Publication and related bias in health services and delivery research (HSDR): a systematic review of literature and evaluation of empirical evidence and methodology, and key informant interview

Protocol

Sponsor:	University of Warwick (Lead), University of Birmingham, University of East Anglia
Version number	2.2
Date	11 April 2017
Summary of amendments from the previous version	Version 2.0: An additional section "Ethical considerations" and related appendices were added to comply with the requirement of the Biomedical & Scientific Research Ethics Committee (BSREC) of the University of Warwick. The configuration and list of members of the project team have been updated to reflect changes made to comply with the NIHR's requirement for the Study Steering Committee Version 2.1: Changes made on request of BSREC, including clarification of outcome measures for Work Package 4 on page 17, details regarding the process of inviting interviewees and obtaining consents on pages 20, and clarification of requirement for HRA approval on page 24. Protocol appendices were detached as separate protocol attachments Version 2.2: an invitation e-mail for potential interviewees was added as an attachment for the protocol, with re-numbering of the protocol attachments

Acknowledgement:

This project is funded by the National Institute for Health Research Health Services and Delivery Research Programme (project number 15/71/06).

The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the HS&DR Programme, NIHR, NHS or the Department of Health.

NHS
National Institute for Health Research

Investigators (including project management group)

Dr Yen-Fu Chen, University of Warwick (Principal Investigator)

Professor Richard Lilford, University of Warwick

Dr Iestyn Williams, University of Birmingham

Dr Abimbola Ayorinde, University of Warwick

Professor Russell Mannion, University of Birmingham

Professor Fujian Song, University of East Anglia

Dr Magdalena Skrybant, Patient & Public Involvement Adviser

Study steering committee

Professor Stephen Sutton, Chair

Reverend Dr Barry Clark, Patient & Public Involvement Adviser

Dr Kaveh Shojania, University of Toronto

Professor Jeremy Grimshaw, University of Ottawa

Professor Timothy Hofer, University of Michigan

Dr Christopher Chiswell, Birmingham Children's Hospital

Mr Tim Sacks, East Leicestershire & Rutland Clinical Commissioning Group

Professor James Thomas, University College London

Professor Richard Lilford, University of Warwick

Address for correspondence

Dr Yen-Fu Chen, Division of Health Sciences, University of Warwick, Coventry CV4 7AL, United Kingdom Email: Y-F.Chen@warwick.ac.uk

Contents

Background, rationale and why this research is needed now

Decision-making at all levels of health care needs to be guided by evidence. Many factors can influence the accessibility of evidence and lead to the information available to decision makers being incomplete or biased. One particular concern is publication bias, which refers to the phenomenon that the likelihood of a research study being published is related to the direction and/or statistical significance of its findings.[1] Other forms of related bias can occur between the generation of research evidence and its uptake by end users. These include issues arising during the analysis and writing up of research data, such as p-hacking (repeating analyses using different methods or subsets of data until statistically significant results are obtained)[2,3] and selective outcome reporting (only reporting favourable outcomes among many outcomes investigated),[4] and potential bias in the dissemination of evidence following the publication of research findings, such as citation bias and media attention bias.[1] A schematic presentation of these various forms of bias, collectively termed 'dissemination bias', and the stages in which they could happen are shown in Figure 1. For this research proposal, we focus on potential bias that occurs up to the stage of publication and use the term 'publication bias' to refer to these forms of publication and related bias for brevity, but do not look into broader dissemination bias following the publication of research findings as the latter can largely be overcome by systematic literature search, which is increasingly being adopted when information is gathered to support decision making.

Publication bias has been documented extensively in clinical research,[1,4,5] and concerns over its impact on scientific integrity and ethical implications have led to mandatory clinical trial registration.[6] By contrast, health services and delivery research (HSDR) is not subject to a similar level of regulation and scrutiny as in pharmaceutical research, and the issue of publication bias seems to be infrequently mentioned/discussed. We undertook a scoping review [7] in health services literature through a search of PubMed and Google without year and language restrictions using combination of terms including 'publication bias' and key terms such as health services research, health services management, service delivery, quality improvement and patient safety. This literature appears to be largely silent on the topic; while the possibility of publication bias is occasionally mooted in systematic reviews of HSDR,[8-11] we found only two sets of empirical studies that specifically focus on the subject itself: one provided two case studies (income elasticities of health care and price elasticities of prescribed drugs)[12-14] while the other examined potential publication bias in health informatics research.[15-17] Although these cases are illustrative, the volume of empirical investigation of publication bias in HSDR is surprisingly small compared with the abundance of literature on this subject in clinical research. The paucity of documented evidence was affirmed in our discussions with some of the leading experts in health services research.

Given that the publication bias has also been documented in many other scientific disciplines, [18-23] and the limited evidence in HSDR mentioned above suggest its existence, there is no obvious reason to believe that HSDR is immune from publication bias. Rather than assuming that the bias does not exist or is largely ignored in HSDR, we posit that the paucity of literature may arise from certain features of HSDR, making its investigation challenging. First, quasi- and non-experimental designs make up a larger proportion of studies in HSDR than in clinical research and there is no requirement to register such studies prospectively. The lack of a comprehensive registration means there is currently no good audit trail for tracking studies and quantifying the extent of publication bias in HSDR except the relatively small proportion of studies funded through competitive grants. Second, the issue is further clouded by lack of a clear boundary between research and routine quality improvement activities. Third, given the complexity of health systems, it is often necessary to investigate the associations between a large number of variables/ outcomes along the service delivery causal pathway. This may reduce the probability of having a study with a completely null result (and associated publication bias) at study level, but nonetheless may increase the possibility of other bias such as p-hacking and selective outcome reporting. Finally, the inherent complexity and fluidity of HSDR interventions and their interaction with contextual factors often result in heterogeneity between studies and create problems in using standard techniques such as funnel plots and regression methods to assess publication bias in systematic reviews of HSDR.

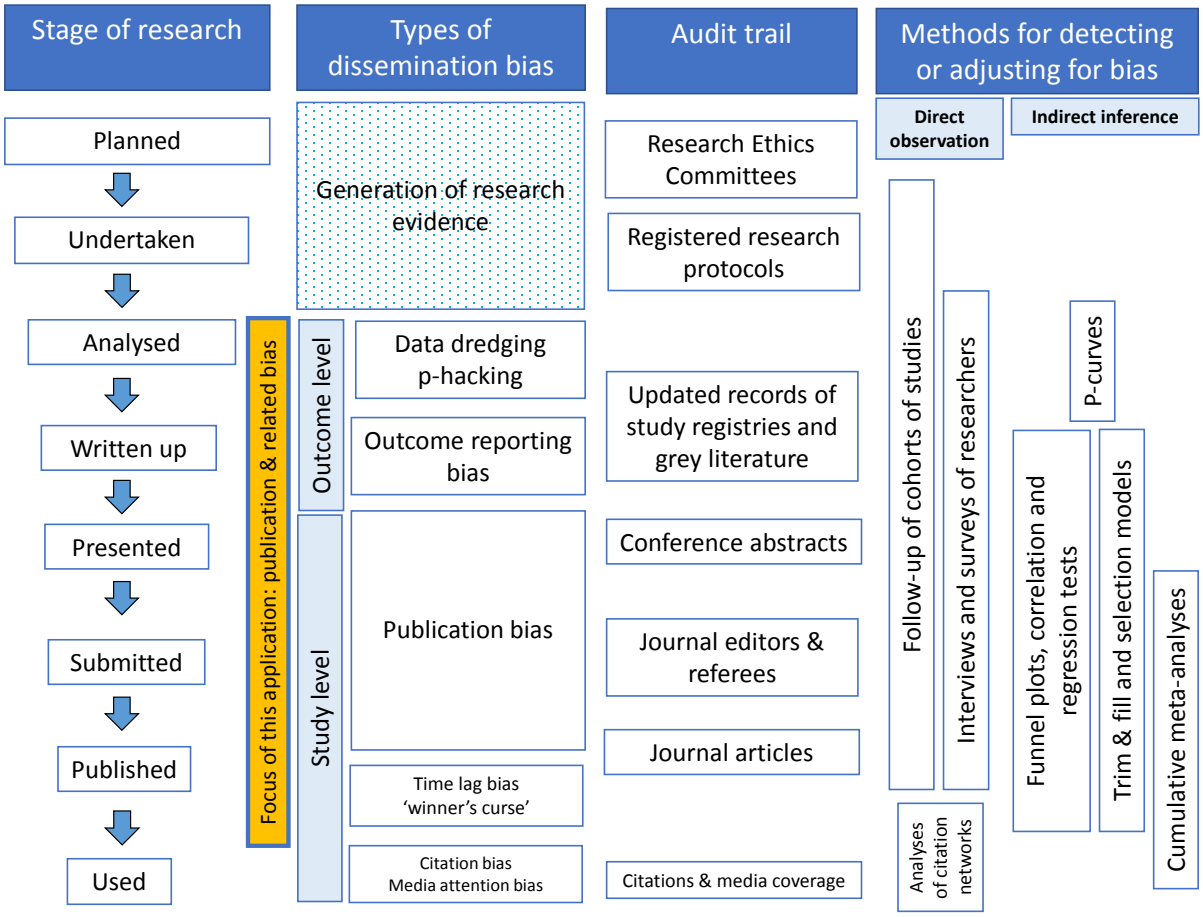


Figure 1. Potential bias in the publication and dissemination of findings at various stages of research and methods used to detect their occurrence or estimate their impact

Despite these challenges, a systematic investigation of publication bias in HSDR is of crucial importance because HSDR frequently informs decisions at institutional and policy levels, and failure to recognise bias in evidence used to inform decisions could have substantial implications for population health and resource allocation. This proposed research aims to address this important issue by gathering *prima facie* evidence of the existence and potential impact of publication bias in HSDR, documenting current practice and exploring common methods in detecting and mitigating the bias and providing recommendations for future research and practice.

