcare: a review and case studies.

By Sculpher MJ, Pang FS, Manca A, Drummond MF, Golder S, Urdahl H, *et al*.

No. 50
Virtual outreach: a randomised controlled trial and economic evaluation of joint teleconferenced medical consultations.

By Wallace P, Barber J, Clayton W, Currell R, Fleming K, Garner P, *et al*.

Volume 9, 2005

No. 1
Randomised controlled multiple treatment comparison to provide a cost-effectiveness rationale for the selection of antimicrobial therapy in acne.

By Ozolins M, Eady EA, Avery A, Cunliffe WJ, O'Neill C, Simpson NB, *et al*.

No. 2
Do the findings of case series studies vary significantly according to methodological characteristics?

By Dalziel K, Round A, Stein K, Garside R, Castelnuovo E, Payne L.

No. 3
Improving the referral process for familial breast cancer genetic counselling: findings of three randomised controlled trials of two interventions.

By Wilson BJ, Torrance N, Mollison J, Wordsworth S, Gray JR, Haites NE, *et al*.

No. 4
Randomised evaluation of alternative electrosurgical modalities to treat bladder outflow obstruction in men with benign prostatic hyperplasia.

By Fowler C, McAllister W, Plail R, Karim O, Yang Q.

No. 5
A pragmatic randomised controlled trial of the cost-effectiveness of palliative therapies for patients with inoperable oesophageal cancer.

By Shenfine J, McNamee P, Steen N, Bond J, Griffin SM.

No. 6
Impact of computer-aided detection prompts on the sensitivity and specificity of screening mammography.

By Taylor P, Champness J, Given-Wilson R, Johnston K, Potts H.

No. 7
Issues in data monitoring and interim analysis of trials.

By Grant AM, Altman DG, Babiker AB, Campbell MK, Clemens FJ, Darbyshire JH, *et al*.

No. 8
Lay public's understanding of equipoise and randomisation in randomised controlled trials.

By Robinson EJ, Kerr CEP, Stevens AJ, Lilford RJ, Braunholtz DA, Edwards SJ, *et al*.

No. 9
Clinical and cost-effectiveness of electroconvulsive therapy for depressive illness, schizophrenia, catatonia and mania: systematic reviews and economic modelling studies.

By Greenhalgh J, Knight C, Hind D, Beverley C, Walters S.

No. 10
Measurement of health-related quality of life for people with dementia: development of a new instrument (DEMQOL) and an evaluation of current methodology.

By Smith SC, Lamping DL, Banerjee S, Harwood R, Foley B, Smith P, *et al*.

No. 11
Clinical effectiveness and cost-effectiveness of drotrecogin alfa (activated) (Xigris®) for the treatment of severe sepsis in adults: a systematic review and economic evaluation.

By Green C, Dinnes J, Takeda A, Shepherd J, Hartwell D, Cave C, *et al*.

No. 12
A methodological review of how heterogeneity has been examined in systematic reviews of diagnostic test accuracy.

By Dinnes J, Deeks J, Kirby J, Roderick P.

No. 13
Cervical screening programmes: can automation help? Evidence from systematic reviews, an economic analysis and a simulation modelling exercise applied to the UK.

By Willis BH, Barton P, Pearmain P, Bryan S, Hyde C.

No. 14
Laparoscopic surgery for inguinal hernia repair: systematic review of effectiveness and economic evaluation.

By McCormack K, Wake B, Perez J, Fraser C, Cook J, McIntosh E, *et al*.

No. 15
Clinical effectiveness, tolerability and cost-effectiveness of newer drugs for epilepsy in adults: a systematic review and economic evaluation.

By Wilby J, Kainth A, Hawkins N, Epstein D, McIntosh H, McDaid C, *et al*.

No. 16
A randomised controlled trial to compare the cost-effectiveness of tricyclic antidepressants, selective serotonin reuptake inhibitors and lofepramine.

By Peveler R, Kendrick T, Buxton M, Longworth L, Baldwin D, Moore M, *et al*.

No. 17
Clinical effectiveness and cost-effectiveness of immediate angioplasty for acute myocardial infarction: systematic review and economic evaluation.

By Hartwell D, Colquitt J, Loveman E, Clegg AJ, Brodin H, Waugh N, *et al*.

No. 18
A randomised controlled comparison of alternative strategies in stroke care.

By Kalra L, Evans A, Perez I, Knapp M, Swift C, Donaldson N.

No. 19
The investigation and analysis of critical incidents and adverse events in healthcare.

By Woloshynowych M, Rogers S, Taylor-Adams S, Vincent C.

No. 20
Potential use of routine databases in health technology assessment.

By Raftery J, Roderick P, Stevens A.

No. 21
Clinical and cost-effectiveness of newer immunosuppressive regimens in renal transplantation: a systematic review and modelling study.

By Woodroffe R, Yao GL, Meads C, Bayliss S, Ready A, Raftery J, *et al*.

No. 22
A systematic review and economic evaluation of alendronate, etidronate, risedronate, raloxifene and teriparatide for the prevention and treatment of postmenopausal osteoporosis.

By Stevenson M, Lloyd Jones M, De Nigris E, Brewer N, Davis S, Oakley J.

No. 23

A systematic review to examine the impact of psycho-educational interventions on health outcomes and costs in adults and children with difficult asthma.

By Smith JR, Mugford M, Holland R, Candy B, Noble MJ, Harrison BDW, *et al*.

No. 24

An evaluation of the costs, effectiveness and quality of renal replacement therapy provision in renal satellite units in England and Wales.

By Roderick P, Nicholson T, Armitage A, Mehta R, Mullee M, Gerard K, *et al*.

No. 25

Imatinib for the treatment of patients with unresectable and/or metastatic gastrointestinal stromal tumours: systematic review and economic evaluation.

By Wilson J, Connock M, Song F, Yao G, Fry-Smith A, Raftery J, *et al*.

No. 26

Indirect comparisons of competing interventions.

By Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R, *et al*.

No. 27

Cost-effectiveness of alternative strategies for the initial medical management of non-ST elevation acute coronary syndrome: systematic review and decision-analytical modelling.

By Robinson M, Palmer S, Sculpher M, Philips Z, Ginnelly L, Bowens A, *et al*.

No. 28

Outcomes of electrically stimulated gracilis neosphincter surgery.

By Tillin T, Chambers M, Feldman R.

No. 29

The effectiveness and cost-effectiveness of pimecrolimus and tacrolimus for atopic eczema: a systematic review and economic evaluation.

By Garside R, Stein K, Castelnuovo E, Pitt M, Ashcroft D, Dimmock P, *et al*.

No. 30

Systematic review on urine albumin testing for early detection of diabetic complications.

By Newman DJ, Mattock MB, Dawnay ABS, Kerry S, McGuire A, Yaqoob M, *et al*.

No. 31

Randomised controlled trial of the cost-effectiveness of water-based therapy for lower limb osteoarthritis.

By Cochrane T, Davey RC, Matthes Edwards SM.

No. 32

Longer term clinical and economic benefits of offering acupuncture care to patients with chronic low back pain.

By Thomas KJ, MacPherson H, Ratcliffe J, Thorpe L, Brazier J, Campbell M, *et al*.

No. 33

Cost-effectiveness and safety of epidural steroids in the management of sciatica.

By Price C, Arden N, Coglan L, Rogers P.

No. 34

The British Rheumatoid Outcome Study Group (BROSG) randomised controlled trial to compare the effectiveness and cost-effectiveness of aggressive versus symptomatic therapy in established rheumatoid arthritis.

By Symmons D, Tricker K, Roberts C, Davies L, Dawes P, Scott DL.

No. 35

Conceptual framework and systematic review of the effects of participants' and professionals' preferences in randomised controlled trials.

By King M, Nazareth I, Lampe F, Bower P, Chandler M, Morou M, *et al*.

No. 36

The clinical and cost-effectiveness of implantable cardioverter defibrillators: a systematic review.

By Bryant J, Brodin H, Loveman E, Payne E, Clegg A.

No. 37

A trial of problem-solving by community mental health nurses for anxiety, depression and life difficulties among general practice patients. The CPN-GP study.

