Improving literature searching in systematic reviews:

the application of tailored literature searching compared to 'the conventional approach'

Volume 1 of 2

Submitted by Chris Cooper to the University of Exeter

as a thesis for the degree of

Doctor of Philosophy in Medical Studies

In December 2017

Signature: Chris Cooper

This thesis and its supporting publications were submitted to the University of Exeter on Friday December 8th 2017.

Publication two (reported in Chapter 6) is subject to an embargo of 12 months from the date of its publication online (March 9th 2018). As the accepted but pre-publication manuscript of publication two forms part of this thesis, the thesis as a whole is embargoed until March 10th 2019.

Date of submission (date embargo begins): December 8th 2017

Date Embargo is lifted: March 10th 2019

Following the viva on March 23rd 2018 this, the corrected version of the thesis, was submitted on Monday, October 22nd 2018.

Abstract

Background

Literature searching is acknowledged as a crucial step in a systematic review. Information professionals, in response to the needs of intervention effectiveness systematic reviews, have developed a systematic process of literature searching which aims to be comprehensive, transparent and reproducible, and to minimise the introduction of bias in systematic reviews. The process which has evolved has not been examined in detail before but it has been adopted as the principal approach to literature searching in other types of systematic review. It is not clear if this is appropriate and if an alternative approach might be more effective.

Thesis aims

The aims of this thesis are to:

1) examine approaches to systematic literature searching for systematic reviews; and

2) propose and test a method of systematic literature searching for reviews which do not focus on the effectiveness of clinical interventions.

Methods

Two literature reviews, one systematic review and two comparative case studies were undertaken to meet the aims of the thesis.

Results

A critical literature review identified and described a conventional approach to literature searching common to nine leading handbooks of systematic review. An alternative, tailored approach to literature searching was developed. Two case studies illustrated that the tailored approach was more effective, and potentially offered better value, than the conventional approach.

Conclusions

Information professionals can develop tailored literature search approaches for use in systematic reviews and as a useful alternative to the conventional approach, particularly for reviews including study designs beyond controlled trials. The role of

the information professional as decision maker, the involvement of the research team and experts, preparing for literature searching and the use of supplementary search methods, are important to the success of tailored literature search approaches.

Table of Contents

List of Tables

List of Figures

Supplementary materials

Reported in volume 2 of this thesis.

Glossary

Case study/studies: A case study is a research method commonly used in social sciences but there is no clear or common methodological definition of what constitutes a case study (1, 2). Broadly, it is an in-depth exploration of a single example case to try to develop an understanding of key issues (1-3).

In this thesis, I present two case studies to compare the conventional approach to the tailored approach. The 'case' evaluated in each of these two examples is the process of literature searching within two different systematic reviews. I undertake a systematic and in-depth evaluation in order to explore the effectiveness and, where possible, the value and efficiency of two approaches to literature searching within the specific setting of these two systematic reviews.

Conventional approach to literature searching: Identified and defined in Chapter 3, the conventional approach is an eight-stage process of literature searching (Figure 1), common to nine leading guidance documents on the process of systematic reviews (Table 3). The conventional approach begins with comprehensive searches of bibliographic databases, followed by supplementary search methods. The same configuration of search methods (databases followed by supplementary search methods) is used irrespective of the research question posed. The aim of the conventional approach is for a comprehensive and transparent identification of relevant studies, aiming to minimise the introduction of bias in study identification (Chapter 3).

Information need: Kuhlthau defines this as 'The gap between the users' knowledge about the problem or topic and what the user needs to know to solve the problem' (4).

Information professional: I use this term to indicate the following professional groups: Information specialists, Information scientists, and librarians.

Research waste: In this thesis, I define research waste as wasting researcher time and resources by over-screening literature search results. I am not aware of a definition of research waste in the context of literature searching and so I use this understanding throughout.

Sensitivity: Sensitivity, in a literature search context, refers to the proportion of studies correctly identified as relevant, relative to the total number of relevant studies that may exist (see Chapter 6) (5).

Specificity: Specificity, in a literature search context, refers to the number of irrelevant studies excluded or not identified by a literature search (Chapter 6) (5).

Supplementary search methods: Supplementary search methods are non-database methods of literature searching, for example: contacting study authors, citation chasing, handsearching or web searching (see Chapter 5) (6).

Tailored literature searching: Tailored literature searching aims to 'tailor' the selection of literature search methods to the research question (or information need(s)), the likely location of the evidence and the publication status of study data. It proposes that the order and extent to which literature search methods are used should be decided by the research team and the information need(s) of the review. Tailored literature searching represents an alternative approach to literature searching when compared to the conventional approach to literature searching and it is what this thesis seeks to explore. I set this out in Chapter 4.

Declarations relating to the candidate's publications and his role in the publications

In reference to 'Chapter 11 - Presentation of theses/dissertations for degrees in the Faculty of Graduate Research: statement of procedures' (published July 2018: accessed August 2018), specifically sub-section 2.2.3, I acknowledge that the thesis that I present includes published material.

I submit four publications as a part of this Ph.D thesis. I set out in Table 1 (below) where these publications are reported in the thesis, the citation, my contribution to the publication and the publication status (i.e. published or accepted).

Pursuant to the examiners' report (received April 23rd 2018), and in reference to sub-section 2.2.5, I heed the examiners' guidance with reference to copyright. I submit the accepted and final pre-publication copy of each of the papers in Microsoft Word format. This ensures that tables and figures are presented in situ within the papers and the broader referencing and page numbering is consistent through-out the thesis. I am grateful for the guidance of the University of Exeter Library and Postgraduate office on this matter.

Nota Bene

Two points are noted here in this corrected version of the thesis:

i) Two of the publications (marked in green and with an asterisk in Table 1) are subject to a publisher embargo of one year from the date that the article was formally published online in its final and fully citable form[1]. Since this thesis includes these publications it, in effect, embargos the thesis too.

The last paper to be accepted was published on-line on March 9th 2018. The embargo on this thesis is therefore lifted on March 10th 2019. I am grateful for the guidance of the University of Exeter Library and Postgraduate office on this matter.

ii) The December 8th 2017 submission of this thesis included an additional publication, a literature review, entitled 'Defining the process to literature searching in systematic reviews: a literature review of guidance and supporting studies.' At the

---

[1] The systematic review published in the journal of clinical of epidemiology was the last paper to be published of the four publications that I submit as part of this corrected thesis. This journal's guidance on embargo periods therefore governs this thesis. The journal of clinical epidemiology is published by Elsevier. Elsevier state an embargo period of 12 months from the date of on-line publication. See: https://www.elsevier.com/journals/journal-of-clinical-epidemiology/0895-4356/guide-for-authors (accessed 27th August 2018)

time of submission, and of the viva, this publication was under review with the journal BMC Medical Research Methodology.

On the advice of the examiners (in the viva and in the examiners' report (18/04/18)), I have amended, and I have further developed, the work presented in that literature review. This revised work now exists as Chapter 3 in this, my corrected thesis, per their guidance.

I would note, however, that the literature review, as submitted in the December 2017 version of this thesis, was accepted by the journal, as submitted to the examiners in 2017, and it was published shortly after the viva (7).

Table 1 Publications included in this thesis

Publication	Where reported in the thesis	Citation	The candidate's contribution to the publication	Publication status
1	Section 1, Chapter 5.	Cooper C, Booth A, Britten N, Garside R. A comparison of results of empirical studies of supplementary search techniques and recommendations in review methodology handbooks: a methodological review. *Syst Rev*. 2017;6(1):234.	CC conceived, designed and undertook the study. AB, NB and RG provided comments on, and discussed, the study in draft as part of CC's Ph.D supervision. All authors have approved this manuscript prior to submission.	Published See: https://systematicreviewsjournal.biomedcentral.com/articles/10.1186/s13643-017-0625-1
2*	Section 1, Chapter 6.	Cooper C, Varley-Campbell J, Booth A, Britten N, Garside R. Systematic review identifies six metrics and one method for assessing literature search effectiveness but no consensus on appropriate use. *J Clin Epidemiol*. 2018;99:53-63.	CC conceived, designed and planned the systematic review. CC was the first reviewer and JVC the second. CC wrote the first draft of the review. JVC, AB, NB and RG provided comments and feedback prior to submission. All authors have approved this manuscript prior to submission.	Published See: https://www.sciencedirect.com/science/article/pii/S0895435617313318
3	Section 2, Chapter 7.	Cooper C, Levay P, Lorenc T, Craig GM. A population search filter for hard-to-reach populations increased search efficiency for a systematic review. *J Clin Epidemiol*. 2014;67(5):554-9.	CC conceived the idea for and design of this study and wrote the population search filter which is evaluated in this publication. CC wrote the first draft of the manuscript with input from PL. CC and PL revised the manuscript with input from TL and GC. All authors have approved this manuscript prior to submission.	Published See: https://www.jclinepi.com/article/S0895-4356(13)00511-8/fulltext
4*	Section 2, Chapter 8.	Cooper C, Lovell R, Husk K, Booth A, Garside R. Supplementary search methods were more effective and offered better value than bibliographic database searching: A case study from public health and environmental enhancement. *Res Synth Methods*. 2018;9(2):195-223.	CC conceived, designed and undertook the literature searching protocols evaluated in this publication. CC undertook all of the analysis reported in the manuscript and he wrote the first draft of manuscript with input from RL and KH. CC, RL, KH, AB and RG revised the manuscript prior to submission. All authors have approved this manuscript prior to submission.	Published See: https://onlinelibrary.wiley.com/doi/abs/10.1002/jrsm.1286

Key: CC = Mr Chris Cooper (the candidate); AB = Dr Andrew Booth (CC's Ph.D supervisor); NB = Prof. Nicky Britten (CC's Ph.D Supervisor); RG = Ruth Garside (CC's Director of Studies); JVC = Dr Joanna Varley-Campbell (colleague of CC); PL = Mr Paul Levay (colleague of CC); Dr Theo Lorenc (colleague of CC); Prof. Gillian Craig (colleague of CC); BL = Dr Rebecca Lovell (colleague of CC) and KH = Dr Kerryn Husk (colleague of CC)

Acknowledgments

I would like to dedicate what follows to Danica, Timothy, and now Sefton.

My thanks to my supervisors: Ruth Garside, Nicky Britten and Andrew Booth.

My thanks to Theo Lorenc, Alan Gomersall, Isaac Marrero-Guillamón, Alison O'Mara and Joanna Varley-Campbell.

I thank Chris Hyde and Martin Hoyle for their support.

I acknowledge the contribution and support of my co-authors for the publications I include in this thesis (and where not named above or below): Paul Levay, Gill Craig, Becca Lovell and Kerryn Husk.

I would like to thank: Ute Schauberger, Sophie Robinson, Jo Thompson Coon, Morwenna Rogers, Alison Bethel, Simon Briscoe, Juan Talens-Bou, Becca Abbott, Becky Whear, Noreen Orr, Sue Whiffin, Jenny Lowe, Linda Dumchen, Stella Taylor, Julia Frost, Sophie Dodman, Becky Hardwick, Tom Monks, Carol Lefebvre, Eleanor Kotas, Pam Gurney, Antonietta Medina-Lara, Suzy Paisley, Rumona Dickson, Mariana Bacelar, Tristan Snowsill, Mary Bond, Damien Hackney, Helen Coelho, Victoria Young, Mathew and Linsey Hussey, Andrew Huntbatch, and Joni O'Sullivan.

Finally, I would like to acknowledge my examiners: Prof. Stein, Prof. Dickson and Prof. Brettle. I thank them for their investment in this Ph.D thesis.

Reflective Statement

Information professionals who have pursued information retrieval research at a Doctorate level either undertake theoretical work with test data sets or examine practical, work- based retrieval problems; the latter are very much in the minority. The literature and evidence available on practical information retrieval problems is limited when compared to the evaluation of other processes in systematic reviews, and it is fragmented and uneven reflecting the fact that much library and information retrieval research is unfunded and undertaken by researchers, as time permits. I note that the work presented in this thesis, and published elsewhere, represents a small contribution to a new and fledging field of study.

I approached this thesis in 2015 with approximately six years working as an information specialist. My work in these preceding years had been broad in scope. It had encompassed literature searching for reviews in topic areas such as: public health guidance, Health Technology Assessment (HTA), terrorism and warfare, social care, and education, to name a few. In this time, I had worked on systematic reviews of intervention or cost effectiveness, reviews to inform public health guidance, theory driven and Realist Reviews, diagnostic and prognostic reviews, scoping reviews and rapid evidence assessments, and qualitative evidence synthesis.

I was struck then, as I am today, by the diversity of necessary experience required to do principally the same thing across different types of review and in different topic areas. The aim of literature searching, to identify relevant studies from non-relevant studies, and to report the process in a clear and transparent way, seldom changed, but there were always a variety of challenges to face and overcome which made each individual review a distinct piece of work.

There is limited guidance to aid the information professional in their work with much of the instrumental work of literature searching being tacit knowledge and learned through experience or from colleagues. The Master of Arts (MA), the prevailing benchmark qualification of the information profession, taught me nothing of the information retrieval field I worked in and that I explore in this thesis. Similarly, training in research methods, and academic research, was not included in the syllabus, perhaps reflecting that the MA is an arts not science-based degree.

My first professional post was at King's College, London. I worked in a small team of five, very experienced researchers. It was here that I learnt to literature search.

Perhaps most importantly, I also learnt to systematic review, to perform basic meta-analysis and build elementary decision models. I believe that this basic grounding in the process and methods of systematic reviews beyond literature searching were crucial to my development as an information professional and it greatly affected my understanding of how-to literature search. It is only after screening 30,000 studies for a systematic review, that you can see and experience the impact made by literature searching. There is no substitute for this practical experience.

The work I present in this thesis reflects the upbringing that I set out above and it reflects my personal journey through the practical information problems I faced in my academic career. This reflection seeks to explain reference to other projects, other citations of my published research or conference presentations, the development and focus of the publications included in this thesis, and the nature of the critique being based on practice and detailed experience. I, necessarily, draw from the tacit knowledge gained through a ten-year career as an information professional and I draw from the published and unpublished literature which forms the evidence-base for this field of study. This also means that the account that follows often uses a first-person narrative to describe what has been done.

The structure of this thesis

I present the work in four sections with a sub-structure of 11 chapters. Each of the sections is introduced and summarised at the start of a new section. The sections are broadly conceived as:

Section One: definitions and preparatory work. In this section, I set out the historical context of literature searching in systematic reviews; I set out the aims of the thesis and its 9 objectives, and I define what I mean by 'the conventional approach' and 'the tailored approach'. I then set out the preparatory work needed to undertake and understand the case studies, namely; a review of supplementary search methods and a review of methods or metrics used to evaluate the effectiveness of literature searching;

Section Two: the case studies. I present two case studies which explore the suitability of the conventional approach compared to the tailored approach. Two different systematic reviews are presented as cases for this work;

Section Three: I set out how I have addressed the aims and objectives of this thesis and I address the issues arising for discussion. My conclusions, contributions to knowledge and areas for future research are reported in this section; and

Section Four: the bibliography.

The structure of the chapters within the sections is broadly similar throughout the thesis.

In each chapter I set out:

- the background to the chapter;

- how the chapter addresses the aims of the thesis;

- the objective(s) of the chapter;

- the publication or work of the chapter (if applicable);

- chapter findings;

- how the chapter contributes to addressing the aims of the thesis; and

- the implications for practice of the work that I present.

The aim of this structure is to clearly set out the aim(s) and objectives of each individual chapter, to present the work of the chapter, and to then situate the work

presented in the context of the thesis, and to ensure that contributions to knowledge that arise, and relevant critique, are clearly presented.

Tables and figures are used throughout the thesis in the usual way.

Section One: Definitions and Preparatory Work

In Section One of this thesis, I set out and I explore the definitions that I will draw upon throughout this thesis. I also set out the preparatory work that is necessary to develop the tailored approach and also to make the comparison with the conventional approach.

In **Chapter 1**, I set out the introductory context within which the work of this thesis is situated. I take a historical perspective to set out and explore the evolution of literature searching in systematic reviews.

In **Chapter 2**, I set out the aims of this thesis. The aims are broken down into research questions, research objectives and research methods. I also indicate where the work is located in the thesis.