Scope

The subject area of HSDR is very broad. In order to draw a boundary which allows a focused investigation, we target: (1) *intervention studies*, which are carried out to evaluate interventions to improve/optimise the effectiveness and/or efficiency of the delivery of health services; (2) *association studies*, which are carried out to evaluate associations between different variables along the service delivery causal chain.[24] A large number of variables can be covered in association studies in HSDR. These include structure variables (e.g. characteristics of a hospital, nurse-patient ratio, etc.), generic processes (e.g. continuous professional development, institutional human resource policy, etc.), intervening variables (e.g. safety culture, staff knowledge and morale etc. that could be influenced by structure and generic processes and then impact upon many downstream processes),[24] targeted processes (e.g. door-to-balloon time for treating myocardial infarction; adherence to guidelines for management of patients with diabetes), health service utilisation, patient, carer or health care provider

outcomes and context (e.g. weekdays vs weekends; low and middle income countries vs high income countries etc).

The criteria for selecting intervention and association studies will be applied within the remit stated by the HS&DR Programme, which concerns research that produces evidence on the quality, accessibility and organisation of health services. Eligible studies may focus on any aspects of health systems and health policy, health care organisations, people who organise and deliver the health services, and users and carers of the services, as well as related processes, outcomes and contextual factors. Studies concerning clinical research and health technology assessment (i.e. those focusing on interventions applied directly to individual patients), disease epidemiology and genetic associations have previously been examined in detail [1] and therefore will not be included in this project. We are aware of potential grey areas where the boundary between HSDR and non-HSDR studies may be vague. These will be dealt with by consulting members of the project steering committee and project advisory committee.

A wide variety of research designs, including quantitative, qualitative and mixed methods research, have been used in HSDR. [25,26] This research proposal will focus on quantitative research and mixed methods research that incorporated an element of quantitative estimation of intervention effects or association, although we acknowledge that qualitative research can also be subject to publication bias.[27] As the mechanisms and manifestation of publication bias for qualitative research are likely to be different and methods for evaluating its occurrence and impact are not well developed, we feel that issues related to qualitative research are beyond the scope of the current proposal and warrant a separate investigation.[28]

Aims and objectives

The aims for the proposed research are to obtain *prima facie* evidence on the existence and extent of publication bias in HSDR, and to document current practice and explore the most suitable methods for detecting and mitigating its occurrence during evidence synthesis. The ultimate purpose is to illuminate the issue of publication bias in the HSDR community and to promote good research practice to minimise its future occurrence and impact.

The above aims will be achieved through five inter-related Work Packages, each with a specific objective:

Work Package 1 (WP1): a systematic review of empirical and methodological studies concerning the occurrence, potential impact and/or methodology related to publication bias in HSDR and cognate fields to provide a summary of what is known from current literature.

Work Package 2 (WP2): an overview of systematic reviews of intervention and association studies in HSDR to describe current practice and potential challenges in assessing publication bias during evidence synthesis.

Work Package 3 (WP3): in-depth case studies to evaluate the applicability of different methods for detecting and mitigating publication bias in HSDR and to provide guidance for future research and practice.

Work Package 4 (WP4): a retrospective study to follow up the publication status of cohorts of HSDR studies to directly observe publication bias in HSDR

Work Package 5 (WP5): semi-structured interviews with health services researchers and commissioners, journal editors and other stakeholders to explore their perception and experience related to publication bias.

These Work Packages will complement each other to provide a full picture of both methodology and empirical evidence related to dissemination bias in HSDR, as shown in Figure 2 below.

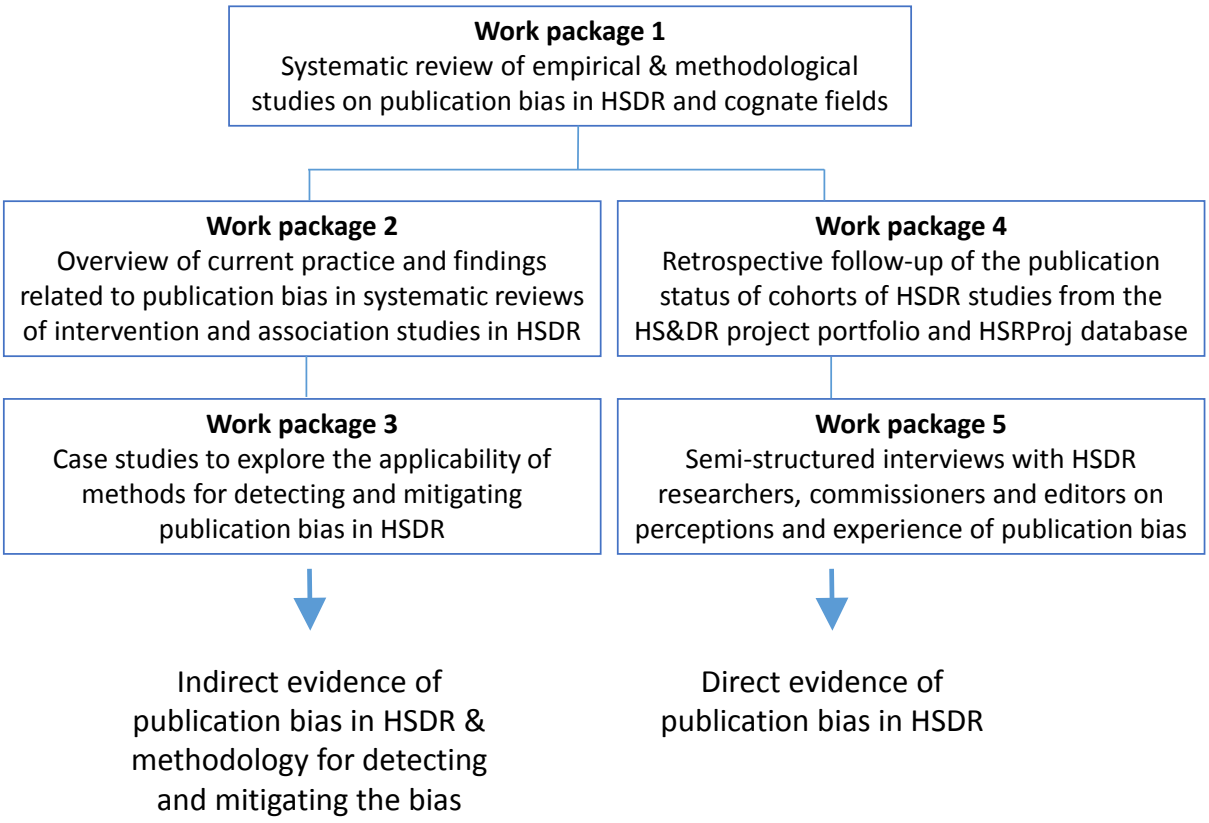


Figure 2 The five Work Packages that constitute the proposed research

Justifications for the chosen approaches

Many methods have been developed in order to detect publication bias and estimate and/or mitigate its potential impact.[1,29] Methods used to detect publication bias can be broadly classified as either making indirect inference or direct observation.

Several methods can suggest that publication bias is present and estimate its potential impact during the synthesis of evidence.[29] These methods only allow indirect inference as they rely on the identification of specific patterns in the literature suggestive of publication bias but cannot rule out alternative causes. Some of the methods such as funnel plots and regression tests have been used routinely in systematic reviews of clinical interventions. However these methods appear to have been used infrequently in systematic reviews of HSDR interventions. For example, a recent overview of 99 Cochrane EPOC reviews found that publication bias was explicitly assessed in only 9% of the reviews and was mentioned in some way in a further 23% of the reviews.[30] Another recent overview of systematic reviews on interventions to improve the quality of care for people with diabetes identified 125 relevant systematic reviews and selected 50 which were judged to be of higher quality for detailed assessment.[31] However even within these better-quality reviews, less than half (22/50) had assessed the likelihood of publication bias. The reasons for this lack of attention were not discussed in these overviews. This proposed research will therefore systematically investigate publication bias in HSDR through three literature-based Work Packages: WP1 will provide a thorough review of literature concerning what is known in HSDR; WP2 will systematically examine current practice and issues in the synthesis of evidence from HSDR studies with regard to detecting publication bias; and WP3 will attempt to obtain further insight through in-depth case studies to explore the applicability of different methods in HSDR and to provide guidance for future research and practice.

In contrast with methods described above which use literature to make indirect inference, direct observation relies on either (1) identification of a cohort of studies and then following them over time to determine whether they are published;, or (2) interrogation of stakeholders who are involved in generating or disseminating research evidence in HSDR through surveys or interviews to find out their experience. While these methods are labour intensive and face some methodological challenges (described later), they provide the strongest evidence on the presence/absence of publication bias. Therefore we will adopt both approaches in the proposed research (WP4 and WP5 respectively). Detailed plans for each of the five Work Packages are presented below.

Research plan / methods for individual Work Packages

Work Package 1 (WP1): A systematic review of empirical and methodological studies on publication bias in HSDR and cognate fields

WP1 is a systematic review of studies that have examined publication bias in HSDR. The objectives are to extend our scoping review:

- To systematically search and summarise studies that have reported empirical evidence of publication bias in HSDR and in cognate fields such as management and behaviour sciences.
- To summarise methodological literature (if found) that has examined or discussed methods for detecting and mitigating publication bias in HSDR and in cognate fields.

WP1 builds upon our initial scoping review described in the background section and will provide a comprehensive overview of what is known on this topic in terms of the occurrence, causes and consequences of publication bias in HSDR and cognate fields. Were available, studies that have examined practical and methodological issues related to the investigation of publication bias in HSDR will also be summarised. The findings of WP1 serve as a 'baseline' knowledge and (where appropriate) can be used to help refine the methodology for the remaining Work Packages.

Methods

Protocol registration

The scope of this systematic review does not fulfil inclusion criteria for registration in the PROSPERO database, which currently focuses on clinical reviews. However we will deposit a protocol of the review in Warwick Research Archive Portal (WRAP), the University of Warwick's open access repository before the review commences.