By Kendrick T, Simons L, Mynors-Wallis L, Gray A, Lathlean J, Pickering R, *et al*.

No. 38

The causes and effects of socio-demographic exclusions from clinical trials.

By Bartlett C, Doyal L, Ebrahim S, Davey P, Bachmann M, Egger M, *et al*.

No. 39

Is hydrotherapy cost-effective? A randomised controlled trial of combined hydrotherapy programmes compared with physiotherapy land techniques in children with juvenile idiopathic arthritis.

By Epps H, Ginnelly L, Utley M, Southwood T, Gallivan S, Sculpher M, *et al*.

No. 40

A randomised controlled trial and cost-effectiveness study of systematic screening (targeted and total population screening) versus routine practice for the detection of atrial fibrillation in people aged 65 and over. The SAFE study.

By Hobbs FDR, Fitzmaurice DA, Mant J, Murray E, Jowett S, Bryan S, *et al*.

No. 41

Displaced intracapsular hip fractures in fit, older people: a randomised comparison of reduction and fixation, bipolar hemiarthroplasty and total hip arthroplasty.

By Keating JF, Grant A, Masson M, Scott NW, Forbes JF.

No. 42

Long-term outcome of cognitive behaviour therapy clinical trials in central Scotland.

By Durham RC, Chambers JA, Power KG, Sharp DM, Macdonald RR, Major KA, *et al*.

No. 43

The effectiveness and cost-effectiveness of dual-chamber pacemakers compared with single-chamber pacemakers for bradycardia due to atrioventricular block or sick sinus syndrome: systematic review and economic evaluation.

By Castelnuovo E, Stein K, Pitt M, Garside R, Payne E.

No. 44

Newborn screening for congenital heart defects: a systematic review and cost-effectiveness analysis.

By Knowles R, Griebsch I, Dezateux C, Brown J, Bull C, Wren C.

No. 45

The clinical and cost-effectiveness of left ventricular assist devices for end-stage heart failure: a systematic review and economic evaluation.

By Clegg AJ, Scott DA, Loveman E, Colquitt J, Hutchinson J, Royle P, *et al*.

No. 46

The effectiveness of the Heidelberg Retina Tomograph and laser diagnostic glaucoma scanning system (GDx) in detecting and monitoring glaucoma.

By Kwartz AJ, Henson DB, Harper RA, Spencer AF, McLeod D.

No. 47

Clinical and cost-effectiveness of autologous chondrocyte implantation for cartilage defects in knee joints: systematic review and economic evaluation.

By Clar C, Cummins E, McIntyre L, Thomas S, Lamb J, Bain L, *et al*.

205

No. 48
Systematic review of effectiveness of different treatments for childhood retinoblastoma.
By McDaid C, Hartley S, Bagnall A-M, Ritchie G, Light K, Riemsma R.

No. 49
Towards evidence-based guidelines for the prevention of venous thromboembolism: systematic reviews of mechanical methods, oral anticoagulation, dextran and regional anaesthesia as thromboprophylaxis.
By Roderick P, Ferris G, Wilson K, Halls H, Jackson D, Collins R, *et al*.

No. 50
The effectiveness and cost-effectiveness of parent training/education programmes for the treatment of conduct disorder, including oppositional defiant disorder, in children.
By Dretzke J, Frew E, Davenport C, Barlow J, Stewart-Brown S, Sandercock J, *et al*.

Volume 10, 2006

No. 1
The clinical and cost-effectiveness of donepezil, rivastigmine, galantamine and memantine for Alzheimer's disease.
By Loveman E, Green C, Kirby J, Takeda A, Picot J, Payne E, *et al*.

No. 2
FOOD: a multicentre randomised trial evaluating feeding policies in patients admitted to hospital with a recent stroke.
By Dennis M, Lewis S, Cranswick G, Forbes J.

No. 3
The clinical effectiveness and cost-effectiveness of computed tomography screening for lung cancer: systematic reviews.
By Black C, Bagust A, Boland A, Walker S, McLeod C, De Verteuil R, *et al*.

No. 4
A systematic review of the effectiveness and cost-effectiveness of neuroimaging assessments used to visualise the seizure focus in people with refractory epilepsy being considered for surgery.
By Whiting P, Gupta R, Burch J, Mujica Mota RE, Wright K, Marson A, *et al*.

No. 5
Comparison of conference abstracts and presentations with full-text articles in the health technology assessments of rapidly evolving technologies.
By Dundar Y, Dodd S, Dickson R, Walley T, Haycox A, Williamson PR.

No. 6
Systematic review and evaluation of methods of assessing urinary incontinence.
By Martin JL, Williams KS, Abrams KR, Turner DA, Sutton AJ, Chapple C, *et al*.

No. 7
The clinical effectiveness and cost-effectiveness of newer drugs for children with epilepsy. A systematic review.
By Connock M, Frew E, Evans B-W, Bryan S, Cummins C, Fry-Smith A, *et al*.

No. 8
Surveillance of Barrett's oesophagus: exploring the uncertainty through systematic review, expert workshop and economic modelling.
By Garside R, Pitt M, Somerville M, Stein K, Price A, Gilbert N.

No. 9
Topotecan, pegylated liposomal doxorubicin hydrochloride and paclitaxel for second-line or subsequent treatment of advanced ovarian cancer: a systematic review and economic evaluation.
By Main C, Bojke L, Griffin S, Norman G, Barbieri M, Mather L, *et al*.

No. 10
Evaluation of molecular techniques in prediction and diagnosis of cytomegalovirus disease in immunocompromised patients.
By Szczepura A, Westmoreland D, Vinogradova Y, Fox J, Clark M.

No. 11
Screening for thrombophilia in high-risk situations: systematic review and cost-effectiveness analysis. The Thrombosis: Risk and Economic Assessment of Thrombophilia Screening (TREATS) study.
By Wu O, Robertson L, Twaddle S, Lowe GDO, Clark P, Greaves M, *et al*.

No. 12
A series of systematic reviews to inform a decision analysis for sampling and treating infected diabetic foot ulcers.
By Nelson EA, O'Meara S, Craig D, Iglesias C, Golder S, Dalton J, *et al*.

No. 13
Randomised clinical trial, observational study and assessment of cost-effectiveness of the treatment of varicose veins (REACTIV trial).
By Michaels JA, Campbell WB, Brazier JE, MacIntyre JB, Palfreyman SJ, Ratcliffe J, *et al*.

No. 14
The cost-effectiveness of screening for oral cancer in primary care.
By Speight PM, Palmer S, Moles DR, Downer MC, Smith DH, Henriksson M, *et al*.

No. 15
Measurement of the clinical and cost-effectiveness of non-invasive diagnostic testing strategies for deep vein thrombosis.
By Goodacre S, Sampson F, Stevenson M, Wailoo A, Sutton A, Thomas S, *et al*.

No. 16
Systematic review of the effectiveness and cost-effectiveness of HealOzone® for the treatment of occlusal pit/fissure caries and root caries.
By Brazzelli M, McKenzie L, Fielding S, Fraser C, Clarkson J, Kilonzo M, *et al*.

No. 17
Randomised controlled trials of conventional antipsychotic versus new atypical drugs, and new atypical drugs versus clozapine, in people with schizophrenia responding poorly to, or intolerant of, current drug treatment.
By Lewis SW, Davies L, Jones PB, Barnes TRE, Murray RM, Kerwin R, *et al*.

No. 18
Diagnostic tests and algorithms used in the investigation of haematuria: systematic reviews and economic evaluation.
By Rodgers M, Nixon J, Hempel S, Aho T, Kelly J, Neal D, *et al*.

No. 19
Cognitive behavioural therapy in addition to antispasmodic therapy for irritable bowel syndrome in primary care: randomised controlled trial.
By Kennedy TM, Chalder T, McCrone P, Darnley S, Knapp M, Jones RH, *et al*.

No. 20
A systematic review of the clinical effectiveness and cost-effectiveness of enzyme replacement therapies for Fabry's disease and mucopolysaccharidosis type 1.
By Connock M, Juarez-Garcia A, Frew E, Mans A, Dretzke J, Fry-Smith A, *et al*.