In **Chapter 3,** I set out the first definition. I introduce, I define and I critique **the conventional approach**. In this chapter, I test my hypothesis that leading guidance documents to the process of systematic reviews advocate the same process to literature searching. The consequence of this is that researchers may be using the same process of literature searching to identify studies for systematic reviews of clinical effectiveness as for qualitative evidence synthesis. I question if this is suitable.

In **Chapter 4,** I set out the second definition**.** I address the question: **what is tailored literature searching?** I set out the issues that led me to consider this question and I set out how tailored literature searching might work in theory.

I then turn my attention to the preparatory work that is necessary to develop and explore my ideas. There are two:

i) a literature review of supplementary search methods

Tailored literature searching draws heavily on supplementary search methods and so it was necessary to develop my understanding of these methods. I identified a number of primary studies reporting evaluations of supplementary search methods but I did not identify a review of supplementary search methods. This represented a gap in research knowledge.

The literature review to address this gap in research knowledge is set out in **Chapter 5** and it is the first publication presented in this thesis.

ii) how is effectiveness in literature searching measured and evaluated

To determine the suitability of the conventional approach compared to the tailored approach it was necessary to understand how researchers measure effectiveness and evaluate it in literature searching. In **Chapter 6**, I undertook a systematic review of metrics and methods used to determine effectiveness in literature searching. This systematic review not only identified a gap in knowledge, namely that there was a need to codify the terminology and methods used to determine effectiveness and report this in one review but also, I respond to the call of other researchers to develop the reporting of effectiveness evaluation. At the end of Chapter 6, I set out how I will compare the conventional approach to the tailored approach, which directly addresses the research aims of this thesis.

These two chapters (Chapter 5 and Chapter 6) informed the case studies I present in **Chapter 7** and **Chapter 8**.

1 Introduction

The aim of this introductory chapter is to set out the history of literature searching as it relates to systematic reviews. I take a historical perspective to briefly describe how and why the research community transitioned from an arguably ad-hoc and often unreported process of identifying studies to a systematic, comprehensive, transparent and replicable process of literature searching.

The purpose of this chapter is to situate this thesis in context. By setting out the development of the methods, technology, and a process to literature searching in systematic reviews, it will help situate the issues that I will examine in this thesis and it will identify the problems that I seek to explore.

The history of literature searching goes hand-in-hand with the development of systematic reviews. I will set out below what I contend are the key elements in the evolution of systematic literature searching. Through this process, I identify some challenges which have arisen and, to some extent, remain in 2018. I will highlight these challenges in this section but address them in subsequent chapters of this thesis.

By way of definition, literature searching in systematic reviews is the systematic process of study identification. The origin of this systematic process can be pin-pointed to 1992 and the work of The Cochrane Collaboration and its systematic reviews of intervention effectiveness. As I will now demonstrate, the work undertaken by The Cochrane Collaboration in the 1990s has since been a major influence on how researchers conduct literature searches in systematic reviews (8).

1.1 The history of literature searching in systematic reviews

1.1.1 Pre-1992

There is no singular or perceptible moment when the idea of literature searching for studies began. Identifying studies or 'evidence', and referencing them to advance learning, has been manifestly a part of science and medicine since the advent of the scientific revelation in the 16th and 17th century. The format for reporting has advanced, from letters and self-published pamphlets then, to scientific journals and reports now, and with many corresponding advances in methodology and reporting. The technology has advanced too, beyond recognition (8, 9). From organised card-catalogues, to CD-Roms, to on-line platforms searchable by anyone with an internet connection (8).

The methods and techniques of literature searching in 1980 and now (2018) are, however, broadly similar. The bibliographic database MEDLINE, in 1983, was regarded as 'state of the art' due to 'excellent controlled vocabulary' (10), a view unchanged now although seldom articulated. MEDLINE is commonly the first bibliographic database of reference when developing a literature search for bio-medical systematic reviews and it is where a systematic search strategy is commonly developed.

Supplementary search techniques, such as handsearching and contacting study authors, were also common-place in identifying relevant studies. Some, such as Eugene Garfield's citation analysis (also known as bibliometrics and now the functionality that makes electronic citation chasing possible), were under development and their potential was, at the time (1972), unrealised (11).

Whilst the literature search methods and techniques then were similar to now (12), literature searching was a largely unsystematic and ad-hoc process, the methods and results of which were often unreported in research (13). In the absence of an organised system to literature searching, or prescribed guidance or standards, it was unclear how study authors and literature searchers had identified and selected the studies they drew upon in reviews, and cited to reach their conclusions (14). It was likely that the results of reviews of the literature were prone to bias, which could impact the reliability of findings and conclusions which could be drawn.

1.1.2 The Cochrane Collaboration

Archie Cochrane's call, in 1979, that there was no 'organised [and] critical summary…of all relevant controlled trials' (15-17) developed the idea for, and methodology of, a systematic review of studies.

The spirit of this idea was captured in a systematic review and meta-analysis by Crowley et al. which identified the effectiveness of prenatal corticosteroids in reducing early neonatal mortality (18). By combining the findings of 12 controlled trials, Crowley et al. were able to demonstrate the effectiveness of corticosteroids in reducing early neonatal mortality, for the first time, and in a way which not only combined the results from a number of similar studies but also reduced bias by making clear statements on the methods used to identify and combine these relevant studies (18, 19). The impact of this systematic review and meta-analysis led to a change in practice and the use of corticosteroids as an intervention (19), which has

remained in place for 24 years until the findings were challenged by a cluster randomised control trial by Althabe et al. in 2014.

Crowley et al's review was important not only due to its findings but also because of what it has come to represent. The UK Cochrane Centre was founded in 1992, and The Cochrane Collaboration was founded in 1993 as an organised group of like-minded researchers (20-22). Crowley et al's forest plot, a graphical display of the estimated results from studies identified and included in the review, was included in the design of the Cochrane Collaboration's logo by David Mostyn. The result of all of these initiatives was Cochrane Reviews, systematic reviews of intervention effectiveness with studies reporting randomised controlled trials as their unit of analysis (23-25).

1.1.3 Towards a systematic process of literature searching: a need for comprehensive literature searches

The stated aim of the methods developed for Cochrane systematic reviews was to minimise the introduction of bias in systematic reviews through a comprehensive search for studies (25, 26). The Cochrane Collaboration focused on systematic reviews of intervention effectiveness, most particularly, determining intervention effectiveness by the evaluation of studies reporting Randomised Controlled Trials (RCTs) as the unit of analysis (25).

As it relates to literature searching for studies reporting RCTs, a key bias to minimise was (and remains) publication bias (26). Publication bias is the recognition that studies with positive clinical or statistically significant findings are more likely to be published than studies reporting null or unfavourable results (27). To reliably determine effectiveness, it is necessary to minimise the effect of publication bias, by identifying both the published and the unpublished studies (28), unpublished study data and studies published in studies other than English (29, 30), in order to ensure that all relevant studies, showing statistically significant, or those with null or unfavourable results, are identified.

In the mid-late 1980s (and today), researchers were able to empirically demonstrate the impact of publication bias in systematic reviews of intervention effectiveness by purposefully excluding studies from statistical meta-analysis, thereby demonstrating the effect of 'missing' studies. These studies demonstrated that the estimate of

intervention effectiveness would be altered, if studies were excluded from meta-analysis (28, 31, 32).

This suggested a need for 'comprehensive' literature searches in which the aim was the complete identification of all published and unpublished studies reporting RCTs relevant to any intervention under review. To ensure the complete identification of studies reporting RCTs, it was necessary to define a process of literature searching which had comprehensive study identification as its aim, that could deliver literature searches which did not miss studies but also did not overwhelm researchers with an unmanageable number of studies to review, and that could be transparently reported so that the process could be validated.

1.1.4 The Cochrane Handbook

The Cochrane Handbook was first published in 1994, and it set out a methodological process of literature searching which Cochrane authors should follow, to complete Cochrane reviews and identify studies reporting RCTs (33). This process focused on the use of bibliographic databases to identify published studies reporting RCTs and it advocated the use of supplementary search methods, such as handsearching, to identity unpublished and unreported RCTs. The handbook also called for transparent reporting of the methods of literature searching, including the full reporting of the search terms used to literature search. The reporting of the methods used and the process followed meant that the process of study identification could be replicated and therefore validated.

The Cochrane Handbook was the first methodological handbook to report a systematic process of literature searching. The reported process reflected the work of The Cochrane Collaboration, which was starting to improve the process of literature searching for studies reporting RCTs, reconciling this with the practical realities of study identification in 1993/1994, coupled with advances in technology (34). This was not straight-forward. As an architect of the process, and author of the handbook, Carol Lefebvre noted at the time, 'databases are inadequate' (24). This statement on the position of bibliographic databases is key to understanding the evolution of literature searching in systematic reviews and the role of technology (8).

1.1.5 Database indexing, the MEDLINE re-tagging project, and handsearching

In 1994, as noted by Lefebvre (above), and colleagues, the position was stark (35). Dickersin et al. demonstrated that relying on indexing to identify RCTs would identify

only 51% of known clinical trials (36) and Adams et al. found that searches of MEDLINE alone would miss relevant studies for systematic reviews of mental health topics (35). There was no easy or rapid way to identify studies reporting RCTs because authors were not consistently reporting their studies using study design terminology (37, 38) and relevant studies were not indexed in leading bibliographic databases (10, 37). Both points prohibited the easy and comprehensive identification of studies (36, 39), a process which was and remains necessary to generate a reliable estimate of intervention effect, and to complete a systematic review of intervention effectiveness.

Lefebvre gained agreement from the National Library of Medicine (NLM) to retrospectively 're-tag,' and prospectively index, studies reporting RCTs from 1993/1994 onwards and all the way back to 1948 in the database PubMed(MEDLINE) (33, 36, 40, 41). The so-called 'MEDLINE re-tagging project' resulted in more than 125,000 additional MEDLINE records being identified and indexed through a comprehensive handsearching effort of journals indexed in MEDLINE (14, 36, 42).

Lefebvre also secured, from Elsevier (the publisher of the bibliographic database Embase), an agreement to introduce an indexing term for RCTs into Embase in 1993 (14, 24) to match the same term in MEDLINE, which had been introduced in 1991. The retrospective 'tagging' of study records for RCTs and Controlled Clinical Trials (CCTs) in Embase only began in 2016 and it is still on-going at this time (2018).

The process of handsearching, the creation of indexing terms, and the 're-tagging' I describe above dramatically increased the visibility of, and access to, studies reporting RCTs in bibliographic databases (34).

1.1.6   The creation of CENTRAL

In 1992, in tandem with the MEDLINE re-tagging project, and in response for calls from researchers (35), work began on a register of studies reporting trials (40). The Cochrane Central Register of Controlled Trials (known as CENTRAL) was conceived as a register of studies reporting RCTs and CCTs derived by searches of MEDLINE and Embase and incorporating studies identified by handsearching (37). CENTRAL was first published in The Cochrane Library in 1996 and it offered Cochrane authors and researchers efficient access to studies reporting RCTs, in one place and for the first time (43-46).

1.1.7 Moving forward: Search filters and the evaluation of 'effectiveness'

Due to the creation of study design indexing terms for RCTs and trials, better indexing in bibliographic databases, and the identification and 're-tagging' of 'untagged' studies reporting RCTs, the number of studies reporting RCTs available in MEDLINE rose by 1250% (20,000 in 1993 to 270,000 in October 2005) (42). Moreover, the CONSORT reporting guidelines were put forward and introduced in 1996 to aid study authors in reporting RCTs (47). CONSORT provides guidance on reporting RCTs, for example, to include study design terminology in the title of the study report. This aimed to increase the effectiveness and efficiency of identifying and indexing studies reporting RCTs (48, 49).

The combination of easier to identify study reports (through the use of CONSORT), and better indexing of studies reporting RCTs due to the work of Lefebvre and colleagues (41), led to the development of methodological literature search filters and an increase in the evaluation of literature searching practice in the context of systematic reviews. Methodological literature search filters are 'lists' of terms most likely to be reported in relevant studies (50). For instance, an RCT search filter focuses on the term 'random' to indicate the process of randomisation in a trial and, per CONSORT, the study author should report that an RCT has been conducted in the title of any study report (47). If both of these points are observed, as a minimum, then the study should 'mesh' with an RCT search filter, and the study in question should be identified (50).

Numerous search filters to identify studies reporting RCTs have been proposed (c.f. (42, 51-53)), with the Cochrane Highly Sensitive Search Strategy (known as the HSSS) being amongst the first and most prominent (25). The operating characteristics of these search filters have been evaluated and derived using the language and process of diagnostic test accuracy – so they are described in terms of their 'sensitivity' and 'specificity' (5). In this introductory chapter, the evaluations of study design literature search filters serve two purposes. First, they indicate the success of the MEDLINE re-tagging project, validating the substantial work done by the handsearchers on that project, since a large number of relevant studies went from not being indexed to being indexed and, secondly, it moved the conversation in literature searching on from 'how to identify RCTs' to 'how can this be done effectively and efficiently'(48), a conversation that was nested in an emerging number of evaluation studies published by information professionals and researchers (54).

The evaluation studies indicated above (c.f. (42, 51-53)), focused on the idea of comprehensive literature searching. Comprehensive literature searches, sometimes called exhaustive literature searching (48), were deemed to be an important and necessary component of a systematic review of intervention effectiveness and are required to produce a reliable estimate of intervention effect. The aim is broadly understood to mean 'find every possible study'. The evaluation studies set out in the paragraph above therefore produced summary estimates of search filter effectiveness based on how effective the search filter in question was at identifying all known studies, and it focused research attention on evaluating literature search practice.

By 2003, researchers had started to question the importance of comprehensive searches in the evaluation of intervention effectiveness (28). Researchers sought to examine if 'complicated' literature searches, which aimed for high sensitivity, were necessary given the advances made by The Cochrane Collaboration (48, 55). Royle and Waugh, for instance, found that a simplified search filter, using the publication type indexing term for RCTs, was effective at identifying studies reporting RCTs, concluding that the 'exhaustive' work of identifying RCTs had now been completed and the results were easily accessible in CENTRAL for intervention effectiveness systematic reviews (56).

1.1.8 2018 and Machine learning

The study identification debate has moved further on in recent years. Researchers are exploring and evaluating the technologies of machine learning and of text mining to identify studies (57-61). Machine learning/text mining occurs at the point of screening, but works on principles similar to study design literature search filters, exploiting the advancements set out above in study indexing and for studies where authors have adhered to CONSORT guidance. As Marshall et al. report, machine learning can effectively and efficiently identity studies which follow CONSORT reporting (62), and that therefore would be identified by study design literature search filters. Machine learning and text mining would again seem to validate the work of Cochrane and the MEDLINE re-tagging project, as I set out above. The difference between the work I set out above (and specifically literature search filters) is that the 'machine' learns as it works, responding to screening decisions made by the researcher and re-prioritising similar records (63).

1.1.9 2018+ and new challenges?

The work of Cochrane and the research community has, in many ways, since come full circle. The chronological advancements in handsearching, indexing, the development of Cochrane CENTRAL, study design literature search filters and advances in technology, have made identifying studies reporting RCTs almost a matter of routine (55). The process of literature searching which has evolved, and which is articulated in guidance documents such as the Cochrane Handbook, has become a by-word for systematic, comprehensive and transparent study identification.

Recent research would suggest that, whilst the advancements in systematic literature searching detailed above have been positive, they have been focused on studies which have a chance of being identified. Jefferson and Doshi have highlighted the differences between relying on published and reported studies, which can be identified and where researchers have focused their methodological effort on improving the effectiveness and efficiency of study identification, and clinical study reports (CSRs), which are not yet commonly or are freely available (64). CSRs are the detailed report of the methods and results of clinical trials which are usually not publicly available (64). Jefferson and Doshi have demonstrated the importance of this finding by comparing a Cochrane review of published studies to a systematic review on the same topic and interventions based-upon unpublished study data found in clinical studies reports (65, 66). They argue that systematic reviews of intervention effectiveness, particularly those focused on pharmaceutical interventions, should not now be based on published study data but instead the CSRs should form the primary unit of analysis (65, 66). This would suggest that there exists another sub-stratum of relevant data for intervention effectiveness reviews: there is data which is relevant but which is almost impossible to identify (67, 68)). A survey of 160 systematic reviewers by Hodkinson et al. found support for the use of CSRs in intervention effectiveness reviews and called for methodological guidance on how to include and use data from CSRs (69). This would suggest that the process of literature searching discussed above might require another modification in 2018, in order to account for this kind of data.