Search strategy

The diverse research disciplines, subject areas and terminologies related to HSDR pose a challenge for searching relevant literature.[32] We will use a combination of different information sources and searching methods to ensure that our coverage of literature is as comprehensive as possible and is inclusive of disciplines closely related to HSDR. These include search of general and HSDR-specific electronic databases, citation search of key papers (snow-balling), search of the internet and contact with experts. Given that three of the Work Packages are literature-based, we provide an overview summarising the strategy for identifying relevant literature across the Work Packages (WP1 to WP3) in Figure 3 below.

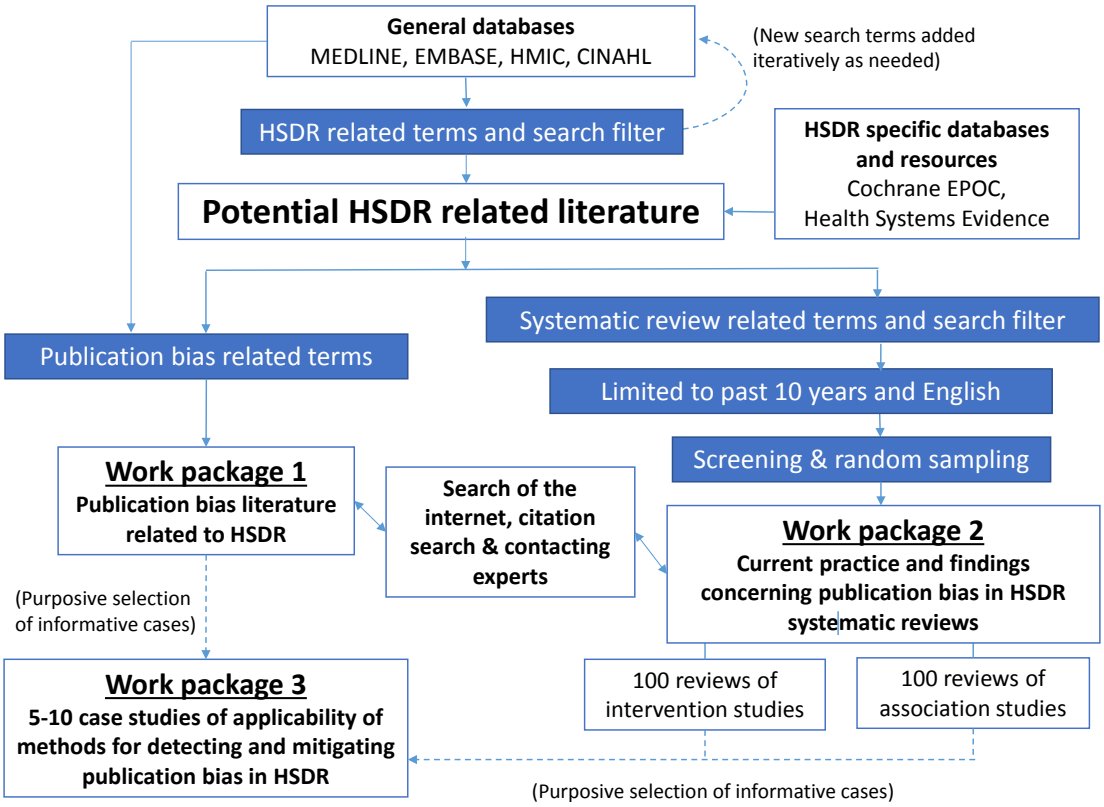


Figure 3. Overview of search strategies for identifying relevant literature for the three literature-based Work Packages

Search of electronic databases

General databases including MEDLINE, EMBASE, HMIC (Health Management Information Consortium), CINAHL, and Web of Science (which includes Social Science Citation Index) will be searched using indexed terms and text words related to HSDR (defined broadly).[32] The retrieved records, along with records contained within HSDR-specific databases including Health Systems Evidence and systematic reviews published by the Cochrane Effective Practice and Organisation of Care (EPOC) Review Group, constitute the initial pool of potential HSDR literature. These records will be searched using publication bias related terms to identify HSDR studies that are potentially relevant to WP1. The searches will be undertaken iteratively (where necessary) so that additional terms related to HSDR that are not covered in the initial search can be added.

Forward and backward citation search ('snow balling')

The essential role of using forward and backward citation search for locating literature for complex topics has been highlighted.[23] We will therefore use these techniques to identify additional studies that may not have been captured in the electronic database searches. Reference lists of all papers judged to be eligible for inclusion in WP1 will be manually examined. Subsequently published papers that have cited these papers will also be located and screened using ISI Web of Science and/or Google Scholar.

Search of the internet

The importance of grey literature in health services research has been highlighted in a report funded by the US National Library of Medicine.[33] Internet will be searched via Google to locate grey literature published by major organisations related to HSDR, such as the Health Foundation, King's Fund, IHI, AHRQ, and RAND. In addition we will search the NIHR HS & DR website and the US

HSRProj (Health Services Research Projects in Progress) database for previously commissioned and ongoing studies.

Consultation with experts

We will contact international experts in HSDR through our existing network associated with Collaboration for Leadership in Applied Health Research and Care West Midlands (CLAHRC WM) and members of the project steering and advisory committees to identify any additional studies that may not have been captured by other means.

Study screening and selection

Records retrieved from electronic databases and subsequently obtained from other sources will be imported into a reference management program to facilitate identification and removal of duplicates. Retrieved records will initially be screened based on titles and abstracts to exclude clearly irrelevant records. Full-text publications will be retrieved for the remaining records and an inclusion/exclusion decision will be made for each study based on the selection criteria described below. The screening and study selection will be carry out be two reviewers independently, with any disagreement resolved by discussion or referring to the wider research team.

Inclusion criteria

Studies of any design that have examined any forms of publication bias within the scope of this project will be included. Specifically, a study needs to:

(1) Have investigated data dredging/p-hacking, selective outcome reporting or publication bias; or evaluated methods for detecting these forms of bias; AND

(2) Have provided empirical, quantitative or qualitative evidence (i.e. not just commentaries or opinions); AND

(3) Be concerned with HSDR related topics.

Data extraction

The following information will be extracted from each included studies:

- Citation details
- Methods of selecting study sample and characteristics of the sample
- Methods for investigating publication bias
- Key findings, limitations and conclusions reported by the authors

Data synthesis and reporting

Included studies will be classified according to the type of bias examined, methods adopted to investigate the bias and HSDR topic involved. Studies characteristics, methods and findings will be tabulated and narrative summaries will be provided. It is expected that there will be insufficient data for quantitative synthesis. However, where multiple studies estimating similar bias are found, results may be pooled across studies using random effects meta-analyses. In such a scenario an analysis plan will be drafted and agreed by the Project Steering Committee before the analysis is carried out.

Work package 2 (WP2): Overview of current practice and findings related to publication bias in systematic reviews of intervention and association studies in HSDR

Systematic reviews have emerged as a key tool for summarising the rapidly expanding evidence base in a way that maximises the completeness while minimises potential bias in their coverage of relevant evidence. Steps to identify and reduce various types of bias are built into the process of a systematic review. The following steps are particularly relevant for publication bias:

- Comprehensive search of literature, including attempts to locate unpublished studies
- Assessment of outcome reporting bias of included studies
- Assessment of potential publication bias using funnel plots, related regression methods or other techniques

While these features are routinely incorporated into systematic reviews of clinical interventions, two recent overviews of systematic reviews of HSDR interventions found that publication bias was formally assessed in less than half of included systematic reviews as described earlier.[1,2] Reasons for the low adoption of formal assessment of publication bias were not explored, although it could either reflect the lack of awareness of the methods or due to various methodological barriers mentioned above. This Work Package intends to document current practice and investigate reasons behind the lack of assessment of publication bias in HSDR systematic reviews by examining and summarising a representative sample of published systematic reviews in HSDR.

We will focus on systematic reviews covering two main types of quantitative study – intervention studies and association studies as described earlier. We will examine whether the practice of assessing publication bias and the findings differ between these two types of studies as they tend to be based upon evidence generated from different study designs. As the approaches to investigation and occurrence and relevance of publication bias may vary according to topics within the broad field of HSDR, we will also explore whether differences exist between different categories of journals reflecting different epistemological propositions.

Objectives

1. To collate a representative sample of systematic reviews of HSDR interventions and associations and to describe their characteristics.
2. To examine whether publication bias was assessed and/or discussed in HSDR systematic reviews, and to summarise the methods adopted and findings reported or reasons stated for not formally assessing the bias.
3. To compare whether the practice of, and findings from assessing publication bias differ according to types of study (intervention vs association) and types of journal in which they are published.

The findings of this work package will provide a fuller picture of current practice and issues in assessing publication and related bias in systematic reviews of HSDR and provide empirical evidence for guiding further research and practice.

Methods

The stages of the proposed methodological overview are summarised in Figure 4 below. A detailed description of each stage will follow.