No. 21
Health benefits of antiviral therapy for mild chronic hepatitis C: randomised controlled trial and economic evaluation.
By Wright M, Grieve R, Roberts J, Main J, Thomas HC, on behalf of the UK Mild Hepatitis C Trial Investigators.

No. 22
Pressure relieving support surfaces: a randomised evaluation.
By Nixon J, Nelson EA, Cranny G, Iglesias CP, Hawkins K, Cullum NA, *et al*.

No. 23
A systematic review and economic model of the effectiveness and cost-effectiveness of methylphenidate, dexamfetamine and atomoxetine for the treatment of attention deficit hyperactivity disorder in children and adolescents.
By King S, Griffin S, Hodges Z, Weatherly H, Asseburg C, Richardson G, *et al.*

No. 24
The clinical effectiveness and cost-effectiveness of enzyme replacement therapy for Gaucher's disease: a systematic review.
By Connock M, Burls A, Frew E, Fry-Smith A, Juarez-Garcia A, McCabe C, *et al.*

No. 25
Effectiveness and cost-effectiveness of salicylic acid and cryotherapy for cutaneous warts. An economic decision model.
By Thomas KS, Keogh-Brown MR, Chalmers JR, Fordham RJ, Holland RC, Armstrong SJ, *et al.*

No. 26
A systematic literature review of the effectiveness of non-pharmacological interventions to prevent wandering in dementia and evaluation of the ethical implications and acceptability of their use.
By Robinson L, Hutchings D, Corner L, Beyer F, Dickinson H, Vanoli A, *et al.*

No. 27
A review of the evidence on the effects and costs of implantable cardioverter defibrillator therapy in different patient groups, and modelling of cost-effectiveness and cost–utility for these groups in a UK context.
By Buxton M, Caine N, Chase D, Connelly D, Grace A, Jackson C, *et al.*

No. 28
Adefovir dipivoxil and pegylated interferon alfa-2a for the treatment of chronic hepatitis B: a systematic review and economic evaluation.
By Shepherd J, Jones J, Takeda A, Davidson P, Price A.

No. 29
An evaluation of the clinical and cost-effectiveness of pulmonary artery catheters in patient management in intensive care: a systematic review and a randomised controlled trial.
By Harvey S, Stevens K, Harrison D, Young D, Brampton W, McCabe C, *et al.*

No. 30
Accurate, practical and cost-effective assessment of carotid stenosis in the UK.
By Wardlaw JM, Chappell FM, Stevenson M, De Nigris E, Thomas S, Gillard J, *et al.*

No. 31
Etanercept and infliximab for the treatment of psoriatic arthritis: a systematic review and economic evaluation.
By Woolacott N, Bravo Vergel Y, Hawkins N, Kainth A, Khadjesari Z, Misso K, *et al.*

No. 32
The cost-effectiveness of testing for hepatitis C in former injecting drug users.
By Castelnuovo E, Thompson-Coon J, Pitt M, Cramp M, Siebert U, Price A, *et al.*

No. 33
Computerised cognitive behaviour therapy for depression and anxiety update: a systematic review and economic evaluation.
By Kaltenthaler E, Brazier J, De Nigris E, Tumur I, Ferriter M, Beverley C, *et al.*

No. 34
Cost-effectiveness of using prognostic information to select women with breast cancer for adjuvant systemic therapy.
By Williams C, Brunskill S, Altman D, Briggs A, Campbell H, Clarke M, *et al.*

No. 35
Psychological therapies including dialectical behaviour therapy for borderline personality disorder: a systematic review and preliminary economic evaluation.
By Brazier J, Tumur I, Holmes M, Ferriter M, Parry G, Dent-Brown K, *et al.*

No. 36
Clinical effectiveness and cost-effectiveness of tests for the diagnosis and investigation of urinary tract infection in children: a systematic review and economic model.
By Whiting P, Westwood M, Bojke L, Palmer S, Richardson G, Cooper J, *et al.*

No. 37
Cognitive behavioural therapy in chronic fatigue syndrome: a randomised controlled trial of an outpatient group programme.
By O'Dowd H, Gladwell P, Rogers CA, Hollinghurst S, Gregory A.

No. 38
A comparison of the cost-effectiveness of five strategies for the prevention of nonsteroidal anti-inflammatory drug-induced gastrointestinal toxicity: a systematic review with economic modelling.
By Brown TJ, Hooper L, Elliott RA, Payne K, Webb R, Roberts C, et al.

No. 39
The effectiveness and cost-effectiveness of computed tomography screening for coronary artery disease: systematic review.
By Waugh N, Black C, Walker S, McIntyre L, Cummins E, Hillis G.

No. 40
What are the clinical outcome and cost-effectiveness of endoscopy undertaken by nurses when compared with doctors? A Multi-Institution Nurse Endoscopy Trial (MINuET).
By Williams J, Russell I, Durai D, Cheung W-Y, Farrin A, Bloor K, *et al.*

No. 41
The clinical and cost-effectiveness of oxaliplatin and capecitabine for the adjuvant treatment of colon cancer: systematic review and economic evaluation.
By Pandor A, Eggington S, Paisley S, Tappenden P, Sutcliffe P.

No. 42
A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness.
By Chen Y-F, Jobanputra P, Barton P, Jowett S, Bryan S, Clark W, *et al.*

No. 43
Telemedicine in dermatology: a randomised controlled trial.
By Bowns IR, Collins K, Walters SJ, McDonagh AJG.

No. 44
Cost-effectiveness of cell salvage and alternative methods of minimising perioperative allogeneic blood transfusion: a systematic review and economic model.
By Davies L, Brown TJ, Haynes S, Payne K, Elliott RA, McCollum C.

No. 45
Clinical effectiveness and cost-effectiveness of laparoscopic surgery for colorectal cancer: systematic reviews and economic evaluation.
By Murray A, Lourenco T, de Verteuil R, Hernandez R, Fraser C, McKinley A, *et al.*

No. 46
Etanercept and efalizumab for the treatment of psoriasis: a systematic review.
By Woolacott N, Hawkins N, Mason A, Kainth A, Khadjesari Z, Bravo Vergel Y, *et al.*

No. 47
Systematic reviews of clinical decision tools for acute abdominal pain.
By Liu JLY, Wyatt JC, Deeks JJ, Clamp S, Keen J, Verde P, *et al.*

No. 48
Evaluation of the ventricular assist device programme in the UK.
By Sharples L, Buxton M, Caine N, Cafferty F, Demiris N, Dyer M, *et al.*

No. 49
A systematic review and economic model of the clinical and cost-effectiveness of immunosuppressive therapy for renal transplantation in children.
By Yao G, Albon E, Adi Y, Milford D, Bayliss S, Ready A, *et al.*

No. 50
Amniocentesis results: investigation of anxiety. The ARIA trial.
By Hewison J, Nixon J, Fountain J, Cocks K, Jones C, Mason G, *et al.*

Volume 11, 2007

No. 1
Pemetrexed disodium for the treatment of malignant pleural mesothelioma: a systematic review and economic evaluation.
By Dundar Y, Bagust A, Dickson R, Dodd S, Green J, Haycox A, *et al.*

No. 2
A systematic review and economic model of the clinical effectiveness and cost-effectiveness of docetaxel in combination with prednisone or prednisolone for the treatment of hormone-refractory metastatic prostate cancer.
By Collins R, Fenwick E, Trowman R, Perard R, Norman G, Light K, *et al.*

No. 3
A systematic review of rapid diagnostic tests for the detection of tuberculosis infection.
By Dinnes J, Deeks J, Kunst H, Gibson A, Cummins E, Waugh N, *et al.*

No. 4
The clinical effectiveness and cost-effectiveness of strontium ranelate for the prevention of osteoporotic fragility fractures in postmenopausal women.
By Stevenson M, Davis S, Lloyd-Jones M, Beverley C.