The work presented above briefly summarises how the research community evolved from an arguably ad-hoc and often unreported process of identifying studies to a systematic, comprehensive, transparent and replicable process of literature

searching. This process has evolved hand-in-hand with the research needs of systematic reviews.

I did not set out to focus this summary on literature searching for reviews of intervention effectiveness. What we can see above though, is that the methodological interest existed, and the resources were available, for investment in developing the literature searching for this style of systematic review.

1.1.10 Beyond reviews of effectiveness and reviews which look beyond RCTs

There is an increasing interest in different types of evidence to address different types of research questions (70-74). The investment and research that I briefly summarise above has been focused on reviews of intervention effectiveness, in particular the identification of studies reporting RCTs.

Other types of systematic review, by method or topic, have not benefitted from the resources and energies of a figure-head such as The Cochrane Collaboration. Advances in literature searching have been slow and uneven without this investment, even though the underlying technology and search methods are the same and have evolved in the same way. This means that literature searching for systematic reviews beyond the evaluation of intervention effectiveness, and beyond the search for studies reporting RCTs, is a more complex task (74).

Taking qualitative evidence synthesis as an example, there have been calls for a controlled indexing term for studies reporting qualitative methods or data (75) but a controlled indexing term for qualitative research was not introduced into MEDLINE until 2003: twelve years later than the indexing term for RCTs. There has been no programme to retrospectively 're-tag' study records for qualitative studies as there has been for studies reporting RCTs and trials. It is not clear how successful such a programme of work would be, since identifying qualitative research is acknowledged as challenging (76, 77). Journal word limits often prevent the full reporting of qualitative research, which only adds to the difficulty of identifying relevant qualitative studies and data, and authors commonly neglect to identify study design or methods of data collection or, more specifically, identify their studies as 'qualitative' research when reporting their work in either title or abstract (6, 78, 79).

This may explain why there have been no correspondingly successful advances in the development of literature search filters for studies reporting qualitative research, despite study authors calling for, and testing, such ideas (75, 77, 80-82). When

compared to similar filters for intervention effectiveness, the operating characteristics do not compare in terms of effectiveness (c.f. (82-84)). Effectiveness, in this case, meaning the successful identification of relevant and known qualitative studies. Without reliable indexing and a programme of handsearching to identity qualitative studies presently overlooked by indexers, combinations of literature search methods are needed to identify qualitative studies (77, 81, 85) and databases are less effective than when used to identity studies reporting RCTs (86, 87).

The problems briefly summarised above for qualitative evidence synthesis are similar in other types of review or topic areas. Briefly, in public health reviews and reviews of social care topics, the indexing of studies is poor and it is acknowledged that researchers must search more broadly for published and unpublished studies and study data (75, 76). It is not uncommon to search numerous databases when compared to only CENTRAL, MEDLINE and Embase for reviews of intervention effectiveness in public health topics and the number of studies identified for screening is often considerably higher when compared to reviews of intervention effectiveness which focus on studies reporting RCTs (50, 77-82).

In addition, in reviews of diagnostics and prognostics, the correct indexing of studies is poor, and study authors use a broad and often inconsistent language to report diagnostic test evaluation, so the literature search filters that exist are poor, and their use is not advised (83, 84). This means that there is no corresponding option for the effective identification of relevant studies when compared to studies reporting RCTs (84-87). Again, without any similar investment in retrospectively indexing studies, searching for studies reporting diagnostic or prognostic accuracy remains challenging.

For each of the examples set out above, there exists a developing body of evaluation studies and research undertaken by information professionals and researchers. These studies examine the methods (79, 88-96) or the tools (97-109) of literature searching and they make individual contributions to the evidence-base of literature searching (110). There has, however, been no distinct attempt to collate the evidence, or to organise a process of literature searching in a way similar to the systematic reviews of intervention effectiveness (explored above), or to reflect the differing needs of these different types of systematic review.

The responses from the research community have taken a similar tack to the solutions identified and implemented in intervention effectiveness literature

searching. This has focused on study design literature search filters but they have been unsuccessful largely (if not wholly) due to an absence of funding and lack of support. It should be remembered that literature search filters to identify studies reporting RCTs are made easier because of the MEDLINE re-tagging project, which retrospectively identified and indexed relevant studies, and the prospective introduction of relevant indexing terms into leading bibliographic databases, which aims to ensure relevant studies were suitably indexed prospectively.

It is not clear if similar investment in the architecture underpinning literature searching for other types of systematic review would improve the process of study identification and literature searching. Where researchers have attempted to replicate the techniques used to evolve literature searching in intervention effectiveness reviews (c.f. (42, 80, 83, 84, 111-114)), there has been no attempt to consider the actual underlying process of literature searching; namely, (i) to consider whether changing the process of literature searching to reflect different research questions and a need for different types of study or data, and (ii) to re-evaluate whether it is necessary to undertake comprehensive literature searching at all, might prove to be another way to resolve the issues that I set out above. I suggest that, in systematic reviews that look beyond the evaluation of intervention effectiveness, the underlying process of literature searching should change.

1.2 Summary

The 1990s defined systematic reviews and systematic literature searching. The work, resources and energy of The Cochrane Collaboration, and researchers engaged with intervention effectiveness systematic reviews, led to the evolution of methods and creation of an infrastructure for studies reporting randomised controlled trials, which is unrivalled in other topics or methodological areas of research.

Advancements in technology were recognised and harnessed early. The inclusion of study design indexing terms for studies reporting RCTs in leading bibliographic databases, and the MEDLINE re-tagging project to correctly index RCTs, combined with the development of bibliographic databases and the development of literature search filters, changed the methods and expectations of literature searching beyond recognition.

The development of methods (i.e. search filters) and process has been uneven and less successful for other types of systematic reviews. Without a similar investment of

time and resources, the process of literature searching remains a challenging undertaking.

1.3   Chapter findings

The crux, as it relates to this thesis, is this: the success of systematic reviews of intervention effectiveness, and the approach to literature searching specifically, has established the methods, the process and rules of literature searching by which all other types of systematic review are measured and compared.

As I set out in this chapter, the advances in literature searching for other types of systematic review, by method or topic, has been uneven. The lack of indexing terms, the absence of programmes of work to retrospectively identify and index studies, inconsistent reporting, and the lack of success in developing study design literature search filters, all mean that literature searching in other types of systematic review, has not kept pace with systematic reviews of intervention effectiveness.

I suggest that the underlying process of literature searching is worthy of examination and this is the focus of this thesis.

Having set out the history and context of literature searching in systematic reviews, I will now set out the aims of this thesis and the objectives that I will address.

2   Aims and objectives of this thesis

In the preceding chapter, I have set out the history and position of literature searching in systematic reviews. The focus was unintentionally but determinedly on systematic reviews of intervention effectiveness, in particular those whose primary unit of analysis is studies reporting randomised controlled trials (RCTs). As I have set out in Chapter 1, this is where the greatest investment in resources, effort and research has been concentrated to date.

In other types of reviews, methodologically or by topic area, there is no corresponding equity in investment. Instead, and in the absence of any similar investment of resources and time, the same process to literature searching has been adopted wholesale from systematic reviews of intervention effectiveness without any question as to if this process is suitable. I seek to question in this thesis if this 'adoption' of methods and of process is suitable.

2.1   The aims of this thesis

The aims of this thesis are to:

1) examine approaches to systematic literature searching for systematic reviews; and 2) propose and test a method of systematic literature searching for reviews that do not focus on the effectiveness of clinical interventions.

2.2   The objectives

I will explore these aims through the 9 objectives that I set out below in Table 2 below.

Table 2 Objectives, research questions, methodology and location

Section	Research question(s)	Objective(s)	Methodology	Where in the thesis
Section 1	Can a shared model of the literature searching process be detected across guidance documents and, if so, how is this process reported and supported?	1) To conduct a critical review of influential systematic review guidance documents to assess whether they share an understanding of the approach to searching for systemic reviews.	Critical literature review.	Chapter 3
	If a shared model can be identified, is it suitable to meet the needs of researchers undertaking systematic reviews?	2) To use the findings of this review to articulate a model of current approach(es) to searching in systematic reviews.		
		3) To conduct a structured critical review of the empirical studies that have investigated each stage of this model in order to understand if there is support for the proposed model and whether it is suitable to meet the needs of researchers.		
Section 1	What is the origin of thinking behind tailored literature searching?	4) To discuss the original thinking behind the need for tailored literature searching.	Literature review	Chapter 4
	What is tailored literature searching?	5) To theorise how tailored literature searching could work conceptually.		
	How does tailored literature searching work in theory?			
Section 1	How do empirical studies of supplementary search techniques compare to the recommendations in the review methodology handbooks?	6) To review the review methodology handbooks and empirical studies on supplementary search methods, in order to determine the advantages, disadvantages and resource requirements of supplementary search methods.	Literature review. Publication one.	Chapter 5
Section 1	What metrics or methods are used to calculate literature search effectiveness? Which metrics or methods are used formatively or summatively?	7) To conduct a systematic review to identify all metrics or methods currently used by researchers to demonstrate literature search effectiveness. 8) To use the findings from this systematic review to identify which methods are formative and summative	Systematic review. Publication two.	Chapter 6

	How is effectiveness defined by study authors?	and how study authors define what effectiveness means in literature searching.		
Section 2	How does the tailored literature search method compare to the conventional approach in practice?	9) to report a case study in which I compare a tailored approach to literature searching to a conventional approach and to examine the findings of this case study in context of the research question of this thesis.	Case study. Publication three and four.	Chapter 7 and Chapter 8

I will revisit this table (Table 2) in Section Three (Chapter 9). In Chapter 9, I will repeat the table reported above (Table 2) but I will include the findings as they relate to the objectives.

I will then set out my contributions to knowledge (Chapter 9.7), conclusions (Chapter 10), and areas for further research (Chapter 11).

In the next Chapter, Chapter 3, I aim to critically review guidance documents and studies to determine if there is a shared process to literature searching and if this is suitable to meet the needs of researchers undertaking systematic reviews.

3 The conventional approach

3.1 Introducing the conventional approach

It was my hypothesis, at the start of this thesis, that there was one common process to literature searching in systematic reviews which was recommended in guidance documents, and that it was followed and used by researchers, irrespective of the type of systematic review being undertaken. If this hypothesis were true, it would raise the question, is it suitable to use the same process of literature searching to identify the range of studies and study designs needed to inform the development of a number of methodologically divergent systematic reviews?

3.2 How this chapter addresses the research aims of this thesis

In this chapter, I aim to test my hypothesis that guidance documents broadly report the same process of literature searching for different types of systematic review. Identifying, establishing and examining this process would make a unique contribution to knowledge because this idea has not been explicitly identified or examined before.

By reviewing the guidance documents, to identify recommend best practice, and supporting studies, to identify evaluations of current practice, I will be able to establish what constitutes recommend best practice in literature searching for systematic reviews. This will form an important contribution to my thesis and broader research aims, since it will establish the benchmark against which I will develop and test my idea for an alternative approach to literature searching in systematic reviews. Namely, the tailored approach to literature searching.

It is important to note a limitation here. This review will only determine how guidance has been adopted into practice and how this guidance is supported by empirical studies. The conventional approach itself been tested or evaluated.

3.3 Chapter objectives and research questions

The objectives of this chapter are:

- To conduct a critical review of influential systematic review guidance documents to assess whether they share an understanding of the approach to searching for systemic reviews.

- To use the findings of this review to articulate a model of current approach(es) to literature searching in systematic reviews.

- To conduct a structured critical review of the empirical studies that have investigated each stage of this model in order to understand if there is support for the proposed model and whether it is suitable to meet the needs of researchers.

To address these objectives, I identified two research questions:

1. Can a shared model of the literature searching process be detected across guidance documents and, if so, how is this process reported and supported?

2. If a shared model can be identified, is it suitable to meet the needs of researchers undertaking systematic reviews?

3.4 Defining the process to literature searching in systematic reviews: a critical literature review of guidance and supporting studies

3.4.1 Background

Systematic literature searching is recognised as a critical component of the systematic review process. It involves a systematic search for studies and aims for a transparent report of study identification, leaving review stakeholders clear about what was done to identify studies, and how the findings of the review are situated in the relevant evidence.

Information professionals and review teams appear to work from a shared and tacit model of the literature search process. How this tacit model has developed and evolved is unclear, and it has not been explicitly examined before. This is in contrast to the information science literature, which has developed information processing models as an explicit basis for dialogue and empirical testing. Without an explicit model, research in the process of systematic literature searching will remain immature and potentially uneven, and the development of shared information models will be assumed but never articulated.

One way of developing such a conceptual model is by formally examining the implicit "programme theory" as embodied in key methodological texts. The research questions of this literature review are therefore to determine:

1. if a shared model of the literature searching process in systematic reviews can be detected across guidance documents and, if so, how this process is reported and supported; and

2. if a shared model can be identified, is it suitable to meet the needs of researchers undertaking systematic reviews?

3.4.2 Methods

Two types of evidence are needed to address the first research question (see 3.4.1): guidance and studies. Below, I set out the methods I used to identify both types of evidence.

I also set out the methods I used to identify models of information retrieval and behaviour. These models will help me to determine if any pre-existing models of information retrieval exist that meet, or could meet, the needs of systematic reviewers. The studies reporting these models will also help inform and develop the broader critique of suitability (research question 2 of this chapter (see 3.4.1)) and the development of the tailored approach in the next chapter.

Identifying guidance

Key texts (henceforth referred to as "guidance") were identified based upon their accessibility to, and prominence within, United Kingdom systematic reviewing practice. The United Kingdom occupies a prominent position in the science of health information retrieval, as quantified by such objective measures as the authorship of papers, the number of Cochrane groups based in the UK, membership and leadership of groups such as the Cochrane Information Retrieval Methods Group, the HTA-I Information Specialists' Group and historic association with such centres as the UK Cochrane Centre, the NHS Centre for Reviews and Dissemination, the Centre for Evidence Based Medicine and the National Institute for Health and Care Excellence (NICE). Coupled with the linguistic dominance of English within medical and health science and the science of systematic reviews more generally, this offers a justification for a purposive sample that favours UK, European and Australian texts.

Nine guidance documents were identified. These documents provide guidance for different types of systematic reviews, namely: reviews of interventions, reviews of health technologies, reviews of qualitative research studies, reviews of social science topics, and reviews to inform guidance (see Table 3).

Whilst these guidance documents occasionally offer additional guidance on other types of systematic reviews, the core and stated aims of these documents as they relate to literature searching have been the focus of this review. Table 3 sets out: the guidance document, the version audited, their core stated focus, and a bibliographical pointer to the main guidance relating to literature searching.

Once a list of guidance documents was determined, it was informally checked by six senior information professionals based in the UK for relevance to current literature searching in systematic reviews.

Table 3 Guidance documents audited for this literature review

Guidance documents	Version: Year	Core focus	Location of guidance in the guidance document
Systematic Reviews: CRD's guidance for undertaking reviews in health care (3).	2009	systematic reviews of health care interventions	1.3 Pages 16-22
The Cochrane Handbook (25).	Version 5.1: June 2017	systematic reviews of interventions	Chapter 6: Searching for studies
Collaboration for environmental evidence: Guidelines for systematic reviews in environmental management (115).	Version 4.2 March 2013	systematic reviews of environmental evidence	Section 4.1 (pages 36-41)
Joanna Briggs Institute Reviewers' Manual (116).	2014 edition	systematic reviews of qualitative studies	Chapter 7 Information Retrieval (pages 28-31)
Institute for Quality and Efficiency in Health Care (IQWiG): IQWiG (117).	2014	systematic reviews of health care interventions	Chapter 7: Information retrieval
Systematic Reviews in the Social Sciences: A Practical Guide (81).	2006	systematic reviews of social science topics	Chapter 4. How to find the studies: the literature search (pages 81- 124)
Process of information retrieval for systematic reviews and health technology assessments on clinical effectiveness. Eunethta (118).	Version 1.1 December 2016.	systematic reviews of health care interventions	Standalone guideline on literature searching
The Campbell Handbook: Searching for studies: a guide to information retrieval for Campbell systematic reviews (119).	Version 1.1. February 2017.	systematic reviews of interventions in social science topics	Standalone guideline on literature searching
Developing NICE guidelines: the manual (120).	2014	systematic reviews to inform health care guidelines	Chapter 5. Identifying the evidence: literature searching and evidence submission.