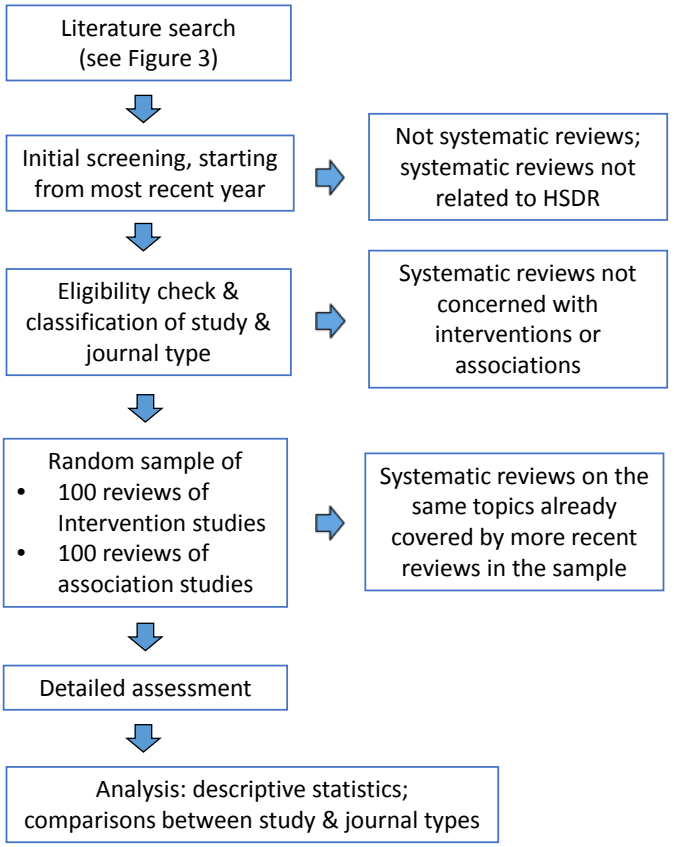


Figure 4: stages of Work Package 2

Protocol registration

The scope of this methodological overview of systematic reviews does not fulfil current inclusion criteria for registration in the PROSPERO database. However we will deposit a protocol of the overview in Warwick Research Archive Portal (WRAP), the University of Warwick's open access repository before the overview commences.

Search strategies

This Work Package shares the initial pool of potential HSDR studies derived from the literature search for WP1 (see Figure 3) but applies terms related to systematic reviews and meta-analysis to retain only studies of these designs. In addition, limits on years of publication (within the last 10 years) and English language will be applied to focus on reflect literature that is likely to be accessed by UK decision makers.

Eligibility check and classification of studies

Records retrieved after applying the study design, year and language restrictions will initially be screened, starting from the most recent year, to eliminate non-systematic reviews and systematic reviews that fall outside the scope of HSDR as defined for this project. For this project, a systematic review is defined as a literature review with explicit statements with regard to research question(s), strategy for literature search and criteria for study selection. Systematic reviews that investigate the effectiveness and cost-effectiveness of clinical interventions (i.e. those traditionally fall under the provenance of health technology assessment) and that explore the association between risk factors and disease conditions (i.e. those fall under the provenance of clinical and genetic epidemiology) will be excluded.

Records that pass through the initial screening will be subject to eligibility check, during which each systematic review will be categorised according to types of study and types of journal in which the review was published. Studies will be classified as either:

- Intervention reviews: reviews of *intervention studies* (as defined in the Scope on pages 2-3) that provide at least one quantitative estimate of an intervention effect); or
- Association reviews: reviews of *association studies* (as defined in the Scope on page 2-3) that provide at least one quantitative estimate of a stated association); or
- Other reviews.

A systematic review that is classified as either an intervention review or an association review will be eligible provided it has a quantitative component, irrespective of whether a meta-analysis was carried out in the review. Systematic reviews classified as 'other reviews' will be excluded. We recognise that intervention studies cannot be completely separated from association studies, as the former is a special case of the latter. Nevertheless such classification reflects how research questions are often asked in HSDR (e.g. whether an intervention works vs. whether certain factors affect one another or influence outcomes in the health system). We will measure the agreement between reviewers in undertaking the classification. We hypothesise that association studies may be more vulnerable to selective publication and reporting than intervention studies because a causal relationship is assumed between an intervention and outcomes whereas relationships between different factors examined in association studies are exploratory and not necessarily causal. In addition, evaluation of interventions may be more likely to be specifically funded with a mandate from the funder to disseminate results, whereas association studies may be carried out without specific funding and related incentive for publication.

The journal/media in which the systematic review was published will be classified based on subject categories of the Journal Citation Reports (ISI Web of Knowledge, Thomson Reuters) as medical journals, health services research and health policy journals, management and social science journals or others (including grey literature) – see Table 1 below for example journals within each category. The study screening, eligibility check and classification will be carried out independently by two reviewers as in WP1. Study types and journals for which the categorisation is ambiguous will be referred to Project Steering Committee for discussion and arbitration.

Table 1 Classification of journal types

Types of journal	**Medical**	**Health services research/health policy**	**Management and social science**	**Other**
Examples	NEJM, JAMA, BMJ, Lancet, Ann Intern Med	Milbank Q,BMJ Qual Saf, J Health Serv Res Policy, Health Policy, Health Serv Res, Health Aff, HS & DR, Implement Sci	Soc Sci Med, Organisation Studies, Human Relations, Sociol Health Illn	Reports from governments, quality improvement institutions, charities

Sampling

Our scoping of literature suggests there exists a large number of systematic reviews in HSDR (e.g. a search of the 'Health Systems Evidence' database using the term 'systematic review' generated 4185 systematic reviews of effects and 1095 systematic reviews addressing other questions). From the pool of studies that have passed through the eligibility check and classification stage described above, we will obtain a random sample of 100 reviews of intervention studies and 100 reviews of association studies for further assessment. Where more than one review within the initially selected samples cover overlapping interventions or associations, only the latest review will be retained to maintain the independence of observations (i.e. reduce overlap of included studies between reviews) and to capture the contemporary practice. Additional studies will be randomly sampled to replace the discarded reviews. While this sample size only has a statistical power of approximately 80% to detect

a large (20%) difference in the characteristics and findings between different types of review (assuming a baseline rate of 32%, the proportion of Cochrane EPOC reviews in which publication bias was formally assessed or partial information was given),[30] the sample size is considered adequate and practical as the main aim is to provide a broad picture of the practice and findings of systematic reviews in HSDR with respect to publication bias, and the comparisons will be exploratory. A record will be kept for reviews checked for eligibility but are not included in the sample due to lack of quantitative synthesis or overlapping scope.

Detailed assessment of sampled systematic reviews

Sampled systematic reviews will be evaluated in detail with regard to the characteristics of the review, the methods and findings related to potential publication bias and issues raised concerning its assessment. The following data will be extracted:

- Key study question(s) for which quantitative estimates were sought, e.g. associations or intervention effects; number of (types of) interventions and/or the structure, intervening, process, outcome and context variables being examined, e.g. a review may have evaluated the association between two measures of a structure variable 'nurse-patient ratio' (using total number of registered nurses and total number of nursing staff including nurse assistant) and three outcome variables (mortality, length of hospital stay and patient satisfaction).

- Databases searched, and whether an attempt was made to search grey literature and unpublished reports, or reasons for not doing this justified.

- Types of studies included in terms of study design.

- Any mentioning of reporting bias (omission from reporting of some measured outcomes; selective reporting of only 'significant' findings from among several analyses undertaken).

- Methods (if used at all) for detecting and/or mitigating potential publication bias (apart from comprehensive search), e.g. funnel plots and related regression methods, trim and fill, and reporting median effects rather than pooled mean effects.

- Findings of assessment of publication bias or reasons for not formally assessing this.

Data extraction will be carried out by one reviewer and be independently checked by another reviewer, with any discrepancies resolved by discussions or contacting authors.

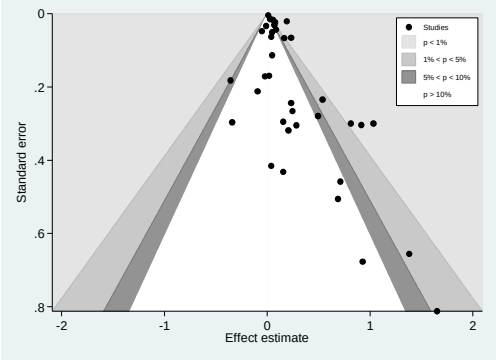
Data analysis

Descriptive statistics will be compiled to summarise the characteristics of HSDR systematic reviews, the practice of assessing publication and related bias among the reviews and their findings. Exploratory comparisons of review characteristics, practice and findings of assessing publication bias will be made between intervention and association reviews, and between reviews published in different categories of journals using both univariate and multivariate logistic regression. Within intervention reviews, we will also make comparisons between reviews which only include randomised controlled trials (RCTs) and those which also include studies of non-randomised designs.

Work package 3 (WP3): Case studies to explore the applicability of methods for detecting and dealing with publication bias

Several methods have been developed to facilitate the detection and potential adjustment of publication and related bias. Among them, funnel plots and related regression methods are most widely used and have been adopted in many systematic reviews. The key assumption for these methods is that the precision of a study (mainly determined by its sample size) is not correlated with the actual size of the intervention effect or association being estimated and hence the results of smaller studies scattered more widely, forming an inverse funnel shape when plotted against precision. Asymmetry in a funnel plot would suggest possible publication bias. Figure 5 shows an example of an asymmetric funnel plot compiled using data from a published systematic review of

mortality risk associated with out of hour admissions in patients with myocardial infarction (Sorita et al. 2014).



The funnel plot was compiled using data from a systematic review investigating the association between out of hour admissions and mortality in patients with acute myocardial infarction (Sorita et al. BMJ 2014;348:f7393). The plot was 'contour enhanced' to facilitate its interpretation. The study author noticed funnel plot asymmetry, reported p=0.052 for the Begg-Mazumdar test and no impact on estimated effect size using the trim and fill method, and stated that "assessment of publication bias was limited in the setting of heterogeneous effect size".

Figure 5 An example of an asymmetric funnel plot

While the assumption behind funnel plots and related regression methods holds for many clinical interventions, this is not necessarily true in many HSDR studies. For example, early evaluation of a quality improvement intervention in a small number of sites may observe a large intervention effect due to the expertise and dedication of the personnel and thoroughness of implementation, which may be difficult to maintain when the intervention is scaled up in a larger study. Alternatively, an intervention that appears to be highly effective in early small studies may have an apparently diminished intervention effect by the time it is subject to a large scale evaluation due to a system-wide improvement triggered by the same social pressure that prompted the intervention.[34] On the other hand, the availability of data from large databases covering nearly the whole population may render the influence of small studies negligible. These different types of heterogeneity arising from the complexity of HSDR interventions/associations and the context in which they are deployed/observed therefore pose a potential threat for the validity of applying these conventional methods. In addition, funnel plots and related regression methods require sufficiently large number of studies (e.g. ≥10), which may not be available for many topics in HSDR.