No. 5
A systematic review of quantitative and qualitative research on the role and effectiveness of written information available to patients about individual medicines.
By Raynor DK, Blenkinsopp A, Knapp P, Grime J, Nicolson DJ, Pollock K, *et al.*

No. 6
Oral naltrexone as a treatment for relapse prevention in formerly opioid-dependent drug users: a systematic review and economic evaluation.
By Adi Y, Juarez-Garcia A, Wang D, Jowett S, Frew E, Day E, *et al.*

No. 7
Glucocorticoid-induced osteoporosis: a systematic review and cost–utility analysis.
By Kanis JA, Stevenson M, McCloskey EV, Davis S, Lloyd-Jones M.

No. 8
Epidemiological, social, diagnostic and economic evaluation of population screening for genital chlamydial infection.
By Low N, McCarthy A, Macleod J, Salisbury C, Campbell R, Roberts TE, *et al.*

No. 9
Methadone and buprenorphine for the management of opioid dependence: a systematic review and economic evaluation.
By Connock M, Juarez-Garcia A, Jowett S, Frew E, Liu Z, Taylor RJ, *et al.*

No. 10
Exercise Evaluation Randomised Trial (EXERT): a randomised trial comparing GP referral for leisure centre-based exercise, community-based walking and advice only.
By Isaacs AJ, Critchley JA, See Tai S, Buckingham K, Westley D, Harridge SDR, *et al.*

No. 11
Interferon alfa (pegylated and non-pegylated) and ribavirin for the treatment of mild chronic hepatitis C: a systematic review and economic evaluation.
By Shepherd J, Jones J, Hartwell D, Davidson P, Price A, Waugh N.

No. 12
Systematic review and economic evaluation of bevacizumab and cetuximab for the treatment of metastatic colorectal cancer.
By Tappenden P, Jones R, Paisley S, Carroll C.

No. 13
A systematic review and economic evaluation of epoetin alfa, epoetin beta and darbepoetin alfa in anaemia associated with cancer, especially that attributable to cancer treatment.
By Wilson J, Yao GL, Raftery J, Bohlius J, Brunskill S, Sandercock J, *et al.*

No. 14
A systematic review and economic evaluation of statins for the prevention of coronary events.
By Ward S, Lloyd Jones M, Pandor A, Holmes M, Ara R, Ryan A, *et al.*

No. 15
A systematic review of the effectiveness and cost-effectiveness of different models of community-based respite care for frail older people and their carers.
By Mason A, Weatherly H, Spilsbury K, Arksey H, Golder S, Adamson J, *et al.*

No. 16
Additional therapy for young children with spastic cerebral palsy: a randomised controlled trial.
By Weindling AM, Cunningham CC, Glenn SM, Edwards RT, Reeves DJ.

No. 17
Screening for type 2 diabetes: literature review and economic modelling.
By Waugh N, Scotland G, McNamee P, Gillett M, Brennan A, Goyder E, *et al.*

No. 18
The effectiveness and cost-effectiveness of cinacalcet for secondary hyperparathyroidism in end-stage renal disease patients on dialysis: a systematic review and economic evaluation.
By Garside R, Pitt M, Anderson R, Mealing S, Roome C, Snaith A, *et al.*

No. 19
The clinical effectiveness and cost-effectiveness of gemcitabine for metastatic breast cancer: a systematic review and economic evaluation.
By Takeda AL, Jones J, Loveman E, Tan SC, Clegg AJ.

No. 20
A systematic review of duplex ultrasound, magnetic resonance angiography and computed tomography angiography for the diagnosis and assessment of symptomatic, lower limb peripheral arterial disease.
By Collins R, Cranny G, Burch J, Aguiar-Ibáñez R, Craig D, Wright K, *et al.*

No. 21
The clinical effectiveness and cost-effectiveness of treatments for children with idiopathic steroid-resistant nephrotic syndrome: a systematic review.
By Colquitt JL, Kirby J, Green C, Cooper K, Trompeter RS.

No. 22
A systematic review of the routine monitoring of growth in children of primary school age to identify growth-related conditions.
By Fayter D, Nixon J, Hartley S, Rithalia A, Butler G, Rudolf M, *et al.*

No. 23
Systematic review of the effectiveness of preventing and treating *Staphylococcus aureus* carriage in reducing peritoneal catheter-related infections.
By McCormack K, Rabindranath K, Kilonzo M, Vale L, Fraser C, McIntyre L, *et al.*

No. 24
The clinical effectiveness and cost of repetitive transcranial magnetic stimulation versus electroconvulsive therapy in severe depression: a multicentre pragmatic randomised controlled trial and economic analysis.

By McLoughlin DM, Mogg A, Eranti S, Pluck G, Purvis R, Edwards D, *et al*.

No. 25
A randomised controlled trial and economic evaluation of direct versus indirect and individual versus group modes of speech and language therapy for children with primary language impairment.

By Boyle J, McCartney E, Forbes J, O'Hare A.

No. 26
Hormonal therapies for early breast cancer: systematic review and economic evaluation.

By Hind D, Ward S, De Nigris E, Simpson E, Carroll C, Wyld L.

No. 27
Cardioprotection against the toxic effects of anthracyclines given to children with cancer: a systematic review.

By Bryant J, Picot J, Levitt G, Sullivan I, Baxter L, Clegg A.

No. 28
Adalimumab, etanercept and infliximab for the treatment of ankylosing spondylitis: a systematic review and economic evaluation.

By McLeod C, Bagust A, Boland A, Dagenais P, Dickson R, Dundar Y, *et al*.

No. 29
Prenatal screening and treatment strategies to prevent group B streptococcal and other bacterial infections in early infancy: cost-effectiveness and expected value of information analyses.

By Colbourn T, Asseburg C, Bojke L, Philips Z, Claxton K, Ades AE, *et al*.

No. 30
Clinical effectiveness and cost-effectiveness of bone morphogenetic proteins in the non-healing of fractures and spinal fusion: a systematic review.

By Garrison KR, Donell S, Ryder J, Shemilt I, Mugford M, Harvey I, *et al*.

No. 31
A randomised controlled trial of postoperative radiotherapy following breast-conserving surgery in a minimum-risk older population. The PRIME trial.

By Prescott RJ, Kunkler IH, Williams LJ, King CC, Jack W, van der Pol M, *et al*.

No. 32
Current practice, accuracy, effectiveness and cost-effectiveness of the school entry hearing screen.

By Bamford J, Fortnum H, Bristow K, Smith J, Vamvakas G, Davies L, *et al*.

No. 33
The clinical effectiveness and cost-effectiveness of inhaled insulin in diabetes mellitus: a systematic review and economic evaluation.

By Black C, Cummins E, Royle P, Philip S, Waugh N.

No. 34
Surveillance of cirrhosis for hepatocellular carcinoma: systematic review and economic analysis.

By Thompson Coon J, Rogers G, Hewson P, Wright D, Anderson R, Cramp M, *et al*.

No. 35
The Birmingham Rehabilitation Uptake Maximisation Study (BRUM). Homebased compared with hospital-based cardiac rehabilitation in a multi-ethnic population: cost-effectiveness and patient adherence.

By Jolly K, Taylor R, Lip GYH, Greenfield S, Raftery J, Mant J, *et al*.

No. 36
A systematic review of the clinical, public health and cost-effectiveness of rapid diagnostic tests for the detection and identification of bacterial intestinal pathogens in faeces and food.

By Abubakar I, Irvine L, Aldus CF, Wyatt GM, Fordham R, Schelenz S, *et al*.

No. 37
A randomised controlled trial examining the longer-term outcomes of standard versus new antiepileptic drugs. The SANAD trial.

By Marson AG, Appleton R, Baker GA, Chadwick DW, Doughty J, Eaton B, *et al*.

No. 38
Clinical effectiveness and cost-effectiveness of different models of managing long-term oral anti-coagulation therapy: a systematic review and economic modelling.

By Connock M, Stevens C, Fry-Smith A, Jowett S, Fitzmaurice D, Moore D, *et al*.

No. 39
A systematic review and economic model of the clinical effectiveness and cost-effectiveness of interventions for preventing relapse in people with bipolar disorder.