Identifying supporting studies

In addition to identifying guidance, I sought to populate an evidence base of supporting studies (henceforth referred to as "studies") that contribute to existing

search practice. Studies were first identified from my knowledge and that of my supervisors on this topic area and, subsequently, through systematic citation chasing key studies ('pearls' (101)) located within each key stage of the search process. These studies are identified in the appendix (volume 2) of this thesis. Citation chasing was conducted by analysing the bibliography of references for each study (backwards citation chasing) and through Google Scholar (forward citation chasing).

A search of PubMed using the systematic review methods filter was undertaken in August 2017. The search terms used were: (literature search*[Title/Abstract]) AND sysrev_methods[sb] and 586 results were returned. These results were sifted for relevance to the key stages in Figure 1 by me.

Identifying models of information retrieval and behaviour

I also aimed to determine if there were any existing information retrieval models which could be used to explain the process of literature searching in a systematic review or help develop an alternative process.

Models were initially identified by pearl growing (101) from leading authors in this field (e.g. Bates' 'Berry picking model' (12)) and additionally through supervisory discussion. I supplemented this, in 2018, with a scoping search to broaden the search for potentially relevant models. This focus on models of information retrieval theory was partly intended to inform my understanding of the information retrieval theory literature for the purposes of critique and partly to explore other models of information retrieval, as I sought to develop my own in the form of the tailored approach (see Chapter 4).

Once a publication reporting a potentially relevant model was identified, two sets of citation chasing were undertaken: (i) backwards and (ii) forwards. Publications found to report models were read and were included if they validated, developed or helped me explore either key stages of the conventional approach or the conventional approach as a process. The method used for this section (3.7.1) was as for a scoping review (121).

Extracting and analysing the data

Guidance and studies

To reveal the implicit process of literature searching within each guidance document, the relevant sections (chapters) on literature searching were read and re-read, with

the aim of determining key methodological stages. A key methodological stage was defined as a distinct step in the overall process for which specific guidance is reported, and action is taken, that collectively would result in a completed literature search.

The chapter or section sub-heading for each methodological stage was extracted into a table using the exact language as reported in each guidance document. I then read and re-read these data, and the paragraphs of the document to which the headings referred, summarising section details. This table was then reviewed, using comparison and contrast to identify agreements and areas of unique guidance. Consensus across multiple texts was used to inform selection of 'key stages' in the process of literature searching.

Having determined the key stages to literature searching, the sections relating to literature searching were read and re-read again, extracting specific detail relating to the methodological process of literature searching within each key stage. Again, the guidance was then read and re-read, first on a document-by-document-basis and, secondly, across all the documents above, to identify both commonalities and areas of unique guidance.

Models of information retrieval and behaviour

For the models identified in scoping searches, I tabulated the model name, author, purpose of the model and I have illustrated the models in the appendix (volume 2) of this thesis.

3.4.3 Findings

Consensus was identified across the guidance on literature searching for systematic reviews suggesting a shared implicit model within the information retrieval and research community. Whilst the structure of the guidance varies between documents, the same key stages are reported, even where the core focus of each document is different (see Table 3).

Specific areas of unique guidance were identified, where a document reported guidance not summarised in other documents, together with areas of consensus across guidance. I set these out below.

Unique guidance

Only '*Systematic Reviews in the Social Sciences: A Practical Guide*' provided guidance on the topic of when to stop searching (81). This guidance from 2006 anticipates a topic of increasing importance with the current interest in time-limited (i.e. "rapid") reviews. Quality assurance (or peer review) of literature searches was only covered in two guidance documents (117, 120). This topic has emerged as increasingly important as indicated by the development of the PRESS instrument (122). Text mining was discussed in four guidance documents (3, 118-120) where the automation of some manual review work may offer efficiencies in literature searching (119).

Areas of consensus across guidance

Where there was agreement on the process, it was determined that this constituted a key stage in the process of literature searching to inform systematic reviews.

From the guidance, eight key stages were determined that relate specifically to literature searching in systematic reviews. These key stages are summarised and illustrated at Figure 1. The data extraction table to inform Figure 1 is reported in the appendix of this thesis (volume 2) and it reports the areas of common agreement and it demonstrates that the language used to describe key stages and processes varies significantly between guidance documents, although they are functionally/ conceptually the same search methods.

The areas of consensus across guidance documents inform development of what I will call 'the conventional approach'. I chose this phrase as common convention across guidance documents appears to support a shared implicit model of information retrieval in systematic reviews. This finding addresses research question one of this chapter (see 3.4.1).

I shall now set out the overall process model and then the individual key stages that make up the process model. Within this, and to address research question two of this chapter (see 3.4.1), I will critique each item to determine the suitability of the overall process.

Figure 1 The conventional approach

| Key stage one: *who should literature search* | Key stage two: *aims and purpose of literature searching* | Key stage three: *preparation* | Key stage four: *search strategy* | Key stage five: *bibliographic database searching* | Key stage six: *supplementary searching* | Key stage seven: *managing references* | Key stage eight: *reporting the search process* |
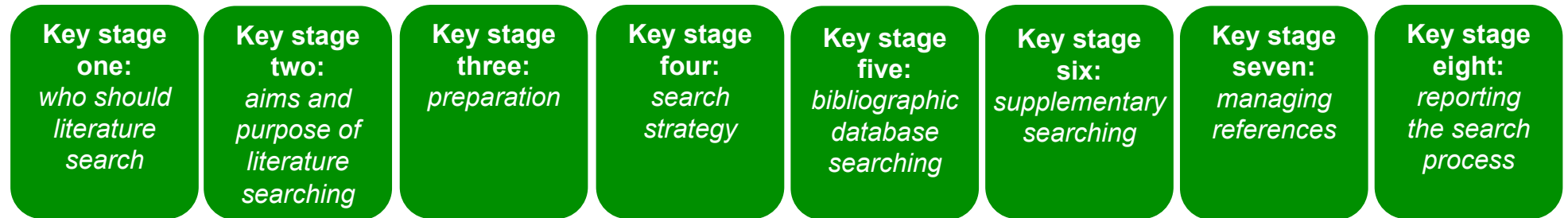
Presenting the findings: introducing the conventional approach

I will present the findings below.

First, I will focus on the top-level process that I have identified. I illustrate the top-level findings as a conceptual process model (see Figure 1) The illustration of the conceptual process model allows me to illustrate how the process of literature searching in systematic reviews is thought to produce its results (123). I will critique the design of the model, and briefly situate it in the context of other process models which seek to explore information retrieval practice or behaviour. This will allow me to explore the suitability of the conventional approach to meet the needs of systematic reviewers (research question 2: see Section 3.4.1).

I will then turn my attention to the eight individual key stages set out in Figure 1. I will set out my findings for each individual key stage to address research question one (see Section 3.4.1) and I will critique my findings to address research question two (see Section 3.4.1).

The conventional approach: the conceptual model

I set out the eight key stages illustrated as a conceptual model in Figure 1.

The conventional approach is manifestly a 'problem-orientated' conceptual model (124, 125). The problem explored by the conventional approach is the need for a process of literature searching that is comprehensive, transparent and replicable, since these factors have been linked to minimising the introduction of bias in systematic reviews (3, 115, 119). I summarise the need for, and understanding of, comprehensive, transparent and replicable literature searching, and I explore these points individually, in Key Stage Two.

The conventional approach, as set out in Figure 1, aims to illustrate the relevant stages and identify the process of literature searching in a systematic review. I have broken the process into key stages, which help demonstrate the inter-related complexity of the process, whilst situating how the 'real-life' need for a process of literature searching has represented itself into this pragmatic model (124). The resulting process model sets out the conceptual response to the problem which the conventional approach seeks to explore (126-128).

Problem-orientated models are primarily dependent on expert input rather than the availability of empirical data (124). Nine leading guidance documents (see Table 3)

serve as a proxy for expert input, although I have also 'tested' the existence of, and need for, the key stages by using empirically derived case-studies (129). I also draw from 10 years' experience as an information professional and systematic reviewer.

Is this suitable? A critique of the conventional approach as a process model

Whilst it is clear that a shared conceptual process can be determined across guidance documents, it is not clear how the conventional approach was derived, or if the chronological process I illustrate in Figure 1, was purposefully planned, or if it emerged organically, to meet the evolving needs of systematic reviewers.

I identified nine potentially relevant models of information retrieval and information behaviour. Table 4 sets out the models identified and the date of the publication. I have also illustrated each of these models, setting them alongside the conventional approach to explore any potential overlap. This work is presented in the appendix of this thesis (see volume 2).

Table 4 Nine conceptual models used to explore the conventional and tailored approach to literature searching

Model name	Author (date)	See figure
The process of asking questions & question negotiation (130).	Taylor (1962)	See appendix: volume 2 of thesis
Anomalous state of knowledge (ASK) (131, 132).	Belkin (1980 + 1982)	See appendix: volume 2 of thesis
Model of information behaviour (133).	Wilson (1981)	See appendix: volume 2 of thesis
Sense-Making triangle (134).	Dervin (1983)	See appendix: volume 2 of thesis
Ellis' model (135, 136).	Ellis (1989) and Ellis, Cox & Hall (1993)	See appendix: volume 2 of thesis
Berry Picking (12).	Bates (1989)	See appendix: volume 2 of thesis
The Information Search Process (ISP) (4).	Kuhlthau (1991)	See appendix: volume 2 of thesis
Information needs, seeking and use (INSU) (137).	Byström (1999)	See appendix: volume 2 of thesis
The information journey model (138).	Du (2014)	See appendix: volume 2 of thesis

The models identified in Table 4 aim to provide a theoretical framework within which information behaviour and user interaction with information retrieval systems can be understood. No other models were identified which would replace or could challenge the conventional approach since no other models share the similar aims and purpose of the conventional approach (see Key Stage 2). Namely: the need for a process of literature searching which is comprehensive, transparent and replicable (3, 115, 119).

Two models promote similar search methods to those indicated for use in the conventional approach. Ellis (135), Ellis, Cox and Hall (136) (two papers presenting one model) and Bates' 'Berry picking' model (12) bore close similarities yet with a different overall focus to the conventional approach. Similar search methods were proposed to identify studies in these models, namely: 'chaining' (aka citation chasing) and browsing (136) and 'footnote/citation chasing', 'journal run' (aka handsearching) and 'subject searching' (aka bibliographic database searching)). Whilst the literature searching methods used to identify studies was similar, the process – or processes within models – were not sufficiently similar, and the aims and purposes were not similar either.

In the other seven models, search methods were not reported (130, 132-134) or searching was represented as a 'black box' function, where searching happened but it was not explained how (4, 137, 138).

I determine, therefore, that no alternative or suitable models existed and that the other models that I did identify (reported in Table 4) did not help articulate or explain the conventional approach since their aims and purpose were different to that of a systematic review.

This is not a surprising finding. With the exception of Byström (137) and Du (138), all of the models identified predate what I determine in Chapter 1 was the start of a process of systematic literature searching (1992). None of the handbooks I reviewed cited, or made reference to, any theoretical models of information retrieval. Even whilst based on a scoping search of information retrieval models, this would appear to suggest that the conventional approach emerged, or it was designed specifically, to meet the needs of systematic reviewers and for use in systematic reviews.

Moreover, the models identified in Table 4 are theoretical models which were designed to illustrate theoretical and library-based information retrieval problems. Specifically, what might be called 'single-use' engagement, in the sense that the models seek to explain how a user defines, identifies and completes an information

retrieval problem, often in the context of finding a single item in a library. It is a difference of a user seeking one item, such as a book, compared to a systematic reviewer or researcher who requires all of the relevant information on a particular topic, and is required to pre-define this request, and transparently report the method and process of retrieval. None of the nine models I identified suited this need. This would appear to suggest that, what I set out as the conventional approach, arose organically and as a process to meet the needs of researchers undertaking systematic reviews. The question now remains, is this process suitable? This is research question 2 (see 3.4.1)

The conventional approach as a process model: summary of findings

It would appear that the conventional approach emerged as a process to meet the needs of researchers undertaking systematic reviews. Further work, in the form of a systematic mapping or review of models, and primary qualitative work, would need to be undertaken to confirm this with greater certainty (see 11.1).

I have not identified any model which could replace the conventional approach, or which may have informed the development of the conventional approach. The requirement for a process of literature searching which is comprehensive, transparent and reproduceable was not an explicit requirement of the models identified in Table 4.

This would appear to suggest that, as least at the stage of critiquing the top-level suitability of the conventional approach, the needs of systematic reviewers and researchers are met by the conventional approach (research question two: see Section 3.4.1). The conclusion could be made firmer by undertaking a full systematic mapping of models which seek to explore information retrieval behaviour. I acknowledge this limitation and I identify it as an area for future research (see 11.1).

Findings: the individual key stages of the conventional approach

I will now turn my attention to each of the eight individual key stages set out in Figure 1. I will set out my findings for each individual key stage to address research question one (defined in Section 3.4.1) and I will critique my findings to address research question two, the suitability of this key stage (defined in Section 3.4.1).

I aim to determine which key stages are suitable to meet the needs of researchers undertaking systematic reviews and which are not. I will summarise agreement and disagreement in my chapter findings (and in Table 7).

Key Stage One: Deciding who should undertake the literature search

The guidance for Key Stage One

Eight guidance documents provided guidance on who should undertake literature searching in systematic reviews (3, 25, 81, 115, 116, 118-120). The guidance states that people with relevant expertise of literature searching should 'ideally' be included within the review team (3). Information professionals (or information scientists), librarians or trial search co-ordinators (TSCs) are indicated as appropriate researchers in six guidance documents (25, 81, 115, 116, 118, 119).

How the published studies correspond to the guidance for Key Stage One

Published studies appear to be consistent with the guidance documents that call for the involvement of information professionals and librarians in systematic reviews (126, 139-151). The studies seek to explore the role of information professionals in systematic reviews; how their training as 'expert searchers' and 'analysers and organisers of data' can be put to good use (143) in a variety of roles (139-142, 144, 146, 152). These arguments make sense in the context of the aims and purposes of literature searching in systematic reviews. The need for 'thorough' and 'replicable' literature searches is fundamental to the guidance and recurs in Key Stage Two.

Studies have found poor reporting, and a lack of replicable literature searches, to be a weakness in systematic reviews (147, 151, 153, 154) and they argue that involvement of information professionals/librarians would be associated with better reporting and better quality literature searching. Meert et al. demonstrated this, finding that involving a librarian as a co-author to a systematic review correlated with a higher score in the literature searching component of a systematic review (155). As 'new styles' of rapid and scoping reviews emerge, where decisions on how to search are more iterative and creative, a clear role for information professionals is evident here too (156).

Knowing where to search for studies was noted as important in the guidance, with no agreement as to the appropriate number of databases to be searched (3, 81).

Database (and resource selection more broadly) is acknowledged as a relevant key skill of information professionals and librarians (139, 145, 146, 157).

Whilst arguments for including information professionals and librarians in the process of systematic review might be considered self-evident, Koffel and Rethlefsen have questioned if the necessary involvement is actually happening (157).

Determining the suitability: A critique of Key Stage One

Guidance documents indicate that information professionals are appropriate researchers to undertake literature searching in systematic reviews (3, 25, 81, 115, 116, 118-120). This is helpful since it indicates clearly that researchers undertaking a systematic review should contact and work with an information professional. The Cochrane Collaboration, and their promotion and support of Trial Search Co-ordinators (information professionals tasked with helping Cochrane authors undertake literature searching for systematic reviews), exemplify this model of working. On starting a Cochrane review, Cochrane authors are referred to a named and expert information professional who is qualified and paid to undertake their literature searching.

Who should literature search?

The guidance does not explain why information professionals are chosen to literature search, nor does it articulate why an information professional is any better placed to literature search than any other type of researcher. I cited 15 studies which appeared to link guidance (information professionals should undertake literature searching) to the studies (and here is why) and I presented this as if it were causal. I have been challenged on the evidence for this link. This is a valid criticism which I shall explore below.

In Table 5, I set out the 15 studies I cite to link the guidance to the published studies, so to address the broader question of this key stage, who should literature search in systematic reviews.