The previous work package (WP2) will allow us to obtain an overview of current practice of examining publication bias in systematic reviews of HSDR, including a description of whether and what methods have been used and the contexts in which existing methods are not applicable. This will help explain the apparently low utilisation of formal methods for detecting publication bias in HSDR reviews. Nevertheless, when formal methods such as funnel plots and related regression methods have been used, there remain potential issues concerning the validity and applicability of these methods. This Work Package (WP3) aims to address these issues through detailed case studies. In addition, WP3 also offers an opportunity to explore novel methods such as the p-curve for identifying p-hacking (details described later), which could be very relevant for HSDR.

Methods

Selection of cases to be studied

Cases to be studied will be selected from systematic reviews identified in either WP1 or WP2. Given that the purpose is to gain insight on potential issues of existing methods when they are applied to HSDR, we will purposively sample five to ten systematic reviews to ensure reasonable coverage of this diverse field. The selection of cases will be guided by the following considerations:

(1) The review includes sufficiently large number of studies (≥10) to meet minimal requirement for using funnel plots and regression methods;
(2) Covering reviews of various sizes (number of included studies)
(3) Including both reviews that evaluate intervention effectiveness and those investigating associations;
(4) The review is judged to be the most comprehensive and up-to-date systematic review for a chosen topic;
(5) The chosen cases cover the major issues and scenarios likely to be encountered during evidence synthesis of HSDR;
(6) The topics are of general interest for health services researchers, practitioners and the general public.

The following topics are provisional examples that will be considered:

- Effectiveness of interventions for improving hand hygiene
- Effectiveness of clinical decision support systems and patient outcomes
- Effectiveness of interventions for improving the quality of disease management for patients with type 2 diabetes
- Effectiveness of incentives for improving prescribing
- The association between days of hospital admissions and mortality
- The association between hospital volume and patient outcomes
- The association between hospital safety culture and patient outcomes
- The association between nurse-patient ratios and patient outcomes

The cases to be studies will be discussed with and chosen in consultation with both the Project Steering Committee and the Project Advisory Committee, taken into account the afore-mentioned criteria, possible saturation of issues and scenarios covered, and practicality within the project timeline.

Estimation of potential publication bias using funnel plots and related methods

In addition to data already collected under WP1 and/or WP2, detailed information on the methods and findings of each systematic review selected for case study will be extracted to a standardised form which will be used for all cases. Data from individual studies included in the review will be taken directly from the review. Original primary studies will be consulted only when judged necessary (e.g. when errors are suspected or when information presented in the review is not sufficient). Data extraction will be carried out by one reviewer and be independently checked by another review, with any discrepancies resolved by discussion or contacting review authors.

A systematic review of HSDR may evaluate several types of interventions, intervention components and features in different settings, or investigate a given association in different patient populations and context. In addition different measurements may have been used in the primary studies included in a review for a given outcome. We will record these details along with the review authors' approaches to deal with these potential sources of heterogeneity (e.g. whether studies/data were grouped and analysed separately or compared using subgroup or regression analyses etc). Review authors' judgement concerning the impact of the potential heterogeneity on the intervention effects or associations being estimated and on the assessment of publication bias will also be noted.

We will compile a funnel plot for each outcome where data are available from ≥10 studies and there is sufficient variation in their sample sizes/precision. Egger's regression tests will be used for continuous variables and the method proposed by Peters et al. (2006) will be used for binary variables to test funnel plot asymmetry as recommended by the Cochrane Collaboration.[35] Where funnel plot asymmetry is suspected (judged by results of above tests or by visual inspection of the plots given the relatively low statistical power of the tests), 'trim and fill' method [24] will be used to estimate the potential impact of small study effects. In the trim and fill method, the asymmetry of a funnel plot is assumed to be caused by publication bias, and alternative estimates correcting for the bias are calculated firstly by trimming out smaller studies with more extreme effect size estimates causing the

asymmetry, and then by re-introducing these studies along with their 'missing' counterparts. The method provides a way to estimate how sensitive the results of meta-analyses are to the small study effects. We recognise that publication bias is just one potential cause of small study effects and will interpret findings with caution accordingly.

Evaluating the association between estimated effect sizes and other potential effect modifiers

The regression methods used alongside funnel plots essentially test the existence of an association between observed effect sizes and sample sizes of studies. As described earlier, heterogeneity in the intervention components, study design, settings and context commonly seen in HSDR may confound this association. One approach to investigate this issue is to use multivariate meta-regression analyses including both sample size and potential confounding factors (e.g. quality of study, or year of publication as a proxy for changes in context) as covariates. If the association between observed effect sizes and sample sizes persists after adjusting for potential confounders, the likelihood of observed funnel plot asymmetry being caused by publication bias increases. Potential confounding factors to be included in the analyses will be determined case by case by the project team, but will be determined *a priori* before initial analyses are carried out. Any further exploratory analyses, if undertaken, will be clearly reported as *post hoc*. Other possible sources of bias include follow-up period; institutional characteristics and, perhaps above all, baseline performance into which many contextual variables are consolidated.

Investigation of p-hacking using p-curves

Repeating analyses using different analytical approaches and datasets until a statistically significant result is obtained – so called 'p-hacking', introduces a bias closely related to publication bias. Recently a novel methodology, termed 'p-curve', that allows the detection of p-hacking from published literature has been developed. [2] The method is based on the fact that, when null hypothesis is true, the distribution of p values is uniform and therefore should take the shape of a straight line when a collection of p values from studies that declare statistical significance are plotted. Where p-hacking exists, however, the distribution of p values will be distorted and a spike in the region just below p=0.05 would be observed.[3] The method has been tested using psychology and biology literature and demonstrated apparent p-hacking in these fields.[2,3] While we are not aware of the application of p-curve in health services research, p-hacking is a possible threat in HSDR, particularly in the increasing number analyses of datasets from routine databases. We therefore propose to use p-curves to explore the potential occurrence of p-hacking in HSDR within the selected case studies.

Work Package 4 (WP4): Follow-up of publication status of cohorts of health services research studies

The previous three Work Packages will draw crucial evidence on issues concerning the extent of publication bias and methods of detecting it in HSDR from the literature. Nevertheless, the evidence gathered is indirect in nature, as observations made (such as asymmetry in funnel plots and significant tests) are indicative of the existence of such bias rather than confirmatory. WP4 consists of a retrospective investigation of cohorts of HSDR studies, which will be followed over time to ascertain whether their publication status was associated with the statistical significance or perceived 'positivity' or interest of their findings. The main objective is to provide a direct observation of the presence or absence of publication bias in HSDR, as measured by the presence or absence of an association between the publication status of HSDR projects and the statistical significance/perceived 'positivity' (defined in further detail below) of their findings. In addition, if publication bias were observed, whether it is associated with study design, study type (intervention vs association) and/or sample size will also be explored.

Methods

Selection of study cohort

The only comprehensive database of UK HSDR studies that we are aware of is the project portfolio of the NIHR HS & DR funded projects, including those previously commissioned under the NIHR Service Delivery and Organisation (SDO) Programme and the NIHR Health Service Research (HSR) Programme. These studies have gone through a highly competitive bidding and selection process and are likely to be most well-funded projects among HSDR studies. In addition, the NIHR has a strong policy to mandate the publication of research findings and indeed the HS&DR Programme has been routinely publishing its funded studies that are submitted from July 2012 onwards in its Health Services and Delivery Research journal within the NIHR Journals Library. Studies included in the HS & DR database are therefore 'atypical' and are least likely to be subject to publication bias. Nevertheless, given the prominence of this portfolio of studies, evaluating the presence/absence of publication bias and documenting the impact of the establishment of the HS&DR journal series on the publication of these studies are both very important.

As of 1 July 2016, the project portfolio contains 525 funded studies, 358 of which have been completed. Allowing sufficient time for publication, we will select a random sample of at least 100 primary quantitative studies (i.e. excluding evidence synthesis based on the literature and studies that adopted exclusively qualitative methods) recorded as completed in or prior to 2012, and retrospectively follow them up to evaluate their publication status (described below).

In order to complement the cohort of studies funded by the HS&DR Programme, we will also identify a cohort of studies from the US-based HSRProj (Health Services Research Projects in Progress) database (https://wwwcf.nlm.nih.gov/hsr_project/home_proj.cfm). The HSRProj is currently hosted within the US National Library of Medicine and is the largest (and the only one that we are aware of) publicly accessible prospective registry of health service and public health research that covers multiple institutions and funding bodies. The database currently holds information on more than 29,000 projects (including both ongoing and completed) funded by over 350 agencies and dating back to 1999. While it is unlikely that projects registered with this database are representative of all HSDR, the coverage in terms of number of funders, years and types of studies make it the best alternative source to assemble a cohort of HSDR studies.

The HSRProj database classifies its project records as ongoing, completed or archived. Records are archived five years after the project's end date. We will take a random sample of at least 100 studies (which will provide confidence limits of under ± 10% assuming a publication rate of 60%; and an 88% power to detect a 20% difference between the two cohorts of HSDR studies) from the 1531 studies recorded as being completed in 2012 (to allow sufficient time for publication). As the HSRProj has a broad scope (e.g. including public health projects and comparative effectiveness research), studies that fall outside the scope of this proposed research will be excluded during the assembly of the study cohort but a list of excluded studies will be compiled and made available to public scrutiny.