By Soares-Weiser K, Bravo Vergel Y, Beynon S, Dunn G, Barbieri M, Duffy S, *et al*.

No. 40
Taxanes for the adjuvant treatment of early breast cancer: systematic review and economic evaluation.

By Ward S, Simpson E, Davis S, Hind D, Rees A, Wilkinson A.

No. 41
The clinical effectiveness and cost-effectiveness of screening for open angle glaucoma: a systematic review and economic evaluation.

By Burr JM, Mowatt G, Hernández R, Siddiqui MAR, Cook J, Lourenco T, *et al*.

No. 42
Acceptability, benefit and costs of early screening for hearing disability: a study of potential screening tests and models.

By Davis A, Smith P, Ferguson M, Stephens D, Gianopoulos I.

No. 43
Contamination in trials of educational interventions.

By Keogh-Brown MR, Bachmann MO, Shepstone L, Hewitt C, Howe A, Ramsay CR, *et al*.

No. 44
Overview of the clinical effectiveness of positron emission tomography imaging in selected cancers.

By Facey K, Bradbury I, Laking G, Payne E.

No. 45
The effectiveness and cost-effectiveness of carmustine implants and temozolomide for the treatment of newly diagnosed high-grade glioma: a systematic review and economic evaluation.

By Garside R, Pitt M, Anderson R, Rogers G, Dyer M, Mealing S, *et al*.

No. 46
Drug-eluting stents: a systematic review and economic evaluation.

By Hill RA, Boland A, Dickson R, Dündar Y, Haycox A, McLeod C, *et al*.

No. 47
The clinical effectiveness and cost-effectiveness of cardiac resynchronisation (biventricular pacing) for heart failure: systematic review and economic model.

By Fox M, Mealing S, Anderson R, Dean J, Stein K, Price A, *et al*.

No. 48
Recruitment to randomised trials: strategies for trial enrolment and participation study. The STEPS study.

By Campbell MK, Snowdon C, Francis D, Elbourne D, McDonald AM, Knight R, *et al*.

209

No. 10
Routine antenatal anti-D prophylaxis
for RhD-negative women: a systematic
review and economic evaluation.
By Pilgrim H, Lloyd-Jones M, Rees A.

No. 11
Amantadine, oseltamivir and zanamivir
for the prophylaxis of influenza
(including a review of existing guidance
no. 67): a systematic review and
economic evaluation.
By Tappenden P, Jackson R, Cooper
K, Rees A, Simpson E, Read R, *et al.*

No. 12
Improving the evaluation of
therapeutic interventions in multiple
sclerosis: the role of new psychometric
methods.
By Hobart J, Cano S.

No. 13
Treatment of severe ankle sprain: a
pragmatic randomised controlled trial
comparing the clinical effectiveness
and cost-effectiveness of three types of
mechanical ankle support with tubular
bandage. The CAST trial.
By Cooke MW, Marsh JL, Clark M,
Nakash R, Jarvis RM, Hutton JL, *et al.*,
on behalf of the CAST trial group.

No. 14
Non-occupational postexposure
prophylaxis for HIV: a systematic
review.
By Bryant J, Baxter L, Hird S.

No. 15
Blood glucose self-monitoring in type 2
diabetes: a randomised controlled trial.
By Farmer AJ, Wade AN, French DP,
Simon J, Yudkin P, Gray A, *et al.*

No. 16
How far does screening women for
domestic (partner) violence in different
health-care settings meet criteria for
a screening programme? Systematic
reviews of nine UK National Screening
Committee criteria.
By Feder G, Ramsay J, Dunne D,
Rose M, Arsene C, Norman R, *et al.*

No. 17
Spinal cord stimulation for chronic
pain of neuropathic or ischaemic
origin: systematic review and economic
evaluation.
By Simpson, EL, Duenas A, Holmes
MW, Papaioannou D, Chilcott J.

No. 18
The role of magnetic resonance
imaging in the identification of
suspected acoustic neuroma: a
systematic review of clinical and cost-
effectiveness and natural history.
By Fortnum H, O'Neill C, Taylor R,
Lenthall R, Nikolopoulos T, Lightfoot
G, *et al.*

No. 19
Dipsticks and diagnostic algorithms in
urinary tract infection: development
and validation, randomised trial,
economic analysis, observational cohort
and qualitative study.
By Little P, Turner S, Rumsby K,
Warner G, Moore M, Lowes JA, *et al.*

No. 20
Systematic review of respite care in the
frail elderly.
By Shaw C, McNamara R, Abrams
K, Cannings-John R, Hood K, Longo
M, *et al.*

No. 21
Neuroleptics in the treatment of
aggressive challenging behaviour for
people with intellectual disabilities:
a randomised controlled trial
(NACHBID).
By Tyrer P, Oliver-Africano P, Romeo
R, Knapp M, Dickens S, Bouras N, *et al.*

No. 22
Randomised controlled trial to
determine the clinical effectiveness
and cost-effectiveness of selective
serotonin reuptake inhibitors plus
supportive care, versus supportive care
alone, for mild to moderate depression
with somatic symptoms in primary
care: the THREAD (THREshold for
AntiDepressant response) study.
By Kendrick T, Chatwin J, Dowrick C,
Tylee A, Morriss R, Peveler R, *et al.*

No. 23
Diagnostic strategies using DNA testing
for hereditary haemochromatosis in
at-risk populations: a systematic review
and economic evaluation.
By Bryant J, Cooper K, Picot J, Clegg
A, Roderick P, Rosenberg W, *et al.*

No. 24
Enhanced external counterpulsation
for the treatment of stable angina and
heart failure: a systematic review and
economic analysis.
By McKenna C, McDaid C,
Suekarran S, Hawkins N, Claxton K,
Light K, *et al.*

No. 25
Development of a decision support
tool for primary care management of
patients with abnormal liver function
tests without clinically apparent liver
disease: a record-linkage population
cohort study and decision analysis
(ALFIE).
By Donnan PT, McLernon D, Dillon
JF, Ryder S, Roderick P, Sullivan F, *et al.*

No. 26
A systematic review of presumed
consent systems for deceased organ
donation.
By Rithalia A, McDaid C, Suekarran
S, Norman G, Myers L, Sowden A.

No. 27
Paracetamol and ibuprofen for the
treatment of fever in children: the
PITCH randomised controlled trial.
By Hay AD, Redmond NM, Costelloe
C, Montgomery AA, Fletcher M,
Hollinghurst S, *et al.*

No. 28
A randomised controlled trial to
compare minimally invasive glucose
monitoring devices with conventional
monitoring in the management of
insulin-treated diabetes mellitus
(MITRE).
By Newman SP, Cooke D, Casbard A,
Walker S, Meredith S, Nunn A, *et al.*

No. 29
Sensitivity analysis in economic
evaluation: an audit of NICE current
practice and a review of its use and
value in decision-making.
By Andronis L, Barton P, Bryan S.

Suppl. 1
Trastuzumab for the treatment of
primary breast cancer in HER2-positive
women: a single technology appraisal.
By Ward S, Pilgrim H, Hind D.

Docetaxel for the adjuvant treatment
of early node-positive breast cancer: a
single technology appraisal.
By Chilcott J, Lloyd Jones M,
Wilkinson A.

The use of paclitaxel in the
management of early stage breast
cancer.
By Griffin S, Dunn G, Palmer S,
Macfarlane K, Brent S, Dyker A, *et al.*

Rituximab for the first-line treatment
of stage III/IV follicular non-Hodgkin's
lymphoma.
By Dundar Y, Bagust A, Hounsome J,
McLeod C, Boland A, Davis H, *et al.*

Bortezomib for the treatment of
multiple myeloma patients.
By Green C, Bryant J, Takeda A,
Cooper K, Clegg A, Smith A, *et al.*

Fludarabine phosphate for the first-
line treatment of chronic lymphocytic
leukaemia.
By Walker S, Palmer S, Erhorn S,
Brent S, Dyker A, Ferrie L, *et al.*

Erlotinib for the treatment of relapsed
non-small cell lung cancer.
By McLeod C, Bagust A, Boland A,
Hockenhull J, Dundar Y, Proudlove C,
et al.