Table 5 Studies which appear to support the need for information professionals in systematic reviews

Citation	RQ (if stated) or aim/objectives of paper	Study Design
Beverley CA, Booth A, Bath PA. The role of the information specialist in the systematic review process: a health information case study. *Health Info Libr J*. 2003;20(2):65-74.	Aim: The aim of this paper is to discuss the role of the information professional in the systematic review process	Case study (n/c)
Harris MR. The librarian's roles in the systematic review process: a case study. *J Med Libr Assoc*. 2005;93(1):81-7.	To chronicle a librarian's required involvement, skills, and responsibilities in each stage of a real-life systematic review.	Case study (n/c)
Koffel JB. Use of Recommended Search Strategies in Systematic Reviews and the Impact of Librarian Involvement: A Cross-Sectional Survey of Recent Authors. *PLoS One*. 2015;10(5):e0125931.	Objective: To more accurately measure the use of recommended search methods in systematic re- views, the levels of librarian involvement, and whether librarian involvement predicts the use of recommended methods.	Cross-sectional survey
Li L, Tian J, Tian H, Moher D, Liang F, Jiang T, et al. Network meta-analyses could be improved by searching more sources and by involving a librarian. *J Clin Epidemiol*. 2014;67(9):1001-7.	Objective: to assess the conduct of literature searches for NMAs.	Case study (n/c)
McGowan J, Sampson M. Systematic reviews need systematic searchers. *J Med Libr Assoc*. 2005;93(1):74-80.	Objective: This paper will provide a description of the methods, skills, and knowledge of expert searchers working on systematic review teams	Literature review
Rethlefsen ML, Farrell AM, Osterhaus Trzasko LC, Brigham TJ. Librarian co-authors correlated with higher quality reported search strategies in general internal medicine systematic reviews. *J Clin Epidemiol*. 2015;68(6):617-26.	To determine whether librarian and information specialist authorship was associated with better reported systematic review (SR) search quality	Case study (n/c)
Weller AC. Mounting evidence that librarians are essential for comprehensive literature searches for meta-analyses and Cochrane reports. *J Med Libr Assoc*. 2004;92(2):163-4.	N/A	Editorial
Swinkels A, Briddon J, Hall J. Two physiotherapists, one librarian and a systematic literature review: collaboration in action. *Health Info Libr J*. 2006;23(4):248-56.	Aim: The aim of this paper is to describe and evaluate some of the processes of collaboration between the three authors (one librarian and two academic physiotherapists) during work on a funded systematic literature review on the topic of hydrotherapy and pain.	Case study (n/c)
Foster M. An overview of the role of librarians in systematic reviews: from expert search to project manager. *J Eur Assoc Health Inf Libr*. 2015;11(3):3-7.	Not stated to detail the content provided in the workshop – introduction to SRs and the role of the librarian and project and data management in SRs	Literature review
Townsend WA, Anderson PF, Ginier EC, MacEachern MP, Saylor KM, Shipman BL, et al. A competency framework for librarians involved in systematic reviews. *J Med Libr Assoc*. 2017;105(3):268-75.	Objective: The project identified a set of core competencies for librarians who are involved in systematic reviews.	Literature review

Cooper ID, Crum JA. New activities and changing roles of health sciences librarians: a systematic review, 1990-2012. *J Med Libr Assoc*. 2013;101(4):268-77.	Objective: The paper identifies and documents new health sciences librarian activities and roles during the period from 1990–2012.	Literature review
Crum JA, Cooper ID. Emerging roles for biomedical librarians: a survey of current practice, challenges, and changes. *J Med Libr Assoc*. 2013;101(4):278-86.	Objective: This study is intended to (1) identify emerging roles for biomedical librarians and determine how common these roles are in a variety of library settings, (2) identify barriers to taking on new roles, and (3) determine how librarians are developing the capacity to take on new roles.	Literature review and Survey of librarians in biomedical setting.
Dudden RF, Protzko SL. The systematic review team: contributions of the health sciences librarian. *Med Ref Serv Q*. 2011;30(3):301-15.v	Objective: to review the contribution of librarians to a systematic review.	Literature review
Meert D, Torabi N, Costella J. Impact of librarians on reporting of the literature searching component of pediatric systematic reviews. *J Med Libr Assoc*. 2016;104(4):267-77.	Objective: The goal of this study was to compare the reporting rigor of the literature searching component of systematic reviews with and without the help of a librarian	Case study (comparative)
Morris M, Boruff JT, Gore GC. Scoping reviews: establishing the role of the librarian. *J Med Libr Assoc*. 2016;104(4):346-54.	To provide an overview of the scoping review and a summary of the current status of its definition and methodology. To suggest how librarians can and should involve themselves in this emerging area of knowledge synthesis landscape.	Literature review

As Table 5 demonstrates, the study designs used in the studies, do not actually help empirically address the question posed in this key stage: who should undertake the literature search? With the exception of the study by Meert et. al., the studies I cite do not seek to address questions of effectiveness or demonstrate the superiority of information professionals as valid researches to undertake literature searching when compared to another group of researchers, or even technology. With the exception of Meert et, al., the study designs are not evaluative or comparative study designs, therefore none explicitly help address the question of this key stage. The absence of such evaluative or comparative studies identifies a gap in the evidence.

This idea, and the gap in evidence it identifies, should be situated in the critique of this key stage and it should be acknowledged that the studies in Table 5 are authored by information professionals. The research community does not appear to dispute the role of information professionals in systematic reviews. There is no clear challenge to the work of an information professional from any other researcher group, and it would seem reasonable that very few researchers would challenge the guidance of the guidance documents on the grounds of appropriateness. Not least as a significant number of completed systematic reviews exist, which have included literature searching undertaken by information professionals. Better quality studies are, or further comparative evidence is, unlikely to sway opinion on including or not including information professionals, and so change the answer to the question posed in this key stage.

In determining the suitability of information professionals to undertake literature searching in systematic reviews, I suggest that the 14 non-comparative studies that cited in (see Table 5) might actually represent a challenge to this gap in evidence. The identification of these 14 non-comparative studies could suggest that the information professional is, or has been, in transition from a librarianship or enquiry service role to an information professional or researcher role. By setting out the tasks the information professional can undertake in a systematic review in the studies (see Table 3), the librarian/information professional may be attempting to locate or make sense of their role and contribution to a systematic review and potentially explain how their skills are relevant to the needs of researchers and systematic reviewers. The 14 studies all share a common theme ('Systematic reviews need systematic searchers' (146) being the title of one paper and the prevailing theme of the 14 studies) and they make sense if seen in this light since, as I set out above, they do not empirically

demonstrate the need for information professionals to literature search rather they set out what they can do.

Information professionals facilitate access to the resources that researchers need to do their work (158) but, as I set out later, in Chapter 4 (see 4.4.1), it is no longer enough to deliver the core information work in a traditional way (158, 159). Brettle sets out a form of transition agenda for a research-focused librarian (158). Through a series of vignettes, Brettle explores the opportunities and challenges of navigating the transition. Notably, information professionals are being encouraged to seek academic recognition and publications (158) and there is a need to 'future-proof' the role of their work (158). There is significant cross-over between the research possibilities of information science and the potential benefits that a research-active information professional academic group can offer to systematic reviews. These points would seem to develop support for an emerging role which the research-focused information professional could transition into.

If this idea of transition is valid, and further primary research is required to explore this, I reflect here that an infrastructure will need to be considered to support or further support it. Support in mentorship of research methods and study writing (perhaps advancing the existing frameworks of the Chartered Institute of Library and Information Professionals (CILIP) chartership portfolio but in an academic or research setting), support to pursue academic and scientific qualifications beyond the standard MA, and a recognition that much library and information-based research is presently unfunded or under-funded will be key (158). Health and medicine has led the way in defining and developing instrumental responses to information retrieval problems. Linking this evident need for robust, transparent and replicable literature searching in systematic reviews to a group of researchers has created an opportunity which these information professionals and researchers have been and are starting to fill in new ways.

I suggest above that the Cochrane Collaboration, through their support for and promotion of Trial Search Co-ordinators (TSCs), have exemplified a model of working with information professionals. Similar access to information professionals, in other types of reviews and topic areas, will be another part of this transition, making access to qualified researchers clear. As part of this, an evaluation of the TSC programme would be valuable.

Access to information professionals is a possible barrier to uptake. Anecdotally, researchers would appear to work with information professionals when they have a track-record of doing so (so they understand the value) and/or they have easy access to staff. A notable barrier, I suggest, is funding since it seems researchers are less willing to pay, or assume that the service provided is free, for a job they believe they can do themselves. Competent information professionals appear to be those who inspire confidence because they are trusted to literature search in an appropriate way (loosely defined by not missing key studies or overwhelming researchers with too many studies to screen) and are sufficiently experienced as a researcher themselves, that they understood the problems researchers face in undertaking and delivering systematic reviews. This would appear to demonstrate the benefit of an information professional to undertake literature searching and the value.

The lack of empirical evidence to firmly link the guidance to empirical studies leaves the question of who should literature search unaddressed. In recognition of the guidance recommendations, and the number of completed systematic reviews which have relied on information professionals, I indicate a potential refinement in the conventional approach. I set out, and I explore, this refinement in Chapter 4, Section 4.4.8.

Instrumental and conceptual tasks and distinctions

To further critique this key stage (Key Stage One), and the question of who should literature search, I explore the idea that there is a distinction between the instrumental tasks of literature searching and the conceptual tasks of literature searching in systematic reviews. Exploring this distinction might be another way to examine the question of who should literature search in systematic reviews. I suggest it may further develop the concept of librarians in transition, identifying potential support and training needs, which I identify above.

Instrumental tasks

The studies included in Table 5 put forward a role and a number of tasks which information professionals have self-identified as relevant to their training and expertise (see Table 5). These tasks are principally practical/instrumental tasks and relate to the delivery of literature searching and managing information in systematic reviews. The question of who should literature search in systematic reviews, in the

practical sense of who facilitates access to and knows how to literature search, would seem neatly aligned and addressed, even in the absence of empirical evidence.

Conceptual tasks

In linking the models of information retrieval theory to the instrumental work of literature searching, and with specific reference to the question of who should literature search in systematic reviews, I started to question if the focus on the instrumental tasks and skills of literature searching in systematic reviews is actually too restrictive a lens through which to address the question. Whilst attempting to determine the suitability of this key stage, there is a broader conceptual issue here which relates to the role of the information professional as a decision maker within a research team.

In literature searching for a systematic review, the instrumental tasks of literature searching and the conceptual task of planning the literature search are separate. The instrumental tasks are defined in the guidance and studies identified in this literature review but the broader conceptual tasks are less clear and less well assigned either to the information professional or another researcher. This represents a challenge and an opportunity.

I define conceptual tasks as: developing the research question into a literature search strategy (160), defining the information needs (4) and developing this in to a search enquiry (161), deciding which elements of a PICO structure (for example) to include in the bibliographic database searching strategy (162), and selecting search methods and planning the literature search process. I would situate these conceptual tasks in the domain of the information professional but as tasks they should lead on in conference with other researchers.

What does this mean for literature searching?

Craven and Levay argue that some of the information tasks set out above are often undertaken by the information professional and quite often in isolation (162). Researchers may not be clear how their research question has been developed into a search enquiry, and stakeholders may question practical matters, such the use of a search filter or selection of search terms. The answer to any of these points may be simple but, if the process is not clearly articulated or reported, or known by the research team, the purpose is lost and it may impact on the review, or reception of the review, later on.

This stimulates a broader question, should the information professional, or anyone, be taking decisions in isolation? If there is a broader recognition that systematic reviews are delivered by teams of researchers, and whilst researchers have allocated practical functions, decision-making should more generally be discussed and shared.

I question now if the question who should literature search in a systematic review should encompass the instrumental and conceptual distinctions briefly set out above.

Testing this idea: intervention effectiveness systematic reviews
A literature search for a clinical effectiveness systematic review broadly corresponds to PICO, making the breaking down of a clinical research question into a search enquiry easier to understand and, in most cases, guidance sets out the resources to be searched to identify studies, many of which are broadly routine. Input is needed on the suitable search terms from clinically trained researchers if the information professional is new to the topic area. Experienced researchers and information professionals may not feel that it is necessary to explore the conceptual aspects of literature searching that I identify above, for such seemingly straight-forward reviews, although I believe that clarity in decision making will support transparent reporting (162, 163).

As set out in Chapter 1, the literature searching component in systematic reviews beyond intervention effectiveness systematic reviews is potentially more complex. As I will explore further in this literature review, and the case studies in Section Two of this thesis, the conventional approach, as I set out in Figure 1, may not meet the needs of researchers undertaking other types of systematic review.

With expert skills in the knowledge of the techniques of identifying studies and data (143, 145, 150), and detailed knowledge of reporting criteria (153, 154), information professionals have trained in the skills required to undertake literature searching (139, 146, 147, 151) but the effectiveness of information professionals in this role has not been formally evaluated. There seems to be a greater need for the exploration of the process of literature searching, which I will explore in my case studies, and consideration of the role of the information professional. I believe that this identifies a genuine gap in knowledge and I will explore this further in Chapter 4, 4.4.8, and it indicates a gap in knowledge which could be explored through a potential refinement to the conventional approach. Namely, should the information professional be cast as the decision maker whilst working with the research team to develop the literature

searching approach? This allies the role with the skills and training of the information professional. I will explore this idea further in Chapter 4, Section 4.4.8.

Key Stage One: a summary of findings Guidance documents clearly state that information professionals are appropriate researchers to undertake literature searching in systematic reviews. The guidance reviewed in this literature review covers a range of review types, so the indication of the guidance documents is that information professionals can literature search for a number of different types of systematic review. The evidence for this guidance is not empirically derived.

The lack of clear empirical evidence is not necessarily a barrier to information professionals undertaking literature searching in systematic reviews. A number of completed systematic reviews exist in which information professionals have undertaken the literature searching and this could be used to demonstrate their suitability for the task and a theoretical alignment with guidance documents in the absence of empirical proof.

In the critique of this key stage, I explore the question of suitability in the context of who should literature search in a systematic review, with reference to the second research question of this chapter (see 3.4.1). Guidance documents suggest that information professionals are suitable researchers to undertake literature searching in systematic reviews and I did not identify a specific challenge to this claim. I suggest the possibility that the information professional is in transition to a new type of information-focused research role in systematic reviews. This transition may, in time, more clearly generate the evidence to match the claims made in the guidance, that information professionals are suitable researchers to undertake literature searching in systematic reviews. I identify a possible refinement to the role of the information professional in this key stage, and a second broader refinement for the information professional as decision maker, which I will explore in the next chapter.

This key stage is suitable in the conventional approach and it would seem to meet the needs of researchers. There is an evident need for literature searching and information professionals are indicated as suitable researchers in the guidance. The empirical evidence for this claim is not clear but the number of completed systematic reviews with the literature searching undertaken by information professionals is a compelling proof of concept.

Key Stage Two: Determining the aim and purpose of a literature search

The guidance for Key Stage Two

The aim of literature searching: Five of the nine guidance documents use adjectives such as 'thorough', 'comprehensive', 'transparent' and 'reproducible' to define the aim of literature searching (3, 25, 115, 118, 119). Analogous phrases were present in a further three texts, namely: 'to identify the best available evidence' (120) or 'the aim of the literature search is not to retrieve everything. It is to retrieve everything of relevance' (81) or '[a] systematic literature search aims to identify all publications relevant to the particular research question' (117). The Joanna Briggs Institute reviewers' manual was the only guidance document where a clear statement on the aim of literature searching could not be identified. **The purpose of literature searching** was defined in three guidance documents, namely to minimise bias in the resultant review (3, 115, 119). Eight of nine documents clearly asserted that thorough and comprehensive literature searches are required as a potential mechanism for minimising bias.

How the published studies correspond to the guidance for Key Stage Two

The need for thorough and comprehensive literature searches is consistent within the eight guidance documents that describe approaches to literature searching in systematic reviews of effectiveness. Reviews of effectiveness, test accuracy and prognosis, require thorough and comprehensive literature searches to transparently produce a reliable estimate of intervention effect. A comprehensive approach to literature searching to identify all relevant studies, and that this process has been 'transparently' reporting this process, increases confidence in the estimate of effect and the conclusions that can be drawn from the review (164). The supporting literature exploring the need for comprehensive literature searches focuses almost exclusively on reviews of intervention effectiveness and meta-analysis where comprehensive literature searching is understood to be an important component of the review process. Different 'styles' of review may have different standards and needs however; and an alternative, offered by purposive sampling, has been suggested in the specific context of qualitative evidence syntheses (114).