Extraction of study information

Information on title, abstract, sample size, sponsoring agency and contact information of lead investigator for each selected study will be downloaded and imported into an Excel spreadsheet. Each study will also be classified according to study type (intervention vs association) and study design features (studies with concurrent controls vs others; method of data collection [bespoke vs routine data vs mixed]) independently by two reviewers according to the abstract initially and through contacting the investigator when needed. Any discrepancies between reviewers in the classification will be resolved by discussions, contacting investigators and/or referring to the wider project team as needed.

Verification of publication status

The publication status for each study will be verified firstly by searching PubMed and Google using information on title and lead investigator from the HS&DR project portfolio and HSRProj records.

Where no publication is identified, we will attempt to contact the investigators by e-mails (see Attachment 1) to verify the status of publication and to request information on published papers or unpublished study results, and reasons for non-publication if this is the case. Where no response is received, at least two reminders will be sent and other means (e.g. search of funding agency's website) may be pursued to enhance the completeness of follow up. Publication status for each study will be categorised as published (in academic journals), grey literature (available on the internet in a form other than an academic paper, such as a technical report or working paper) or unpublished.

Classification of study findings

Based on information available, each study will be classified according to statistical significance (with a p value ≤0.05 considered as statistically significant). First we will consider studies with one outcome or with one pre-specified primary outcome. Where results were reported for more than one outcome or association, studies will be further classified as 'all or mostly significant', 'mixed (≥1 significant result but for less than two-thirds of the outcomes/associations)' or 'all or mostly non-significant'. The same classification procedure involving at least two reviewers will be followed.

In addition to statistical significance, the findings of research studies have also been classified in various way such as being 'positive' vs 'negative', 'striking' vs 'less or not important' etc.[1] For example the findings of a study may be regarded as positive or favourable if a cheaper way to deliver a service is as effective as a more costly option (i.e. no significant difference in outcomes between the options). We will adopt the method used by Song et al. and classify the findings of each study as 'positive' or 'non-positive':[1] positive results include those that were considered (by the original study authors) as being 'positive', 'favourable', 'significant', 'important', 'striking', 'showed effect' and 'confirmatory'. Non-positive result refers to other results labelled as being 'negative', 'nonsignificant', 'less or not important', 'invalidating', 'inconclusive', 'questionable', 'null' and 'neutral'. The 'positivity' classification will be used only in a sensitivity analysis in place of 'statistical significance' given that the two measures are likely to be highly correlated.

Data analysis

Descriptive statistics for study type, study design, sample size, study findings and publication status will be computed. We will conduct a narrative synthesis and quantitative examination. The narrative synthesis will seek patterns in the data in a 'realist' mode to see whether there are emerging patterns that we had not considered in addition to those mentioned below. Univariate and multivariate logistic regression will be carried out to explore the association between publication status and statistical significance and other variables including funding source (no specific funding; local funding; national funding – HS&DR Programme; national funding - others), size of study (number of institutions; number of individuals) and design of study (concurrent control vs other; method of data collection [routine data, bespoke data collection, mixed]). The sizes of the studies will be grouped according to quarterlies for this purpose.

Work Package 5 (WP5): Semi-structured interviews with health services researchers and journal editors

This work package seeks to complement direct evidence assimilated from the retrospective cohort study in WP4 by exploring the perceptions and first-hand experiences of health services researchers and commissioners, journal editors, service managers and users with regard to the occurrence and impact of publication and related bias. It contributes to the overall aim of obtaining (qualitative) evidence on the extent and existence of publication bias. It will also contribute to the development of methods for the detection and mitigation of publication bias in HSDR. As well as generating important data on the perspectives of key actors in the HSDR process, this work package is designed to support analysis of results deriving from prior work packages.

Objectives

Interviews will be designed to:

1. Enable qualitative exploration of quantitative findings derived from WPs 1-4 for example in relation to current rates and types of publication bias in HSDR;
2. Gauge the views of a sample of those currently commissioning, publishing or conducting HSDR as to the prevalence or otherwise and perceived impact of publication bias;
3. Identify and explore current and future strategies for prevention, detection and mitigation of any bias detected, and
4. Explore the experiences and views of service managers and patient and user experts involved in HSDR.

Methods

Selection and recruitment of key informants

We will undertake in-depth interviews with 20 to 30 key informants in the field of health services research to explore their perceptions, experience and preferred solutions to overcoming problems associated with publication bias in HSDR. Key informants will be identified in consultation with the Study Steering Committee and Project Management Group and will be purposively sampled to include researchers reflecting different epistemological traditions such as improvement science, healthcare organisation and management, health psychology, systems research, health economy and policy, and health informatics. We will ensure that the sample includes researchers at various stages in their careers (e.g. from Research Fellow to Professor). Editors/assistant editors of key UK health services journals (e,g. BMJ Quality and Safety, Journal of Health Services Research and Policy) will be included, as will journal editors from outside the UK (example journals include: Milbank Quarterly, Social Science and Medicine, Healthcare Policy) as articles published in the latter outlets will also impact on evidence use in the UK. UK funders and patient/service user groups are included in the interview sample. Funders include: the Health Foundation, the Economic and Social Research Council, the Joseph Rowntree foundation, NIHR and The European Research Council. Examples of patient and user experts to be included are: Healthwatch England, CCG Patient & Public Involvement Lay Members Network, the Community Hospitals Association and Citizens UK. We will also include senior service managers to explore their awareness of publication bias in the evidence they use and how, if at all, this influences their decision making. We do not require that those included have a specific research interest in publication bias but instead will design our interview schedule to enable them to reflect on publication bias from their standpoints and experiences.

The interviews will be continued until a saturation of emerging themes is achieved or the maximum number (30) within the allocated resource is reached. The indicative numbers for each category of informants are shown below:

- Researchers at various stages in their careers (e.g. from Research Fellow to Professor): 10
- Editors/assistant editors UK: 4
- Journal editors from outside the UK: 4
- Research commissioners/funders: 5
- Patient/service user groups: 4
- Senior managers from commissioning and provider organisations: 3

Potential interviewees will be invited by the lead researcher for Work Package 5 (IW) via email in the first instance. The invitation email (see Attachment 2) will include a short summary of the project with more detail attached (i.e. participant information leaflet, Attachment 3). Those agreeing to take part will be requested to return a signed consent form (Attachment 4) either by email (scanned signed copy or provision of an electronic signature) or by post, in which case their postal address will be sought and pre-paid envelopes will be sent by the research team. Those declining to take part will be asked to give reasons for declining. Non-responders will be sent a reminder email within two week of the initial email, and this will be indicated in the initial invitation.

Assurances will be given that where possible all steps will be taken to ensure anonymity. It will however be made clear that in a relatively small sample of high profile interviewees full anonymity may be somewhat compromised. In acknowledgment of the sensitivity of the subject material we will put safeguards and assurances in place so that respondents feel able to speak freely and candidly. For example, we will assure interviewees that as well as anonymising transcripts, steps will be taken to ensure that any identifying details are redacted in any subsequent reports. Participants will be given the opportunity to comment on a draft report of this work package so that they can be assured that all identifying features are removed.

Data collection

Semi-structured interviews will be undertaken face-to-face or by telephone depending on interviewees' preference and accessibility. The structured questions for the interviews will be informed by emerging findings from other Work Packages and will focus on the informants' perception and past experience of publication bias in HSDR, and their opinions on possible approaches to mitigate this problem. The interviewer will then explore emerging issues in more depth based on the initial responses. An interview schedule can be found in Attachment 5. Permission to voice-record the interviews will be sought whenever possible, and recorded interviews will be fully transcribed for subsequent analysis.

Data analysis

Interviews will be subject to thematic analysis to identify emerging themes. Findings from earlier interviews will be used to inform subsequent interviews to facilitate the exploration of different perceptions, experiences and opinions among the interviewees. Data will be analysed inductively to gauge participants' perspectives and experiences within the framework provided by the research aims as well as issues identified in prior work packages. For internal validity all interviews will be fully transcribed and we will use qualitative coding software (NVIVO) to facilitate data storage and retrieval during analysis.[36] Two members of the research team will contribute to the building of thematic coding frames from qualitative data, and will share independent coding of a data subset in order to ensure consistency. Identified themes will then be discussed at meetings of the core project team. External validity and transferability of analysis will be addressed through detailed description and data-triangulation between work packages.[37]

Plan of investigation and timetable

The project is expected to start from January 2017 and last for 24 months. A Gantt chart is uploaded as an additional supporting document. The five Work Packages (WPs) will be rolled out in a staggered manner. WP1 (systematic review of relevant literature) will be undertaken during the first 5 months so that any initial findings can be used to inform the development and execution of the other WPs. Preparatory work for other WPs will also be started during this time. WP2 (overview of systematic reviews) and WP4 (retrospective cohort study) will commence in April and June 2017 respectively and will run in parallel for 10 months each. The selection of cases for detailed case study for WP3 needs to be informed by WP1 and WP2 and therefore it will commence later in August 2017. WP3 will also be spread over 10 months to allow initial case studies to be refined based on emerging results from other WPs. WP5 (interviews with health services researchers and journal editors) will commence in August 2017 and be spread over 12 months again to allow initial findings from the project and earlier interviews to inform later interviews. Additionally, we will hold two project dissemination event 3-4 months before the conclusion of the project, in which relevant stakeholders and interested public will be invited to comment on project findings and discuss practical implications and future recommendations. We have included 3 months towards the end of the project to allow final updating of literature-based WPs and 3 further months to integrate findings from all five WPs and discussions in the dissemination event into a coherent final report.

Research team expertise and project management arrangements

The project team consists of a core research & administrative team, a steering committee and a advisory committee. It brings together international experts with a wealth of practical and methodological expertise in HSDR, evidence synthesis, publication bias, and patient and public involvement (PPI). The team covers a wide range of perspectives and have been working together for many years. Team members will meet regularly through teleconference and face to face meetings, complemented by communications through emails.

The expertise and specific role of each team member is shown in Figure 6, with further details described in the online application form. University of Warwick finance department will administer the project budget which will be overseen by Dr Chen. A contract will be drawn up between Warwick and external partners.