Cetuximab plus radiotherapy for the
treatment of locally advanced squamous
cell carcinoma of the head and neck.
By Griffin S, Walker S, Sculpher M,
White S, Erhorn S, Brent S, *et al.*

Infliximab for the treatment of adults
with psoriasis.
By Loveman E, Turner D, Hartwell
D, Cooper K, Clegg A.

No. 30
Psychological interventions for postnatal depression: cluster randomised trial and economic evaluation. The PoNDER trial.
By Morrell CJ, Warner R, Slade P, Dixon S, Walters S, Paley G, *et al.*

No. 31
The effect of different treatment durations of clopidogrel in patients with non-ST-segment elevation acute coronary syndromes: a systematic review and value of information analysis.
By Rogowski R, Burch J, Palmer S, Craigs C, Golder S, Woolacott N.

No. 32
Systematic review and individual patient data meta-analysis of diagnosis of heart failure, with modelling of implications of different diagnostic strategies in primary care.
By Mant J, Doust J, Roalfe A, Barton P, Cowie MR, Glasziou P, *et al.*

No. 33
A multicentre randomised controlled trial of the use of continuous positive airway pressure and non-invasive positive pressure ventilation in the early treatment of patients presenting to the emergency department with severe acute cardiogenic pulmonary oedema: the 3CPO trial.
By Gray AJ, Goodacre S, Newby DE, Masson MA, Sampson F, Dixon S, *et al.*, on behalf of the 3CPO study investigators.

No. 34
Early high-dose lipid-lowering therapy to avoid cardiac events: a systematic review and economic evaluation.
By Ara R, Pandor A, Stevens J, Rees A, Rafia R.

No. 35
Adefovir dipivoxil and pegylated interferon alpha for the treatment of chronic hepatitis B: an updated systematic review and economic evaluation.
By Jones J, Shepherd J, Baxter L, Gospodarevskaya E, Hartwell D, Harris P, *et al.*

No. 36
Methods to identify postnatal depression in primary care: an integrated evidence synthesis and value of information analysis.
By Hewitt CE, Gilbody SM, Brealey S, Paulden M, Palmer S, Mann R, *et al.*

No. 37
A double-blind randomised placebo-controlled trial of topical intranasal corticosteroids in 4- to 11-year-old children with persistent bilateral otitis media with effusion in primary care.
By Williamson I, Benge S, Barton S, Petrou S, Letley L, Fasey N, *et al.*

No. 38
The effectiveness and cost-effectiveness of methods of storing donated kidneys from deceased donors: a systematic review and economic model.
By Bond M, Pitt M, Akoh J, Moxham T, Hoyle M, Anderson R.

No. 39
Rehabilitation of older patients: day hospital compared with rehabilitation at home. A randomised controlled trial.
By Parker SG, Oliver P, Pennington M, Bond J, Jagger C, Enderby PM, *et al.*

No. 40
Breastfeeding promotion for infants in neonatal units: a systematic review and economic analysis
By Renfrew MJ, Craig D, Dyson L, McCormick F, Rice S, King SE, *et al.*

No. 41
The clinical effectiveness and cost-effectiveness of bariatric (weight loss) surgery for obesity: a systematic review and economic evaluation.
By Picot J, Jones J, Colquitt JL, Gospodarevskaya E, Loveman E, Baxter L, *et al.*

No. 42
Rapid testing for group B streptococcus during labour: a test accuracy study with evaluation of acceptability and cost-effectiveness.
By Daniels J, Gray J, Pattison H, Roberts T, Edwards E, Milner P, *et al.*

No. 43
Screening to prevent spontaneous preterm birth: systematic reviews of accuracy and effectiveness literature with economic modelling.
By Honest H, Forbes CA, Durée KH, Norman G, Duffy SB, Tsourapas A, *et al.*

No. 44
The effectiveness and cost-effectiveness of cochlear implants for severe to profound deafness in children and adults: a systematic review and economic model.
By Bond M, Mealing S, Anderson R, Elston J, Weiner G, Taylor RS, *et al.*

Suppl. 2
Gemcitabine for the treatment of metastatic breast cancer.
By Jones J, Takeda A, Tan SC, Cooper K, Loveman E, Clegg A.

Varenicline in the management of smoking cessation: a single technology appraisal.
By Hind D, Tappenden P, Peters J, Kenjegalieva K.

Alteplase for the treatment of acute ischaemic stroke: a single technology appraisal.
By Lloyd Jones M, Holmes M.

Rituximab for the treatment of rheumatoid arthritis.
By Bagust A, Boland A, Hockenhull J, Fleeman N, Greenhalgh J, Dundar Y, *et al.*

Omalizumab for the treatment of severe persistent allergic asthma.
By Jones J, Shepherd J, Hartwell D, Harris P, Cooper K, Takeda A, *et al.*

Rituximab for the treatment of relapsed or refractory stage III or IV follicular non-Hodgkin's lymphoma.
By Boland A, Bagust A, Hockenhull J, Davis H, Chu P, Dickson R.

Adalimumab for the treatment of psoriasis.
By Turner D, Picot J, Cooper K, Loveman E.

Dabigatran etexilate for the prevention of venous thromboembolism in patients undergoing elective hip and knee surgery: a single technology appraisal.
By Holmes M, C Carroll C, Papaioannou D.

Romiplostim for the treatment of chronic immune or idiopathic thrombocytopenic purpura: a single technology appraisal.
By Mowatt G, Boachie C, Crowther M, Fraser C, Hernández R, Jia X, *et al.*

Sunitinib for the treatment of gastrointestinal stromal tumours: a critique of the submission from Pfizer.
By Bond M, Hoyle M, Moxham T, Napier M, Anderson R.

No. 45
Vitamin K to prevent fractures in older women: systematic review and economic evaluation.
By Stevenson M, Lloyd-Jones M, Papaioannou D.

No. 46
The effects of biofeedback for the treatment of essential hypertension: a systematic review.
By Greenhalgh J, Dickson R, Dundar Y.

No. 47
A randomised controlled trial of the use of aciclovir and/or prednisolone for the early treatment of Bell's palsy: the BELLS study.
By Sullivan FM, Swan IRC, Donnan PT, Morrison JM, Smith BH, McKinstry B, *et al.*

Suppl. 3
Lapatinib for the treatment of HER2-overexpressing breast cancer.
By Jones J, Takeda A, Picot J, von Keyserlingk C, Clegg A.

Infliximab for the treatment of ulcerative colitis.
By Hyde C, Bryan S, Juarez-Garcia A, Andronis L, Fry-Smith A.

213

Rimonabant for the treatment of overweight and obese people.

By Burch J, McKenna C, Palmer S, Norman G, Glanville J, Sculpher M, *et al.*

Telbivudine for the treatment of chronic hepatitis B infection.

By Hartwell D, Jones J, Harris P, Cooper K.

Entecavir for the treatment of chronic hepatitis B infection.

By Shepherd J, Gospodarevskaya E, Frampton G, Cooper, K.

Febuxostat for the treatment of hyperuricaemia in people with gout: a single technology appraisal.

By Stevenson M, Pandor A.

Rivaroxaban for the prevention of venous thromboembolism: a single technology appraisal.

By Stevenson M, Scope A, Holmes M, Rees A, Kaltenthaler E.

Cetuximab for the treatment of recurrent and/or metastatic squamous cell carcinoma of the head and neck.

By Greenhalgh J, Bagust A, Boland A, Fleeman N, McLeod C, Dundar Y, *et al.*

Mifamurtide for the treatment of osteosarcoma: a single technology appraisal.

By Pandor A, Fitzgerald P, Stevenson M, Papaioannou D.

Ustekinumab for the treatment of moderate to severe psoriasis.

By Gospodarevskaya E, Picot J, Cooper K, Loveman E, Takeda A.

No. 48
Endovascular stents for abdominal aortic aneurysms: a systematic review and economic model.