Whilst the guidance calls for thorough and comprehensive literature searches, it lacks clarity on what constitutes this, beyond the implication that all of the literature search methods in Table 6 should be used to identify studies. Egger et al., in a case

study evaluating the importance of comprehensive literature searches for trials in systematic reviews, defined a comprehensive search for trials as:

- a search not restricted to English language;

- where Cochrane CENTRAL or at least two other electronic databases had been searched (such as MEDLINE or Embase); and

- at least one of the following search methods has been used to identify unpublished trials: searches for (I) conference abstracts, (ii) theses, (iii) trials registers; and (iv) contacts with experts in the field (28).

Tricco et al. used a similar threshold of bibliographic database searching AND a supplementary search method in a review when examining the risk of bias in systematic reviews. Their criteria were: one database (limited using the Cochrane Highly Sensitive Search Strategy (HSSS)) and handsearching (165).

Together with the guidance, this would suggest that comprehensive literature searching requires the use of BOTH bibliographic database searching AND supplementary search methods.

Comprehensiveness in literature searching, in the sense of how much searching should be undertaken, remains unclear. Egger et al. recommend that 'investigators should consider the type of literature search and degree of comprehension that is appropriate for the review in question, taking into account budget and time constraints' (28). This view tallies with the Cochrane Handbook, which stipulates clearly, that literature searching should be undertaken 'within resource limits' (25). This would suggest that the limitations to comprehensiveness are recognised but it raises questions on how this is decided and reported (113).

The purpose of thorough and comprehensive literature searches is to avoid missing key studies and to minimize the introduction of bias in literature searching (3, 28, 74, 92, 95, 115, 119) since a systematic review and meta-analysis based only on published (or easily accessible) studies may produce an exaggerated effect size (165). Felson sets out potential biases that could affect the estimate of effect in a meta-analysis (166) and Tricco et al. summarize the evidence concerning bias and confounding in systematic reviews (165). Egger et al. point to non-publication of studies, publication bias, language bias and MEDLINE bias, as key biases (28, 30, 165-170). Comprehensive literature searches are not the sole factor to mitigate these biases but their contribution is thought to be significant (28, 81, 164). Fehrmann

suggests that 'the search process being described in detail' and that, where standard comprehensive search techniques have been applied, increases confidence in the search results (164).

Egger et al., and other study authors, have demonstrated a change in the estimate of intervention effectiveness where relevant studies were excluded from meta-analysis (28, 171). This would suggest that missing studies in literature searching alters the reliability of effectiveness estimates. This is an argument for comprehensive literature searching. Conversely, Egger et al. also found that 'comprehensive' searches still missed studies and could, in fact, introduce bias into a review rather than preventing it, through the identification of low quality studies then being included in the meta-analysis (28). Studies query if identifying and including low quality or grey literature studies changes the estimate of effect (170, 172) and whether time is better invested updating systematic reviews rather than searching for unpublished studies (173), or mapping studies for review as opposed to aiming for high sensitivity in literature searching (174).

The need for comprehensive literature searches is less certain in reviews of qualitative studies, and for reviews where a comprehensive identification of studies is difficult to achieve (for example, in Public health) (112, 114, 175-178). There is also a distinction (too) between aggregative systematic reviews and configurative or interpretative reviews. Literature searching for qualitative studies, and in public health topics, typically generate a greater number of studies to sift (screen) than in reviews of effectiveness (92) and demonstrating the 'value' of studies identified or missed is harder (179), since the study data does not typically support meta-analysis, or meta-analysis may not help explore the problem faced where aggregation of studies is not the primary objective of synthesis.

Nussbaumer-Streit et al. have evaluated whether abbreviated literature searches (as opposed to comprehensive literature searches) have an impact on conclusions across multiple bodies of evidence, not only on effect estimates (180). Nussbaumer-Streit et al. conclude that abbreviated literature searches are viable options for rapid evidence syntheses, if decision-makers are willing to trade the certainty from a comprehensive literature search and systematic review, but that decision-making which demands detailed scrutiny should still be based on comprehensive literature searches (180, 181). This would appear to suggest that decision makers and users of systematic reviews are willing to trade the certainty from a comprehensive literature

search and systematic review in exchange for different approaches to evidence synthesis (181), and that comprehensive literature searches are not necessarily a marker of literature search quality, as previously thought (113). Different approaches to literature searching (48, 80, 92, 94, 95, 182) and developing the concept of when to stop searching are important areas for further study (94, 113).

Determining suitability: A critique of Key Stage Two

Guidance documents and studies shared a generally unified understanding of the aim and purpose of systematic literature searching in a systematic review. These were for a literature searches that are: 'comprehensive', 'transparent' and 'reproducible' (3, 115, 118, 119) with the aim of minimising the introduction of bias (3, 115, 119). I will adopt the concepts of 'comprehensive' and 'transparent' and 'reproducible' to structure my critique of this key stage as I determine the suitability of the conventional approach in systematic reviews.

Comprehensive literature searches: a definition?

A comprehensive search for studies represents what is possibly best-known about systematic reviews. Comprehensive literature searching is informally understood to mean a literature search which does not miss any relevant studies. This has previously, and variously, been referred to as 'exhaustive' literature searching (48). Both comprehensive and exhaustive literature searching mean to convey a similar concept: that the findings of the systematic review can be trusted because they are based on a literature search which has exhaustively identified all the relevant studies.

The guidance and studies identified in the literature review appear to support the understanding of a need for comprehensive or exhaustive literature searches. A criticism could, however, be made of the lack of clarity in defining what constitutes a comprehensive literature search and how a researcher can measure if a literature search is a comprehensive literature search or not. Criticism might also be raised if this understanding is still relevant generally, and if it applies equally, to all types of systematic review.

The Cochrane Handbook appears to approach the question of what constitutes a comprehensive literature search by stipulating a mandatory minimum set of bibliographic resources which must be searched in a Cochrane review (25). A Cochrane author cannot publish a Cochrane review which does not include – as a minimum – a search of their group's trials register, a search of the Cochrane Central

register of trials and MEDLINE. Embase is highly desirable, where access is possible (25). This would appear to suggest that the number of resources searched, or the number of 'appropriate' databases searched, is a marker of a comprehensive literature search.

The studies by Egger et. al. and Tricco et al. helpfully define what they consider a comprehensive literature (28, 165). Egger et al. and Tricco et al. both consider a comprehensive literature search to be where a researcher searches bibliographic databases AND uses a supplementary search method in combination (28, 165). There is support for this in the Cochrane handbook too, where the aim of supplementary search methods is to identify unpublished studies and data (25). This would appear to suggest that the combination of bibliographic database searching AND the use of supplementary search methods is a marker of a comprehensive literature search.

The concept of searching a minimum number, or minimum of appropriate databases, combined with a search approach which utilises both bibliographic databases AND supplementary searching offers a working definition of what might constitute a comprehensive literature search. This combination approach is implicitly acknowledged in guidance too (25), and so it is illustrated in the conventional approach in Key Stages 4 and 5. It would seem likely that, if this understanding were tested against a sample of Cochrane or other bio-medical systematic reviews, it would appear as valid.

This still seems an inadequate definition since it is not really clear how many databases are enough and which are the appropriate ones (172). The guidance on supplementary search methods also notably varies between guidance documents (see Table 6 and Key Stage Five) and there is no single and clear review of supplementary search methods combining the findings in one place. This is a gap in research knowledge I fill in Chapter 5.

Other factors may also develop this definition of comprehensive literature searching. In considering the suitability of this key stage, further factors may be worthy of critique since they may help elucidate how researchers understand and measure comprehensive literature searching. These factors are not yet acknowledged as important in appraising the comprehensiveness of literature searches and I believe this critique makes a unique contribution in developing this field.

Quality

The quality of a literature search does not appear to form part of understanding the comprehensiveness of a literature search. This is initially surprising since a literature search with typographical errors, or with limited search terms, is likely to miss studies, thereby affecting retrieval and the comprehensiveness of the literature search. What constitutes a 'good' literature search is unclear (113), but peer review can be used to determine the quality of a literature search with feed-back helping to identify errors or deficiencies in the search strategy or search approach (183). This suggests that peer review should be part of determining the quality and comprehensiveness of a literature search.

This issue is considered later in the thesis: I identified peer review as a critical factor in my systematic reviews of effectiveness measures (Chapter 6) which found that the quality of literature searches was rarely appraised when determining the effectiveness of one search compared to another (52% of studies in the reference standard group, and 6% in the index test group were peer reviewed) (5). This in spite of the fact that validated checklists have been developed and are available (e.g. PRESS (183)). The complexity of reviewing a search approach may be barrier to up-take of peer review (163) and researchers have called for 'search narratives' (162) to help explain the contextual and conceptual structure of a search and inform peer review (163). It is hoped that these search narratives will develop peer review, potentially removing a barrier to its uptake and increasing the speed of review, but the use of search narratives has yet to be appraised. Whilst any form of peer review is useful, peer review against a structured checklist (such as PRESS (122, 183)) makes clearer and more transparent statements about the appraisal of the underlying literature search.

In determining search quality, an additional and clear report or statement about the potential limitations of a literature search would also inform a notion of quality. A researcher could identify any practical issues arising with the literature search which might impact on the comprehensiveness of the literature search. For instance, setting out that the literature search was undertaken in time-limited circumstances, or the fact that they did not have access to and did not search a particular database. This would help inform the understanding of a literature search. By clearly grounding limitations in an accessible way, it removes the chance that limitations might be read as errors or inadequate searching. It is important to note that the literature searches

reported in a systematic review represent the final draft of the search strategy, often without context (163). Peer review and a clear statement of limitations would help a reader explore the 'quality' of a literature search or at least understand the context.

Time

Neither guidance nor studies report the amount of time or resource use which should be invested in literature searching in order to produce a comprehensive literature search (6). In a previous edition of NICE Public Health guidance, NICE suggested that 3-4 weeks was an appropriate amount of time to develop and undertake a literature search for systematic reviews to inform public health guidance (184). This marker on timing was not repeated when the unified guide to NICE methods was published in 2014 (120). It is not clear why.

To minimise the introduction of bias in literature searching for systematic reviews, it is implied (if not stated) that equal amounts of time and resource will be invested in searching bibliographic databases as on supplementary search methods. Spending more time on bibliographic database searching than any other method would seem to privilege studies published and indexed in bibliographic databases, compared to unpublished studies, potentially introducing bias (publication bias most specifically) into a review (6, 185). I have proposed the idea that we were over-searching and over-prioritising bibliographic databases elsewhere (185). I argued, specifically, that researchers spend too long searching bibliographic databases leaving not enough time to undertake supplementary search methods (185). As part of the thesis, I identified a gap in knowledge that there was not a review of supplementary search methods and so no report of the time needed to effectively undertake these types of search methods. This review is reported in Chapter 5 but I found the reporting of timing in the primary studies to be limited (6).

The amount of time invested in literature searching would seem linked to understanding the comprehensiveness of a comprehensive literature search. A literature search undertaken and developed in one hour will likely produce different results than a literature search developed over two weeks. Intuitively, the literature search developed over two weeks would feel more reliable simply because more time has been invested in ensuring it identifies relevant studies, that it does not identify too many irrelevant studies, and that the approach has been tested in a variety of bibliographic databases and supplementary search methods have had time

to be developed. It feels important to state here that the context of this discussion is of systematic reviews and not rapid reviews or clinical enquires.

In the absence of clarity on how much time should be invested in literature searching, and more generally, the absence of any guidance – either in the form of guidance or retrospective, generalised summaries of the time spent searching per search method – it is not clear how time could be used to develop an understanding of comprehensive literature searching. This indicates a gap in knowledge simply because it is not clear how long literature searching in systematic reviews should take and how long individual tasks within a systematic review do take. Researchers are starting to explore this topic area, notably Bullers et al. and Haddaway et al. (186, 187). Their research could produce summary estimates which could inform resource allocation in systematic reviews. The search summaries proposed by Bethel could be further developed to capture timing (188), as I have identified in Chapter 5 of this thesis (6).

Within the context of time there are numerous confounders to consider, such as relevant experience of the information professional, access to resources, and the amount of time available to literature search generally due to funding. These confounders should form part of the wider discussion above and also as a part of the limitations of any literature search.

Measuring comprehensive literature searches
As I will explore in Chapter 6, common metrics to evaluate comprehensive literature searches focus on effectiveness. The metrics of sensitivity, specificity and accuracy are used to determine the effect of one search when compared to another. In using these metrics to understand comprehensive literature searches I see some issues which I identified in my systematic review (reported in Chapter 6) but which I summarise here for context.

First, these metrics used (sensitivity, specificity and accuracy) are summative metrics (5) which can only be calculated when the process of literature searching and screening is completed and so at the end of the process of review. It is not clear what represents a good (or good enough) score in determining sensitivity, specificity or accuracy, or what happens if a score is considered poor (see Chapter 6). This leaves the notion of what constitutes a comprehensive literature search unclear (5, 189).

Secondly, the analysis required to calculate these metrics is necessarily sample dependant (190). A crux of determining sensitivity in an information retrieval context is the recognised impossibility of not knowing definitively if all relevant studies have been identified. With this in mind, it is impossible to know if a search is truly comprehensive or has identified all possible studies or study data (189). Doshi and Jefferson demonstrate this point, and they set out the consequences too, by comparing their Cochrane review, which followed the guidance of the Cochrane handbook, to a systematic review based entirely on published clinical study reports. Doshi and Jefferson found significant differences in the review based on CSRs which had not been observed in the original Cochrane review (64-66, 68). This illustrates not only the point that it is impossible to comprehensively identity all relevant data (since there is some data which is difficult to access without due process) and it also illustrates the potential changes of finding new study data.

This also relates to the question of when to stop searching in order to ensure comprehensiveness, since there is no agreed standard, metric or stopping rule, recommend by guidance or accessible in other studies reviewed in this chapter (6, 113). Researchers presently appear to work from the same informal understanding, broadly aligned with Egger et al. and Tricco et al., that a search is stopped when a certain number of named databases have been searched and a number of supplementary search methods have been undertaken. This does not seek to locate the idea of comprehensiveness in the studies identified and the question whether all the relevant studies have been identified. It is not clear that there is a satisfactory way to address this issue. Having explored this topic in greater detail in this thesis (Chapter 6), I now suggest face validity is truthfully the prevailing test of comprehensive literature search at this time. By 'face validity' I mean, does the literature search look comprehensive; either through the variety of search terms used, the length of the search strategy (i.e. the number of lines in a MEDLINE search strategy), or the number of databases searched.

In the absence of an agreed understanding of what a comprehensive literature search means, it is left to researchers and users of evidence to determine their own standards, and there is evidence that researchers and research users may have different priorities with respect to comprehensiveness (180, 181). Linking comprehensiveness to the amount of searching, and use of resources or search methods, is intuitive, but it implies that searching more databases and undertaking

more supplementary searching (potentially linked to spending more time literature searching) produces a more comprehensive and therefore a better systematic literature search. Decision makers may be willing to accept less rigour (i.e. less comprehensive literature searches) in exchange for more rapid results (180). It would then fall to the researcher to articulate the risks in accepting a lower degree of rigour.

Is comprehensives context specific?

The sources I cite to develop a definition of comprehensive literature searching (The Cochrane handbook, Egger et al. and Tricco et al.) are all taken from and based on intervention effectiveness systematic reviews. This definition of comprehensive literature searching, and what constitutes 'comprehensiveness', is therefore limited to that specific context. Beyond the evaluation of intervention effectiveness, the need (and possibility) of achieving a comprehensive identification of studies is questioned. The emerging recognition is that it may not be possible to demonstrate that a comprehensive literature search has been undertaken (5, 113) and the mechanism to demonstrate or examine this is less accessible. Researchers have used statistical meta-analysis to demonstrate the change in summary estimates in intervention effectiveness systematic reviews. In other types of systematic reviews, it is less clear how researchers can examine and illustrate the effect of 'missing' studies in the same way and so it is harder to explore an understanding of the 'true' effect of missing studies. I identified this as a gap of knowledge in my second case study (Chapter 8) and it represents one of the novel contributions to knowledge arising from this thesis (see 9.7). I used CERQual to determine the confidence in findings between two different literature search approaches, which is a potential method to examine the effect of 'missing' studies or determining the 'value' of studies on the synthesis (see Chapter 8) (77). This is the first time that CERQual has been used in this way and I explore this in greater detail later in the thesis.