Project team

Project management group
Prof Richard Lilford (University of Warwick) – HSDR methodology & senior management
Prof Russell Mannion (University of Birmingham) – HSDR methodology & mix methods
Prof Fujian Song (University of East Anglia) – expertise in publication bias & statistical advice
Dr Magdalena Skrybant (University of Birmingham PPIE lead) – patient and public perspective

Core research & administrative team
Dr Yen-Fu Chen (University of Warwick) – lead applicant, evidence synthesis & project management
Dr Iestyn Williams (University of Birmingham) – in-depth interviews (WP5)
Dr Abimbola Ayorinde (University of Warwick) – evidence synthesis, data collection, day to day running of project
Research administrator (University of Warwick) – administrative support

Study steering committee
Prof Stephen Sutton (University of Cambridge) – Chair of the committee
Revd Dr Barry Clark (PPI representative) – patient and public opinion & input
Dr Kaveh Shojania (University of Toronto) – HSDR methodology & journal editor perspective
Prof Jeremy Grimshaw (University of Ottawa) – HSDR methodology & evidence synthesis perspective
Prof Timothy Hofer (University of Michigan) – HSDR methodology & health professional perspective
Dr Christopher Chiswell (Birmingham Children's Hospital) – health professional and service provider perspective
Mr Tim Sacks (East Leicestershire & Rutland CCG) – service provider and management perspective
Prof James Thomas (University College London) – Evidence synthesis & policy perspective
Prof Richard Lilford (University of Warwick) – HSDR methodology & senior management

Figure 6. Composition and roles of the research team

Ethical considerations

Contacting health services researchers

In Work Package 4, we will retrospectively follow up two cohorts of previously registered HSDR studies to verify their subsequent publication status. For those studies for which we cannot identify any publications of study findings through PubMed and Google search, we will try to contact the investigator(s) and ask them to verify the publication status of their studies. As the main intention is to confirm the non-publication of HSDR studies, we intend to do this through a relatively short e-mail

(see Attachment 1) in order to maximise the response rate. We will invite the respondents to share the reasons for non-publication and any other comments that they might wish to offer, and will explain in the e-mail that their response is voluntary and any comments will be anonymised before they are used. We do not envisage any particular ethical issues with this approach, as the respondents can choose not to reply and have full control of the timing and contents of their replies if they choose to do so.

In the rest of this section we will explain the approach to ethical considerations for the key informant interviews to be undertaken in Work Package 5 of the project.

Informed consent for participants of the interview

When the initial approach is made to potential participants it will be made clear that we are looking for volunteers and there is no obligation to take part. The invitation to take part will include provision of an information sheet giving full details of what participation will involve. Consent forms will be completed at the beginning of each interview by each participant. Participants will be informed that they can withdraw from the study at any time up until the point at which the data are anonymised.

Participant confidentiality

Interview transcripts will be anonymised and names and details of interviewees will be stored separately from associated data. As we will interview a relatively small sample of interviewees, some of whom may have a high profile, we cannot guarantee that respondents will be entirely unidentifiable. In recognition of this, participants will be given the opportunity to comment on a draft report of the interview findings to help ensure that identifying features are removed.

Data security

Research information – including data, consent forms and administrative records - will be kept in a password-protected file saved on a secure University server for a period of 10 years with access limited to key study personnel. Paper records will be stored in locked cabinets at the Health Services Management Centre, University of Birmingham and shredded following submission of the final NIHR report. In all outputs from the research, participants will be referenced via a unique identifier code only.

Right of Withdrawal

Participants will be informed that they may withdraw from the study at any time up until the point at which the data from the interview are anonymised, beyond which time we will be unable to disaggregate data.

Process for dealing with sensitive disclosures

The study is not intended to involve discussion of sensitive or potentially sensitive topics. We are aware that under-reporting (including non-publication) of research has been described as “scientific misconduct” and discussion of such issues may be considered as sensitive disclosures. Based on literature in biomedical research, however, we know that under-reporting and non-publication of research could arise for many different reasons, some of which may be seen as poor academic practice but most would not be construed as misconduct and/or as being serious enough to warrant further actions to be taken. We therefore intend ***not*** to pre-occupy interviewees with such a concern in order to encourage open discussions of relevant issues. However we will assure interviewees that if they mention any identifying information (e.g. names, institutions and job titles), it will be removed before the data is included within the study.

Benefits and risks for participation

There are few if any risks in taking part in this study as all participants are required to do is share thoughts and experiences on a matter related to research practice.

Given the potentially significant impact of publication bias on decisions made on the organisation and delivery of health services and the paucity of existing evidence in this field, the findings will help us to understand the nature and scale of the problem in HSDR and to develop recommendations to help prevent publication bias and/or minimise its impact. We will offer to share early findings with participants.

Approval by ethics committees

We will seek formal ethical approval for the retrospective cohort study (WP4) and key informant interviews (WP5) from the University of Warwick Biomedical & Scientific Research Ethics Committee (BSREC). We will in particular seek input from our PPI representatives who can discuss and advise the wider ethical implications of potential findings of the project. An application for ethics will be submitted at least three months before the start date of the Work Packages. It will normally take the committee 30 days to review the ethics application and provide approval. Given that senior managers from commissioning and provider organisations within the NHS will be included in the interviews, we have registered the study with the Integrated Research Application System (IRAS) and have clarified with NHS Health Research Authority (HRA), who advised us that an HRA approval is not required for undertaking the research provided that our interviewees are not identified through an NHS organisation.

Patient and public involvement

We will include two PPI representatives in the project team. In addition to helping design the project from the start, their involvement will be continued throughout the life of the project in several ways: (1) representing patients and the public in the project steering and advisory committees; (2) reviewing and commenting on project documents such as review protocols and report drafts; (3) leading on the production of plain English summaries; (4) planning and attending project dissemination events.

PPI will be embedded in the project and there will be patient and public representation (Revd Dr Barry Clark and Ms Magdalena Skrybant) on project advisory and steering committees. We have taken this approach because publication bias is an issue that, while not confined to specific patient groups, is relevant to all. Both PPI representatives are involved in activities associated with CLAHRC West Midlands (WM) and therefore are familiar with research process associated with publication bias. They will have the opportunities to read and comment on important project documents such as detailed protocols for individual work packages and draft reports and their involvement will allow a fresh and independent voice to be heard at the highest level.

We will also work with PPI representatives to produce a plain English summary for each of the work packages and disseminate these through various outlets and in a range of formats such as the project website and relevant circulation lists (e.g. CLAHRC WM and Health Services Management Centre) to ensure that the findings reach researchers and managers across the NHS and are accessible for patients and the public. We plan to hold two dissemination events towards the end of the project to allow a full range of stakeholders including those who commission and use HSDR and interested researchers, patients and the public to share their views and discuss practical implications for future practice and research. Our PPI representatives will play an important role in planning and participating in these events.

Pathway to impact

We believe that findings from this work could have far-reaching impact both within and beyond the HSDR community. The pathways from research findings to their ultimate impact are summarised in Figure 7. Findings from individual work packages will initially be communicated to the HSDR community through presentations in conferences and publications in academic journals. Where

feasible, we will take the opportunities of attending HSDR conferences to run workshops in which the findings of the project will be presented and feedback from HSDR researchers sought. For each work package, we plan to produce a lay summary of the findings in partnership with our PPI representatives. We will distribute these summaries through various channels such as CLAHRC West Midlands' news blogs, Health Services Management Centre's distribution list and the project's web page to be hosted on the University of Warwick's web site to increase the reach beyond academic researchers. In addition, three to four months prior to the end of the project we will host two one-day dissemination events to provide further opportunities to engage with various target audiences. The dissemination events will target health services researchers, journal editors, HSDR funders, service managers and health care professionals, patients and service users but will also be open to interested members of the public. In addition to presenting findings from the project in these events, we will discuss preliminary recommendations for future research and practice with participants of the dissemination events to ensure the acceptability and comprehensiveness of the recommendations made. The project findings along with refined recommendations will be collated in our final report to the HS&DR Programme and published in the NIHR Journals Library.

Offering practical guidance on preventing and detecting publication bias and minimising its impact plays a crucial role in the pathway to impact. We will formulate preliminary recommendations covering the following items:

- What is publication bias and why is it important?
- What is known about publication bias in HSDR?
- When should we worry about publication bias?
- How can publication bias be prevented during the generation of research evidence?
- How can publication bias be minimised during the publication / dissemination of research evidence?
- What tools can/should be used for detecting and/or correcting publication bias?
- What to do when interpreting and using evidence in the face of potential publication bias?
- What are gaps in our current knowledge concerning publication bias in HSDR that warrants further research?

The coverage of the items, the depth of the contents and the use of language (technical vs plain) will be tailored for each target audience, which includes researchers who generate new evidence, researchers who critically appraise and synthesise evidence, research funders, journal editors, and evidence users (service managers and health professionals; patients, carers and the public). Recommendations will firstly be drafted by the core research team based on findings from individual work packages. Comments and feedback will be sought from the project steering and advisory committees. The revised recommendations will then be discussed and further feedback sought during the project dissemination events before final recommendations are drawn as described above.

It is hoped that the multiple channels of dissemination, opportunities to engage and practical recommendations for a wide variety of audiences will help improve the processes for research commissioning, monitoring and publication to minimise publication bias. This in turn should lead to better decisions informed by unbiased evidence and ultimately more efficient delivery of effective services, which are the ultimate impact that this project aims to achieve.