By Chambers D, Epstein D, Walker S, Fayter D, Paton F, Wright K, *et al.*

No. 49
Clinical and cost-effectiveness of epoprostenol, iloprost, bosentan, sitaxentan and sildenafil for pulmonary arterial hypertension within their licensed indications: a systematic review and economic evaluation.

By Chen Y-F, Jowett S, Barton P, Malottki K, Hyde C, Gibbs JSR, *et al.*

No. 50
Cessation of attention deficit hyperactivity disorder drugs in the young (CADDY) – a pharmacoepidemiological and qualitative study.

By Wong ICK, Asherson P, Bilbow A, Clifford S, Coghill D, R DeSoysa R, *et al.*

No. 51
ARTISTIC: a randomised trial of human papillomavirus (HPV) testing in primary cervical screening.

By Kitchener HC, Almonte M, Gilham C, Dowie R, Stoykova B, Sargent A, *et al.*

No. 52
The clinical effectiveness of glucosamine and chondroitin supplements in slowing or arresting progression of osteoarthritis of the knee: a systematic review and economic evaluation.

By Black C, Clar C, Henderson R, MacEachern C, McNamee P, Quayyum Z, *et al.*

No. 53
Randomised preference trial of medical versus surgical termination of pregnancy less than 14 weeks' gestation (TOPS).

By Robson SC, Kelly T, Howel D, Deverill M, Hewison J, Lie MLS, *et al.*

No. 54
Randomised controlled trial of the use of three dressing preparations in the management of chronic ulceration of the foot in diabetes.

By Jeffcoate WJ, Price PE, Phillips CJ, Game FL, Mudge E, Davies S, *et al.*

No. 55
VenUS II: a randomised controlled trial of larval therapy in the management of leg ulcers.

By Dumville JC, Worthy G, Soares MO, Bland JM, Cullum N, Dowson C, *et al.*

No. 56
A prospective randomised controlled trial and economic modelling of antimicrobial silver dressings versus non-adherent control dressings for venous leg ulcers: the VULCAN trial

By Michaels JA, Campbell WB, King BM, MacIntyre J, Palfreyman SJ, Shackley P, *et al.*

No. 57
Communication of carrier status information following universal newborn screening for sickle cell disorders and cystic fibrosis: qualitative study of experience and practice.

By Kai J, Ulph F, Cullinan T, Qureshi N.

No. 58
Antiviral drugs for the treatment of influenza: a systematic review and economic evaluation.

By Burch J, Paulden M, Conti S, Stock C, Corbett M, Welton NJ, *et al.*

No. 59
Development of a toolkit and glossary to aid in the adaptation of health technology assessment (HTA) reports for use in different contexts.

By Chase D, Rosten C, Turner S, Hicks N, Milne R.

No. 60
Colour vision testing for diabetic retinopathy: a systematic review of diagnostic accuracy and economic evaluation.

By Rodgers M, Hodges R, Hawkins J, Hollingworth W, Duffy S, McKibbin M, *et al.*

No. 61
Systematic review of the effectiveness and cost-effectiveness of weight management schemes for the under fives: a short report.

By Bond M, Wyatt K, Lloyd J, Welch K, Taylor R.

No. 62
Are adverse effects incorporated in economic models? An initial review of current practice.

By Craig D, McDaid C, Fonseca T, Stock C, Duffy S, Woolacott N.

Volume 14, 2010

No. 1
Multicentre randomised controlled trial examining the cost-effectiveness of contrast-enhanced high field magnetic resonance imaging in women with primary breast cancer scheduled for wide local excision (COMICE).

By Turnbull LW, Brown SR, Olivier C, Harvey I, Brown J, Drew P, *et al.*

No. 2
Bevacizumab, sorafenib tosylate, sunitinib and temsirolimus for renal cell carcinoma: a systematic review and economic evaluation.

By Thompson Coon J, Hoyle M, Green C, Liu Z, Welch K, Moxham T, *et al.*

No. 3
The clinical effectiveness and cost-effectiveness of testing for cytochrome P450 polymorphisms in patients with schizophrenia treated with antipsychotics: a systematic review and economic evaluation.

By Fleeman N, McLeod C, Bagust A, Beale S, Boland A, Dundar Y, *et al.*

No. 4
Systematic review of the clinical effectiveness and cost-effectiveness of photodynamic diagnosis and urine biomarkers (FISH, ImmunoCyt, NMP22) and cytology for the detection and follow-up of bladder cancer.

By Mowatt G, Zhu S, Kilonzo M, Boachie C, Fraser C, Griffiths TRL, *et al.*

No. 5

Effectiveness and cost-effectiveness of arthroscopic lavage in the treatment of osteoarthritis of the knee: a mixed methods study of the feasibility of conducting a surgical placebo-controlled trial (the KORAL study).

By Campbell MK, Skea ZC, Sutherland AG, Cuthbertson BH, Entwistle VA, McDonald AM, *et al.*

No. 6

A randomised 2 × 2 trial of community versus hospital pulmonary rehabilitation, followed by telephone or conventional follow-up.

By Waterhouse JC, Walters SJ, Oluboyede Y, Lawson RA.

No. 7

The effectiveness and cost-effectiveness of behavioural interventions for the prevention of sexually transmitted infections in young people aged 13–19: a systematic review and economic evaluation.

By Shepherd J, Kavanagh J, Picot J, Cooper K, Harden A, Barnett-Page E, *et al.*

215

Health Technology Assessment programme

Director,
Professor Tom Walley,
Director, NIHR HTA
programme, Professor of
Clinical Pharmacology,
University of Liverpool

Deputy Director,
Professor Jon Nicholl,
Director, Medical Care Research
Unit, University of Sheffield

Prioritisation Strategy Group

Members

Chair,
Professor Tom Walley,
Director, NIHR HTA
programme, Professor of
Clinical Pharmacology,
University of Liverpool

Deputy Chair,
Professor Jon Nicholl,
Director, Medical Care Research
Unit, University of Sheffield

Dr Bob Coates,
Consultant Advisor, NETSCC,
HTA

Dr Andrew Cook,
Consultant Advisor, NETSCC,
HTA

Dr Peter Davidson,
Director of Science Support,
NETSCC, HTA

Professor Robin E Ferner,
Consultant Physician and
Director, West Midlands Centre
for Adverse Drug Reactions,
City Hospital NHS Trust,
Birmingham

Professor Paul Glasziou,
Professor of Evidence-Based
Medicine, University of Oxford

Dr Nick Hicks,
Director of NHS Support,
NETSCC, HTA

Dr Edmund Jessop,
Medical Adviser, National
Specialist, National
Commissioning Group (NCG),
Department of Health, London

Ms Lynn Kerridge,
Chief Executive Officer,
NETSCC and NETSCC, HTA

Dr Ruairidh Milne,
Director of Strategy and
Development, NETSCC

Ms Kay Pattison,
Section Head, NHS R&D
Programme, Department of
Health

Ms Pamela Young,
Specialist Programme Manager,
NETSCC, HTA

HTA Commissioning Board

Members

Programme Director,
Professor Tom Walley,
Director, NIHR HTA
programme, Professor of
Clinical Pharmacology,
University of Liverpool

Chair,
Professor Jon Nicholl,
Director, Medical Care Research
Unit, University of Sheffield

Deputy Chair,
Dr Andrew Farmer,
Senior Lecturer in General
Practice, Department of
Primary Health Care,
University of Oxford

Professor Ann Ashburn,
Professor of Rehabilitation
and Head of Research,
Southampton General Hospital

Professor Deborah Ashby,
Professor of Medical Statistics,
Queen Mary, University of
London

Professor John Cairns,
Professor of Health Economics,
London School of Hygiene and
Tropical Medicine

Professor Peter Croft,
Director of Primary Care
Sciences Research Centre, Keele
University

Professor Nicky Cullum,
Director of Centre for Evidence-
Based Nursing, University of
York

Professor Jenny Donovan,
Professor of Social Medicine,
University of Bristol

Professor Steve Halligan,
Professor of Gastrointestinal
Radiology, University College
Hospital, London