There is further and evolving discussion that a comprehensive literature search may not be an attainable goal in some types of systematic reviews (112, 191). A new or alternate definition, and an evaluation of what a relevant rather than comprehensive literature search could look like and mean for the synthesis of studies, would be a possible starting point. This may lead researchers to evaluate their understanding of what a systematic review means in an alternative context and whether the 'traditional' understanding of comprehensive literature searching is necessary or important. A review that follows an *a priori* system to literature searching, that does not perhaps

focus on an exhaustive list of bibliographic databases and supplementary searching, could still be a systematic review. There is, after all, a system of methodological investigation that was pre-specified and followed. So long as the limitations of the approach are made clear, this would seem to meet the needs for decision makers who might be willing to accept less certainty associated with a traditional understanding of comprehensiveness (180).

It is likely that comprehensive literature searching means different things in different topic areas and for different types of systematic review, possibly linked to the purpose of the review. In intervention effectiveness reviews, the prevailing marker of a comprehensive literature search is not missing studies. As I explore in this critique, and in Chapter 6, it is not presently possible to know if studies have been missed until the end of literature searching, or more often, the end of a systematic review. It may never be known if a systematic review has missed a relevant study.

Different types of review may have different levels of tolerance or comfort for missing studies and are therefore not necessarily focused on a comprehensive literature search at all. This might suggest that other markers, such as quality, time or suitability, may increasingly play a role in determining the role or purpose of a literature search. A focus on the number of databases searched (172), and the number of supplementary search methods used in combination, is not a helpful marker for a comprehensive literature search. It possibly encourages over-searching, where additional databases are searched to give the impression of a comprehensive literature search, but with little corresponding clarity on the utility and value in terms of identifying additional and relevant studies. Without clearer guidance on the use of supplementary search methods (see Key Stage Five and Chapter 5) it is possible that these search methods are not used to their best advantage either.

I identify comprehensiveness as an important question in literature searching. I explore measuring it in Chapter 6 of this thesis and I question alternative ways to measure and value literature searches in Chapter 6 and the case studies presented in Chapter 7 and Chapter 8, in response to the calls of other researchers (192, 193). In this key stage, and my thesis, I am starting to question if the idea of value, in terms of the value of studies identified, and any corresponding effect on the synthesis, is a potential way to re-orientate the discussion on comprehensive literature searching.

Comprehensiveness Summary: what I seek to explore is the idea that comprehensiveness, or comprehensive literature searching, is understood in a potentially limited context. Systematic reviews of intervention effectiveness may have fixed on the idea of not missing studies. Whilst it is important to generate a reliable estimate of intervention effectiveness, how this is evaluated, and if it is reliable, is in question. As the work of Doshi and Jefferson illustrates, it is impossible to identify all the relevant studies and study data, and so comprehensiveness as an on-going marker of literature search quality, is worthy of being reconsidered (113).

It is also not clear if comprehensive literature searches are necessary in all types of systematic reviews yet the guidance documents and studies appear to work to this standard or expectation. Researchers have been questioning the suitability of this (112) but, to explore this idea, researchers need to better understand the risk of missing studies on the synthesis of studies. I identify this as gap in knowledge and I seek to explore it in my case studies presented in Chapter 7 and Chapter 8 through the identification and illustration of value, based on the value of studies identified by literature searching to the synthesis of studies.

On this basis, I am not clear that a uniform understanding of comprehensiveness in literature searching is suitable. Whilst the understanding might meet the needs of reviews of intervention effectiveness, and therefore the conventional approach might be appropriate in this context, I am not certain that it is suitable in all types of review. This has identified a need for further work, both to develop my understanding of this issue, and also to examine how measurement of comprehensive literature searching is understood. I will explore these points in Chapter 6.

Transparent literature searches

This section considers transparency and reproducibility in the context of developing and understanding a comprehensive literature search. Transparency and reporting (including reproducibility) are discussed more completely in Key Stage Eight.

Transparency and reproducibility are considered key markers of a systematic review (194). Transparently reporting the literature searches in such a way that they could be repeated makes a clear statement on the methods used to identify studies and the extent to which this has been done. This has been linked to increasing confidence in systematic reviews (164). Whether this is actually the case has been little explored.

Studies that link transparent searching to increased confidence identify both a limitation and an area for further research. Whilst they suggest that increased confidence is an outcome of transparently reported literature searches, they have not themselves repeated any literature searches in their analysis, to see if the increase in confidence is merited. That the search is reported does not mean that it is of good quality, or that it is sufficient, or that it is valid, or that it can be replicated in practice (151).

Replicating literature searches may not be straight-forward. The guidance on reporting literature searches in PRISMA guidance is the least detailed of step in the reporting process (194). Rader et al., in the findings of a survey of 260 information professionals to determine issues with following PRISMA reporting guidance (I was a co-author of this work), found that what to report and to what level of detail was a confusing area of PRISMA guidance (195). A study that sought to examine the possibility of replicating literature searches could not only examine confidence, as I suggest above, but also, in replicating someone else's searches, begin to make judgments about what literature searches should be reporting to ensure transparency and to ensure that literature searches are truly replicable. It would be my suspicion that, whilst literature searches are reported, many could not be replicated in practice, or without contacting the original author for further input. I have developed Craven and Levay's idea for search narratives with this in mind to address this gap (163).

The ability to transparently report literature searches is potentially different for searches of bibliographic databases and supplementary search methods yet the same standards of reporting appear to apply for all literature search methods and reporting guidelines, such as PRISMA (196). Bibliographic databases are, generally speaking, easy to export studies from and export the literature search strategies used. This functionality gives the clearest possible sense of transparently reporting literature searches because there is a clear association between the search terms used and the results found. In spite of being the 'easiest' search method to transparently report, study authors still find limited or a lack of sufficient reporting (194, 197, 198).

Moving beyond bibliographic databases, supplementary search methods are less easy to transparently report (199) and may be impossible in some cases (see Chapter 5 for further discussion) (6). Guidance is lacking, and studies too, to determine how to report the methods and results of author contact, for instance, but

developments are being made in some supplementary search methods. For example, Briscoe has set out guidance for web-searching in HTA (200) and he has evaluated web-search reporting undertaken in Cochrane reviews (201). Haddaway et al. have put forward a rapid method to increase transparency of web-searching (202). Haddaway et al's method requires validation, but would seem to offer at least another method to raise the standard of reporting in web-searching.

There is a potential tension in reporting between searches that can be transparently reported and repeated (i.e. bibliographic database, citation chasing) and searches which cannot be transparently reported and repeated with the same effect (author contact, handsearching). Method-by-method reporting guidance would potentially be a useful mechanism to account for the difficulties between some search methods and as a way to generalise and standardise literature search reporting. It may be time to acknowledge that there are some searches which it would be impossible to recreate and produce the same effect but that this does not lessen the importance of describing the methods used and what was found. Perhaps a summary of guiding principles would be a useful place to start, so it is clear why reporting is important, identifying which methods are problematic, and what as a minimum can be reported so at least future researchers can understand what was done even if they cannot exactly replicate it.

Transparently reporting literature searches also relates to an understanding of the comprehensiveness of individual literature searches in context. By transparently reporting the literature searches used, researchers also demonstrate the comprehensiveness of their literature searching. This makes clear statements not only about the extent to which they have literature searched but also the reliability of their findings. Other researchers can view the underlying searches and make their own judgements about the suitability of the searches linked to the purpose of the review. This could usefully be extended further if more reflective detail on the searches was provided in an annex or on-line supplement with a clear statement on any potential limitations.

Transparency Summary: The case for transparency in the reporting of literature searching is clearly made in the guidance. The dual roles are demonstrating rigour/quality (and possibly comprehensiveness) and replicability. It is situated in the idea that the underlying literature searches could be repeated to verify the findings and conclusions of the review in question. Rader et al. identified a number of areas

where researchers are not clear about what should be reported in literature searching (195). This suggests that the guidance documents do not yet provide enough detail.

What is also unclear is the extent to which searches that are transparently reported are actually replicable. It seems an assumption that because the literature searches are reported that they could be repeated and this point has not been subject to evaluation. It may prove worthwhile to test and it could inform the development of updates to guidance on what to report and how. This will be picked up further in Key Stage 8.

Key Stage Two: a summary of findings

Comparing guidance with studies elucidated a definition of what constitutes a comprehensive literature search. This is a valuable definition but there is emerging evidence that comprehensive literature searching may mean different things depending on the purpose of the review and it may be context dependant depending on the possibility of identifying studies.

What is still not clear is how to measure a comprehensive literature search, or if a comprehensive literature search is a valid marker to determine search quality. Without a clear understanding of what comprehension in literature searching means, or looks like, it is unclear how this should be appraised.

On this basis, I explore comprehensive literature searching further in Chapter 6 and consider new ways to examine the effect of literature searching in my case studies in Chapter 7 and Chapter 8.

It is not, therefore, clear if the idea of comprehensiveness in this key stage meets the needs of and is suitable for researchers undertaking systematic reviews. For transparency and reproducibility, these are core requirements of a systematic review. Whilst studies identify confusion in understanding reporting guidance, the concept of transparent reporting of literature searches would seem an important and suitable requirement of the conventional approach and other approaches to literature searching too.

Key Stage Three: Preparing for the literature search

The guidance for Key Stage Three

Six guidance documents provided guidance on preparing for a literature search (3, 25, 81, 115, 117, 118). The Cochrane Handbook clearly stated that Cochrane authors (i.e. reviewers) should seek advice from a trial search co-ordinator (i.e. a person with specific skills in literature searching) 'before' starting a literature search (25).

Two key tasks were perceptible in preparing for literature searching according to five guidance documents (3, 25, 81, 116, 118). First, to determine if there are any existing or on-going reviews, or if a new review is justified (3, 116); and, secondly, to develop an initial literature search strategy to estimate the volume of relevant literature (and quality of a small sample of relevant studies (115)) and indicate the resources required for literature searching and the review of the studies that follows (115, 118). This work is sometimes known as scoping.

Three guidance documents summarised guidance on where to search to determine if a new review was justified (3, 81, 116). These focused on searching databases of systematic reviews, namely: The Cochrane Database of Systematic Reviews (CDSR) and the Database of Abstracts of Reviews of Effects (DARE), institutional registries (including PROSPERO), and MEDLINE (3, 116). It is worth noting, however, that as of 2015, DARE (and the NHS Economic Evaluations Database (NHS EEDs)) are no longer being updated and so the relevance of these resources will diminish over-time (203). One guidance document, 'Systematic reviews in the Social Sciences', noted, however, that databases are not the only source of information and unpublished reports, conference proceedings and grey literature may also be required, depending on the nature of the review question (81).

Two guidance documents reported clearly that this preparation (or 'scoping') exercise should be undertaken before the actual search strategy is developed (115, 118) and before the final topic is finalised.

How the published studies correspond to the guidance for Key Stage Three

The guidance documents offer the best available source on preparing the literature search with the published studies not typically reporting how their scoping informed the development of their search strategies nor how their search approaches were developed. Text mining has been proposed as a technique to develop search

strategies in the scoping stages of a review, by 'mining' relevant documents and suggesting search terms by frequency, although this work is still exploratory (60). 'Clustering documents' and word frequency analysis have also been tested to identify search terms and studies for review (204, 205). Effective strategies for preparing for literature searches and scoping constitutes an area for future research.

Determining the suitability: A critique of Key Stage Three

Two tasks were perceptible from the guidance on preparing to literature search: The first task was to determine if there are any existing or on-going reviews and the second task was to develop an initial literature search to estimate the number of studies likely to be returned and indicate the resources (selection of databases and other sources to search) required for review. The studies provided little guidance or support for the development of this key stage.

Perhaps the first criticism is the naming of this key stage. I did not name the key stages myself (see 3.4.2) rather, the naming of the key stages was drawn from the guidance documents. 'Preparing for the literature search' is not a tittle which accurately or helpfully articulates the various tasks a researcher would undertake in this key stage. This may mean that it does not help researchers understand the purpose or importance of this key stage, or what is expected of them in the wider context of the conventional approach, as it is reported in the literature review and illustrated in the process model.

This stage is colloquially known as scoping (206). A similar concept but scoping is a more accessible and familiar handle to name this key stage, given the tasks to be undertaken. The term 'scoping' somehow seems to better capture the uncertainty of preparation which is reflected in the literature explored below.

Does guidance reflect research needs and practice?

The guidance on preparing to literature search for a systematic review was surprisingly limited. This is perhaps explained by the fact that, at key stage one, guidance supports involving information professionals in the process of literature searching. The tasks identified in the studies cited in the literature review (see Table 5) are an acknowledged part of the information professional's role in a systematic review (142, 143, 150, 152) so the knowledge required for this key stage is tacit, being part of what information professionals know how to do, although this is not articulated in detail. It is not clear how information professionals acquire the

knowledge to undertake this task other than through experience. This might explain the lack of detailed methods and guidance available and it sets out a challenge to information professionals to articulate their role at this key stage of undertaking a systematic review.

Another possible explanation for the limited guidance, is that the guidance documents that make the clearest statements on preparing to literature search, are for intervention effectiveness systematic reviews. The guidance on literature searching for intervention effectiveness systematic reviews has benefited from greater methodological input both in terms of developing the methodology of literature searching and undertaking systematic reviews more generally (see Chapter 1). The lack of guidance could be explained, in short, because the guidance documents feel that this key stage is simple and straight-forward and that the work indicated is simply procedural. This may, therefore, represent a separation between planning in intervention effectiveness systematic reviews and other types of systematic reviews. In my experience, this would be a rational explanation. As I found in my case studies (see Section Two), these needed far more time for planning than for a systematic review of clinical effectiveness.

Whatever the explanation for the limited guidance, it is, however, a criticism of the guidance documents, studies, and this key stage generally, that no general guidance is given for the tasks identified. Specifically, in relation to the tasks identified by guidance at this stage, namely: in estimates of the time needed to sufficiently prepare to literature search and the methods used to estimate the likely number of studies a literature search will identify and which will need screening. As I explore and find in my case studies (Chapters 7 & 8), the guidance does not accurately convey the amount of work, and team work specifically, which this section of developing a systematic review necessitates. It might be that information professionals are not clear about the amount of work that goes into this aspect of a systematic review. A lack of clarity on the work and time needed is then reflected in the lack of guidance and minds of researchers.

Time

There is a temptation to see the scoping stage of a systematic review as a rapid process (207). Arskey and O'Malley suggest that the temptation is to rush the scoping stage of a review (207). Read alongside the work of Armstrong et. al., the prevailing wisdom appears to be to 'slow down' (207, 208). Guidance documents

could play an important role here by providing estimates of the time indicated for scoping and for the tasks of literature searching in a systematic review. This would develop the key stage by providing an indication of the time needed to scope, develop and finalise a literature search. This would also provide a mechanism by which researchers could judge if appropriate time has been invested in developing the approach to literature searching.

Evaluating the time spent literature searching would benefit from examination. There comes a point where too much time might be invested in planning a literature search and the realisation that a simpler search, with a higher volume of studies to screen, may have been a better use of time. This is a trade-off between planning time and screening time. I recognise that this was a potential risk in my case studies (see Section Two). By comparing systematic reviews that invest time in the preparing to literature search, to reviews which 'just got on with it', and looking at outcomes such as: number of studies identified v number of studies uniquely identified (77), the time taken (6, 187), and quality markers such as rigour in reporting (155), researchers could begin to examine if allocating additional time to prepare a literature search is a valuable use of researcher time, or not.

A recent study by Bullers et al., based on a survey of 185 librarians who have undertaken literature searching for systematic reviews, examined the time that medical librarians spent on specific information retrieval tasks within a systematic review (187). Whilst the study makes a unique contribution in answer to the question 'how long do *librarian* indicated tasks take in a systematic review', the study authors have adopted a novel categorisation to stratify the timings that they report. Instead of describing their findings in the language and commonly understood tasks of systematic review methodologically, and with reference to leading guidance documents specifically, they report timings by 'sections' and then by tasks. The tasks indicated included 'interviews' and 'instruction', and the purpose and context of these are not made clear, nor are they situated in common systematic review practice. It is unclear how useful this paper is therefore as the results are not directly generalisable to guidance-informed common practice.

Researchers may be better placed to fill this evidence gap by providing empirical evidence on the time taken to prepare to literature search within the context of their systematic reviews. If researchers reported in the supplementary material of their reviews the time taken in preparation (and for other research tasks such as

screening), and any lessons learned of their preparation stage which might aid future reviews, in time, the results of timing from similar reviews could be pooled and used to broker knowledge. This would permit the generation of approximate timings for tasks, perhaps closer to the informal suggestion that I make above, that scoping and developing literature searches for public health reviews may take between 2-4 weeks (184).