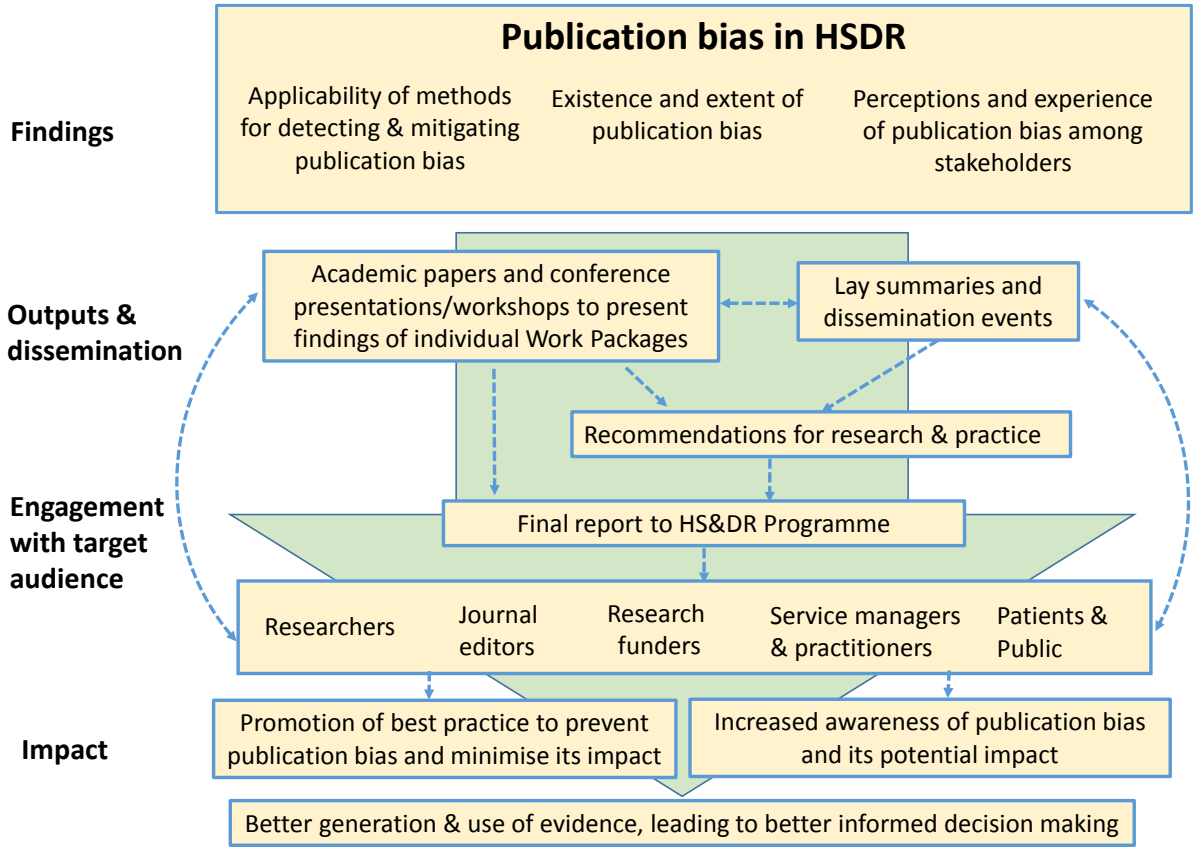


Figure 7. Pathway to impact for this project

References

1. Song F, Parekh S, Hooper L, et al. Dissemination and publication of research findings: an updated review of related biases. Health Technol Assess 2010;**14**(8):1-193.
2. Simonsohn U, Nelson LD, Simmons JP. P-curve: a key to the file-drawer. J Exp Psychol Gen 2014;**143**(2):534-47.
3. Head ML, Holman L, Lanfear R, Kahn AT, Jennions MD. The extent and consequences of P-hacking in science. PLoS Biology 2015;**13**(3):e1002106.
4. Dwan K, Gamble C, Williamson PR, Kirkham JJ, the Reporting Bias G. Systematic review of the empirical evidence of study publication bias and outcome reporting bias — an updated review. PLoS ONE 2013;**8**(7):e66844.
5. Kicinski M, Springate DA, Kontopantelis E. Publication bias in meta-analyses from the Cochrane Database of Systematic Reviews. Stat Med 2015;**34**(20):2781-93.
6. Gulmezoglu AM, Pang T, Horton R, Dickersin K. WHO facilitates international collaboration in setting standards for clinical trial registration. Lancet 2005;**365**(9474):1829-31.
7. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International Journal of Social Research Methodology 2005;**8**(1):19-32.
8. Knight K, Badamgarav E, Henning JM, et al. A systematic review of diabetes disease management programs. The American Journal of Managed Care 2005;**11**(4):242-50.
9. Shojania KG, Ranji SR, McDonald KM, et al. Effects of quality improvement strategies for type 2 diabetes on glycemic control: a meta-regression analysis. JAMA 2006;**296**(4):427-40.
10. Shojania KG, Ranji SR, Shaw LK, et al. Closing the quality gap: a critical analysis of quality improvement strategies (Vol. 2: Diabetes Care). Technical Review No. 9. In: Shojania KG, McDonald KM, Wachter RM, Owens DK, eds. Rockville (MD): Agency for Healthcare Research and Quality (US), 2004.
11. Stone EG, Morton SC, Hulscher ME, et al. Interventions that increase use of adult immunization and cancer screening services: a meta-analysis. Ann Intern Med 2002;**136**(9):641-51.
12. Costa-Font J, McGuire A, Stanley T. Publication selection in health policy research: the winner's curse hypothesis. Health policy 2013;**109**(1):78-87.
13. Gemmill MC, Costa-Font J, McGuire A. In search of a corrected prescription drug elasticity estimate: a meta-regression approach. Health Economics 2007;**16**(6):627-43.
14. Costa-Font J, Gemmill M, Rubert G. Biases in the healthcare luxury good hypothesis?: a meta-regression analysis. Journal of the Royal Statistical Society: Series A (Statistics in Society) 2011;**174**(1):95-107.
15. Machan C, Ammenwerth E, Bodner T. Publication bias in medical informatics evaluation research: is it an issue or not? Studies in Health Technology and Informatics 2006;**124**:957-62.
16. Vawdrey DK, Hripcsak G. Publication bias in clinical trials of electronic health records. Journal of Biomedical Informatics 2013;**46**(1):139-41.
17. Ammenwerth E, de Keizer N. A viewpoint on evidence-based health informatics, based on a pilot survey on evaluation studies in health care informatics. J Am Med Inform Assoc 2007;**14**(3):368-71.
18. Franco A, Malhotra N, Simonovits G. Publication bias in the social sciences: Unlocking the file drawer. Science 2014;**345**(6203):1502-05.
19. Torgerson CJ. Publication bias: the Achilles' heel of systematic reviews? British Journal of Educational Studies 2006;**54**(1):89-102.
20. Banks GC, Kepes S, Banks KP. Publication Bias: The Antagonist of Meta-Analytic Reviews and Effective Policymaking. Educational Evaluation and Policy Analysis 2012;**34**(3):259-77.
21. Harrison JS, Banks GC, Pollack JM, O'Boyle EH, Short J. Publication bias in strategic management research. Journal of Management 2014:doi: 10.1177/0149206314535438.

22. Kepes S, Banks GC, McDaniel M, Whetzel DL. Publication bias in the organizational sciences. Organizational Research Methods 2012;**15**(4):624-62.
23. Jennions MD, Moller AP. Publication bias in ecology and evolution: an empirical assessment using the 'trim and fill' method. Biological reviews of the Cambridge Philosophical Society 2002;**77**(2):211-22.
24. Lilford RJ, Chilton PJ, Hemming K, Girling AJ, Taylor CA, Barach P. Evaluating policy and service interventions: framework to guide selection and interpretation of study end points. BMJ 2010;**341**:c4413.
25. Brown C, Hofer T, Johal A, et al. An epistemology of patient safety research: a framework for study design and interpretation. Part 2. Study design. Qual Saf Health Care 2008;**17**(3):163-69.
26. Brown C, Hofer T, Johal A, et al. An epistemology of patient safety research: a framework for study design and interpretation. Part 4. One size does not fit all. Qual Saf Health Care 2008;**17**(3):178-81.
27. Petticrew M, Egan M, Thomson H, Hamilton V, Kunkler R, Roberts H. Publication bias in qualitative research: what becomes of qualitative research presented at conferences? J Epidemiol Community Health 2008;**62**(6):552-4.
28. Lewin S, Glenton C, Munthe-Kaas H, et al. Using qualitative evidence in decision making for health and social interventions: an approach to assess confidence in findings from qualitative evidence syntheses (GRADE-CERQual). PLoS Medicine 2015;**12**(10):e1001895.
29. Rothstein HR, Sutton AJ, Borenstein M, editors. *Publication bias in meta-analysis: prevention, assessment and adjustments*. Chichester, England: John Wiley & Sons, 2005.
30. Li X-X, Zheng Y, Chen Y-L, Yang K-H, Zhang Z-J. The reporting characteristics and methodological quality of Cochrane reviews about health policy research. Health policy 2015;**119**(4):503-10.
31. Worswick J, Wayne SC, Bennett R, et al. Improving quality of care for persons with diabetes: an overview of systematic reviews - what does the evidence tell us? Systematic Reviews 2013;**2**:26.
32. Wilczynski NL, Haynes RB, Lavis JN, Ramkissoonsingh R, Arnold-Oatley AE, The HSRHT. Optimal search strategies for detecting health services research studies in MEDLINE. CMAJ : Canadian Medical Association Journal 2004;**171**(10):1179-85.
33. AcademyHealth. Health services research and health policy grey literature project: summary report. 2006. https://www.nlm.nih.gov/nichsr/greylitreport_06.html Accessed 29 November 2015.
34. Chen YF, Hemming K, Stevens AJ, Lilford RJ. Secular trends and evaluation of complex interventions: the rising tide phenomenon. BMJ Qual Saf 2015:doi:10.1136/bmjqs-2015-004372.
35. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Comparison of two methods to detect publication bias in meta-analysis. JAMA 2006;**295**(6):676-80.
36. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013;**13**:117.
37. Lincoln YS, Guba EG. *Naturalistic Inquiry*. Newbury Park, California: Sage Publications, 1985.