Professor Freddie Hamdy,
Professor of Urology,
University of Sheffield

Professor Allan House,
Professor of Liaison Psychiatry,
University of Leeds

Dr Martin J Landray,
Reader in Epidemiology,
Honorary Consultant Physician,
Clinical Trial Service Unit,
University of Oxford

Professor Stuart Logan,
Director of Health & Social
Care Research, The Peninsula
Medical School, Universities of
Exeter and Plymouth

Dr Rafael Perera,
Lecturer in Medical Statisitics,
Department of Primary Health
Care, Univeristy of Oxford

Professor Ian Roberts,
Professor of Epidemiology &
Public Health, London School
of Hygiene and Tropical
Medicine

Professor Mark Sculpher,
Professor of Health Economics,
University of York

Professor Helen Smith,
Professor of Primary Care,
University of Brighton

Professor Kate Thomas,
Professor of Complementary &
Alternative Medicine Research,
University of Leeds

Professor David John
Torgerson,
Director of York Trials Unit,
University of York

Professor Hywel Williams,
Professor of Dermato-
Epidemiology, University of
Nottingham

Observers

Ms Kay Pattison,
Section Head, NHS R&D
Programme, Department of
Health

Dr Morven Roberts,
Clinical Trials Manager,
Medical Research Council

Current and past membership details of all HTA programme 'committees' are available from the HTA website (www.hta.ac.uk)

Therapeutic Procedures Panel

Disease Prevention Panel

Expert Advisory Network

Members

Professor Douglas Altman,
Professor of Statistics in
Medicine, Centre for Statistics
in Medicine, University of
Oxford

Professor John Bond,
Professor of Social Gerontology
& Health Services Research,
University of Newcastle upon
Tyne

Professor Andrew Bradbury,
Professor of Vascular Surgery,
Solihull Hospital, Birmingham

Mr Shaun Brogan,
Chief Executive, Ridgeway
Primary Care Group, Aylesbury

Mrs Stella Burnside OBE,
Chief Executive, Regulation
and Improvement Authority,
Belfast

Ms Tracy Bury,
Project Manager, World
Confederation for Physical
Therapy, London

Professor Iain T Cameron,
Professor of Obstetrics and
Gynaecology and Head of the
School of Medicine, University
of Southampton

Dr Christine Clark,
Medical Writer and Consultant
Pharmacist, Rossendale

Professor Collette Clifford,
Professor of Nursing and
Head of Research, The
Medical School, University of
Birmingham

Professor Barry Cookson,
Director, Laboratory of Hospital
Infection, Public Health
Laboratory Service, London

Dr Carl Counsell,
Clinical Senior Lecturer in
Neurology, University of
Aberdeen

Professor Howard Cuckle,
Professor of Reproductive
Epidemiology, Department
of Paediatrics, Obstetrics &
Gynaecology, University of
Leeds

Dr Katherine Darton,
Information Unit, MIND – The
Mental Health Charity, London

Professor Carol Dezateux,
Professor of Paediatric
Epidemiology, Institute of Child
Health, London

Mr John Dunning,
Consultant Cardiothoracic
Surgeon, Papworth Hospital
NHS Trust, Cambridge

Mr Jonothan Earnshaw,
Consultant Vascular Surgeon,
Gloucestershire Royal Hospital,
Gloucester

Professor Martin Eccles,
Professor of Clinical
Effectiveness, Centre for Health
Services Research, University of
Newcastle upon Tyne

Professor Pam Enderby,
Dean of Faculty of Medicine,
Institute of General Practice
and Primary Care, University of
Sheffield

Professor Gene Feder,
Professor of Primary Care
Research & Development,
Centre for Health Sciences,
Barts and The London School
of Medicine and Dentistry

Mr Leonard R Fenwick,
Chief Executive, Freeman
Hospital, Newcastle upon Tyne

Mrs Gillian Fletcher,
Antenatal Teacher and Tutor
and President, National
Childbirth Trust, Henfield

Professor Jayne Franklyn,
Professor of Medicine,
University of Birmingham

Mr Tam Fry,
Honorary Chairman, Child
Growth Foundation, London

Professor Fiona Gilbert,
Consultant Radiologist and
NCRN Member, University of
Aberdeen

Professor Paul Gregg,
Professor of Orthopaedic
Surgical Science, South Tees
Hospital NHS Trust

Bec Hanley,
Co-director, TwoCan Associates,
West Sussex

Dr Maryann L Hardy,
Senior Lecturer, University of
Bradford

Mrs Sharon Hart,
Healthcare Management
Consultant, Reading

Professor Robert E Hawkins,
CRC Professor and Director
of Medical Oncology, Christie
CRC Research Centre,
Christie Hospital NHS Trust,
Manchester

Professor Richard Hobbs,
Head of Department of Primary
Care & General Practice,
University of Birmingham

Professor Alan Horwich,
Dean and Section Chairman,
The Institute of Cancer
Research, London

Professor Allen Hutchinson,
Director of Public Health and
Deputy Dean of ScHARR,
University of Sheffield

Professor Peter Jones,
Professor of Psychiatry,
University of Cambridge,
Cambridge

Professor Stan Kaye,
Cancer Research UK Professor
of Medical Oncology, Royal
Marsden Hospital and Institute
of Cancer Research, Surrey

Dr Duncan Keeley,
General Practitioner (Dr Burch
& Ptnrs), The Health Centre,
Thame

Dr Donna Lamping,
Research Degrees Programme
Director and Reader in
Psychology, Health Services
Research Unit, London School
of Hygiene and Tropical
Medicine, London

Mr George Levvy,
Chief Executive, Motor
Neurone Disease Association,
Northampton

Professor James Lindesay,
Professor of Psychiatry for the
Elderly, University of Leicester

Professor Julian Little,
Professor of Human Genome
Epidemiology, University of
Ottawa

Professor Alistaire McGuire,
Professor of Health Economics,
London School of Economics

Professor Rajan Madhok,
Medical Director and Director
of Public Health, Directorate
of Clinical Strategy & Public
Health, North & East Yorkshire
& Northern Lincolnshire
Health Authority, York

Professor Alexander Markham,
Director, Molecular Medicine
Unit, St James's University
Hospital, Leeds

Dr Peter Moore,
Freelance Science Writer,
Ashtead

Dr Andrew Mortimore,
Public Health Director,
Southampton City Primary
Care Trust

Dr Sue Moss,
Associate Director, Cancer
Screening Evaluation Unit,
Institute of Cancer Research,
Sutton

Professor Miranda Mugford,
Professor of Health Economics
and Group Co-ordinator,
University of East Anglia

Professor Jim Neilson,
Head of School of Reproductive
& Developmental Medicine
and Professor of Obstetrics
and Gynaecology, University of
Liverpool

Mrs Julietta Patnick,
National Co-ordinator, NHS
Cancer Screening Programmes,
Sheffield

Professor Robert Peveler,
Professor of Liaison Psychiatry,
Royal South Hants Hospital,
Southampton

Professor Chris Price,
Director of Clinical Research,
Bayer Diagnostics Europe,
Stoke Poges

Professor William Rosenberg,
Professor of Hepatology
and Consultant Physician,
University of Southampton

Professor Peter Sandercock,
Professor of Medical Neurology,
Department of Clinical
Neurosciences, University of
Edinburgh

Dr Susan Schonfield,
Consultant in Public Health,
Hillingdon Primary Care Trust,
Middlesex

Dr Eamonn Sheridan,
Consultant in Clinical Genetics,
St James's University Hospital,
Leeds

Dr Margaret Somerville,
Director of Public Health
Learning, Peninsula Medical
School, University of Plymouth

Professor Sarah Stewart-Brown,
Professor of Public Health,
Division of Health in the
Community, University of
Warwick, Coventry

Professor Ala Szczepura,
Professor of Health Service
Research, Centre for Health
Services Studies, University of
Warwick, Coventry

Mrs Joan Webster,
Consumer Member, Southern
Derbyshire Community Health
Council

Professor Martin Whittle,
Clinical Co-director, National
Co-ordinating Centre for
Women's and Children's
Health, Lymington

Current and past membership details of all HTA programme 'committees' are available from the HTA website (www.hta.ac.uk)