Estimating the number of studies

Guidance documents recommend undertaking scoping searches to estimate the volume of studies likely to be retrieved for the systematic review but there is a lack of specific guidance or recommendation of a method on how to estimate the volume of studies. Information professionals appear to approach scoping and the estimation of studies to screen informally, using the number of studies identified from bibliographic databases as a proxy, informally discounting possible de-duplication between resources (209). Haddaway and colleagues recently launched The PredicTER (Predicting time requirements for evidence reviews) decision-support tool. The tool aims to estimate the likely screening and time requirements for researchers undertaking Collaboration for Environmental Evidence (CEE) systematic reviews, based on data from previously completed CEE systematic reviews (186). Whilst specific to CEE reviews, and presently under peer-review, this could be an appropriate solution to a problem that presently lacks a clear method or guidance. A method or tool to estimate the likely retrieval of studies would be a valuable contribution for knowledge. The certainty with which estimates can be predicated is likely to vary between review types and the contribution of search methods.

With reference to time, and the work undertaken in this scoping stage, the guidance offered, and studies identified, I suggest does not go far enough in realistically setting out and recognising the range and amount of work that happens during this key stage. It should be noted that this might vary between reviews and the experience of the researchers involved. Armstrong et al. encourage researchers to consider scoping as an opportunity to collect and organise important background information (208). For the information professional, this means scoping and working with the review team to determine a unified understanding of the research question (4) and which studies meet and do not meet inclusion, and why, as the scoping progresses. This work should be situated in the guidance of The Cochrane Handbook which suggests that scoping should ideally happen before a research question is firmly

fixed (25). This scoping time should also include developing definitions of key terms, which may be of particular importance for complex interventions where a lack of standardised language is present in the studies (50).

Recording the work undertaken and these decisions as it relates to the development of the literature searching is important (206, 210). Depending on the nature of the review, and the resources available, this preparation and scoping stage can be taken further still. O'Brien et al., in a web-based survey of scoping study methodology, identified a need for a consultation phase (211). This consultation phase involves sharing findings from the literature and additional expertise from stakeholders and it is recognised by other researchers as an important part of developing a scope (211, 212). This consultation stage better reflects my experience of scoping and developing a protocol for a complex systematic review and it broadly encapsulates the other issues that I explore above but that are omitted from the guidance in handbooks. A consultation approach was used in both of the case studies reported in Section Two of this thesis.

Key Stage Three: a summary of findings A number of gaps in knowledge exist in guidance and understanding for this key stage. The findings suggest that the guidance and knowledge within studies is limited in terms of helping researchers prepare to literature search. The guidance provided may support reviews of intervention effectiveness but further guidance on timing, estimating yield, and working with researchers and stakeholders is indicated for other types of systematic reviews. I explore these ideas further in the case studies that I present in Chapter 7 and Chapter 8.

It is unclear whether the information on this key stage is suitable to meet the needs of researchers. The suitability of this key stage is clearest for reviews of intervention effectiveness. Elsewhere, and potentially in other types of systematic review, there seems a clearer further and future role for the development of this key stage to explore how researchers plan for reviews, scope the evidence, and approach more complicated reviews.

Key Stage Four: Designing the search strategy

The guidance for Key Stage Four

The Population, Intervention, Comparator, Outcome (PICO) structure was the most commonly reported structure promoted to design a literature search strategy. Five guidance documents suggested that the eligibility criteria or review question will determine which concepts of PICO will be populated to develop the search strategy (25, 81, 118-120). NICE guidance promoted multiple structures, namely PICO for effectiveness questions, SPICE (Setting, Perspective, Intervention, Comparison, Evaluation) for social science questions, and multi-stranded approaches for complex systematic reviews (120).

With the exception of The Joanna Briggs Institute reviewers' manual, the guidance offered detail on selecting key search terms, synonyms, Boolean language, selecting database indexing terms and combining search terms. The CEE handbook suggested that 'search terms may be compiled with the help of the commissioning organisation and stakeholders' (115).

The use of limits, such as language or date limits, were discussed in all documents (3, 25, 81, 115-120).

How the published studies correspond to the guidance in Key Stage Four

Structure

The guidance documents described above typically relate to reviews of intervention effectiveness, so PICO – with its focus on intervention and comparator - is the dominant model used to structure literature search strategies (160). PICOs – where the S denotes study design - is also commonly used in effectiveness reviews, where the S relates to study design literature search filters (3, 160). As the NICE handbook notes, alternative models to structure literature search strategies have been developed and tested. Booth provides an overview on formulating questions for evidence based practice (160) and he has developed a number of alternatives to the PICO structure, which may be more appropriate for reviews including study designs beyond RCTs of intervention effectiveness. Namely: BeHEMoTh (Behaviour of interest; Health context; Exclusions; Models or Theories) for use when systematically identifying theory (176); SPICE (Setting, Perspective, Intervention, Comparison, Evaluation) for identification of social science and evaluation studies (213) and, working with Cooke and colleagues, SPIDER (Sample, Phenomenon of Interest,

Design, Evaluation, Research type) for qualitative evidence synthesis (214). SPIDER has been compared to PICO and PICOs in a study by Methley et al. (215).

The NICE handbook also suggests the use of multi-stranded approaches to developing literature search strategies (120). Glanville developed this idea in a study by Whiting et al. (216) and a worked example of this approach is included in the development of a search filter by Cooper et al. (50). Multi-stranded approaches have been tested in public health to capture related issues of reference to the population under review. For instance, in a review of tuberculosis testing in hard-to-reach populations, a multi-stranded search was used to identify not only relevant population terms (e.g. prisoners) but also search terms for potentially relevant settings (e.g. prisons) and any inter-linked concepts (e.g. substance abuse) too (50).

Conceptual and objective approaches

Hausner et al. provide guidance on writing literature search strategies, delineating between conceptually and objectively derived approaches (217). The conceptual approach, advocated by and explained in the guidance documents, relies on the expertise of the literature searcher to identify key search terms and then develop key terms to include synonyms and controlled syntax. Hausner and colleagues set out the objective approach (217) and describe what may be done to validate it (218). Objective approaches develop search syntax from studies that are known to be of relevance to the research question and develop a search strategy and approach on this objective basis (218).

Limits

All of the guidance documents offer direction on the use of limits within a literature search. Limits can be used to focus literature searching to specific study designs or by other markers (such as by date) which limits the number of studies returned by a literature search. The use of limits should be described and the implications explored (165) since limiting literature searching can introduce bias. Craven et al. have suggested the use of a supporting narrative to explain decisions made in the process of developing literature searches and this advice would usefully capture decisions on the use of search limits (162). This work has recently been updated by Cooper et al. with publication of a worked example (163).

Determining the suitability: A critique of Key Stage Four

The prevailing focus of the guidance documents reviewed was on effectiveness evaluation. The PICO mnemonic, with its focus on intervention and comparator, dominated the guidance which suggests that PICO is aligned most closely to intervention effectiveness systematic reviews. This suggests that the guidance available for this key stage is suitable for systematic reviews of intervention effectiveness since PICO was the dominant format used to structure literature searches.

The supporting studies set out alternative structures to search strategy design, such as SPICE and SPIDER for structuring qualitative research questions and searches. The majority of the studies which explore alternative search strategy designs (4/5) have been published since the majority of the handbooks were last published.

The studies identified and included in this review highlight that other approaches to structuring the design of a literature search strategy exist beyond PICO. This is potentially helpful, and potentially a helpful separation and useful theoretical grounding, since it not only acknowledges the broad suitability of PICO in searching for systematic reviews of intervention effectiveness, but it also acknowledges that different types of systematic review require different approaches to literature search structure, since they rely on different types of studies and study data. This potentially questions the structure and approach of the conventional approach, as I set it out in Figure 1, since it acknowledges that different types of systematic review have different information and data needs.

If the purpose of alternative search structures to PICO is to identify different types of study and study data, it is logical that, just as the approach to search strategy design should change, so should the approach to literature searching. There is support for this idea in the studies reporting theoretical models of information retrieval. Belkin (131), Wilson (133) and Taylor (130) suggest that information retrieval should match the information needs of the search enquiry. In this way, the conventional approach, with its emphasis on PICO and bibliographic databases, may not meet the needs of researchers who use alternative search strategy designs and seek to identify different types of studies and data in different ways. This raises broader questions on the suitability of the conventional approach. The issue of identifying different types of study and study data is likely more complicated than simply adjusting the design of

the search strategy in a bibliographic database, it calls for new approaches to literature searching.

Using these approaches

There is insufficient guidance on how PICO can be used to helpfully and logically categorise research questions into literature search strategies and also how to set out which elements of the research question will form the basis of the literature search strategy (219). It is also unclear who decides how to translate the research question into a literature search strategy and also how this should be done (162).

The information professional commonly, but perhaps informally, decides which elements of the research question's PICO structure will be searched (162). This decision is often made in context of the number of studies retrieved by various combinations of searches. For example, a search strategy focused on Population AND Intervention AND Outcome might be too specific, missing studies that do not report outcomes in a homogenous way (219), but the volume of studies to screen using only Population and Intervention might be sufficiently low that the risk of missing studies by focusing on outcomes is removed. The information professional would make this decision on which option (PIO v PI) through scoping Key Stage Three.

Craven and Levay note that it is not common to involve or explain how the information professional designs a search strategy based on a research question and there should perhaps be greater clarity on the decision making here, since it might impact the transparency of decision making and the process of literature searching (162). Craven and Levay propose the use of search narratives as a method to communicate the decision-making behind the search and the development of a search strategy but they stopped short of explaining how a narrative should be designed and what detail it should offer (162). This is an idea that I have explored elsewhere by offering a worked example of a search narrative that seeks to explain the conceptual (the purpose and aims of a literature search) and contextual detail (any technical structure within the search strategy) of a literature search strategy (163). This extension would seem to be a logical next step in improving the transparency of decision-making.

Bibliographic v Supplementary

The handbooks and supporting studies ground search strategy design in the context of searching bibliographic databases. So, whilst it is acknowledged in the literature that different approaches to literature search strategy design are possible, and potentially desirable, the context of their use is indicated predominantly for use in bibliographic databases. This underlines the dominance of guidance and practice in bibliographic database searching in systematic reviews (see Key Stage Five, below). There is no further guidance (handbooks) or evaluation (studies) on how to adopt an alternative search strategy design into supplementary search methods and beyond bibliographic database searching. This is a notable limitation and further research is indicated (see below) as researchers start to navigate beyond the use of bibliographic databases to identify studies (8).

Evaluation

There is also a need for further evaluation of alternative search strategy design and alternative approaches to literature searching since it is not clear that alternative search strategy designs work. Methley et al. compared the number of studies identified and the sensitivity and specificity of a qualitative literature search using PICO, PICOS and SPIDER. They conclude that SPIDER has the potential for greater specificity but they could not recommended its use due to the risk of missing relevant studies (215). Whilst a valid conclusion, based on reliable report of methods, it is a conclusion that should be viewed on its limitations too, seeing as it is based on one case study, and it is not clear if comprehensive literature searching is the aim of literature searching in qualitative evidence synthesis, so missing studies may not be an appropriate metric to evaluate the success of SPIDER (112, 114). The work presented in Chapter 8, where I use CERQual to determine the reliability of the studies and their likely contribution to the qualitative synthesis, would be a way to re-analyse Methley et al's study, moving it on from missing studies as an outcome of interest to question does it matter that studies were missed. This would develop Methely's work, and my own (as presented in Chapter 8), and it would better situate the outcomes in the emerging discussion on the need for comprehensive literature searches, since this is presently unclear.

Further evaluation of alternate approaches to literature searching is indicated if researchers are to manage calls for greater efficiency with no loss of rigour. I suggest that researchers should be looking further than simply changing the database search

strategy design towards altering the entire process of literature searching. I will set this proposal out based on the refinements that I identify here and in the next chapter.

Key Stage Four: a summary of findings Guidance on PICO and PICOs dominate guidance documents but alternative search strategy designs are starting to emerge. This is potentially a helpful distinction since it begins to separate approaches for literature searching in systematic reviews of intervention effectiveness from other types of systematic reviews and for other types of studies or study data. It would also appear to suggest that, as the type of study or study data changes, so should the approach to literature searching.

This key stage is broadly suitable to the needs of systematic reviewers (research question two) since researchers are essentially provided with options when structuring their approach to literature searches. These structures are loosely allied to different types of systematic review. It would appear most useful for searches of bibliographic databases and this has been the focus of evaluation to date. Further research is called for on improving the transparency in decision making within this key stage and an evaluation agenda is indicated to further explore the suitability of alternative search strategy designs.

Key Stage Five: Determining the process of literature searching and deciding where to search (bibliographic database searching)

The guidance for Key Stage Five

Table two summarises **the process** of literature searching as reported in each guidance document. Searching bibliographic databases was consistently reported as the 'first step' to literature searching in all nine guidance documents.

Three guidance documents reported specific guidance on **where to search**, in each case specific to the type of review their guidance informed, and as a minimum requirement (25, 116, 120). Seven of the guidance documents suggest that the selection of bibliographic databases depends on the topic of review (3, 81, 115, 117-120), with two documents noting the absence of an agreed standard on what constitutes an acceptable number of databases searched (3, 81).

How the published studies correspond to the guidance for Key Stage Five

The guidance documents summarise 'how to' search bibliographic databases in detail and this guidance is further contextualised above in terms of developing the search strategy. The guidance documents provide guidance on selecting bibliographic databases, in some cases stating acceptable minima (i.e. The Cochrane Handbook states Cochrane CENTRAL, MEDLINE and Embase), and in other cases simply listing bibliographic databases available to search. Studies have explored the value in searching specific bibliographic databases; Wright et al. note the contribution of CINAHL in identifying qualitative studies (107); Beckles et al. question the contribution of CINAHL for identifying clinical studies for guideline development (98), and; Cooper et al. explored the role of UK-focused bibliographic databases to identify UK-relevant studies (78). Younger and Boddy examined the impact of the database host (e.g. OVID or ProQuest) on search returns, finding differing search returns from the same database (AMED) but where the 'host' was different (109).

When to database search is another question posed in the literature. Beyer et al. report that databases can be prioritised for literature searching which, whilst not addressing the question of which databases to search, may at least bring clarity as to which databases to search first (99). Paradoxically, this links to studies that suggest PubMed should be searched in addition to MEDLINE (OVID interface) since this improves the currency of systematic reviews (220, 221). In Chapter 8 of this thesis, I have tested the idea of database searching not as a primary search method (as suggested in the guidance) but as a supplementary search method in order to manage the volume of studies identified for an environmental effectiveness systematic review. I compared the effectiveness of database searching versus a protocol using supplementary search methods and found that the latter identified more relevant studies for review than searching bibliographic databases (77).

Determining the suitability: A critique of Key Stage Five

Bibliographic database searching was consistently identified as the primary search method in the guidance reviewed. There was no explicit acknowledgement of this primacy in the studies, where discussion explored the use of databases, in particular how to 'manage' the volume of studies returned through the use of literature search filters and an emerging topic, the optimal number of databases a researcher needs to

search in a systematic review. This primacy was presented as established but it is not clear why or how databases have gained this status.

The average number of bibliographic database searched in systematic reviews has risen in the period 1994-2014 (from 1 to 4) (222) but there remains (as attested to by the guidance) no consensus on what constitutes an acceptable number of databases searched (172). This is perhaps because thinking about *the number* of databases searched is the wrong question, researchers should be focused on which databases were searched and why, and which databases were not searched and why. The discussion should re-orientate to the differential value of sources but researchers need to think about how to report this in studies to allow findings to be generalised. Bethel has proposed 'search summaries', completed by the literature searcher after literature searching is completed, to record where included studies were identified, whether from database (and which databases specifically) or supplementary search methods (188). Search summaries document both yield and accuracy of searches (see Chapter 6 for a discussion of the ways in which these concepts are presently used to evaluate literature searches), which could prospectively inform resource use and decisions to search or not to search specific databases in topic areas. The prospective use of such data presupposes, however, that past searches are a potential predictor of future search performance (i.e. that each topic is to be considered representative and not unique). In offering a body of practice, this data would be of greater practicable use than current studies which are considered as little more than individual case studies (35, 223-230).

I explored the development of bibliographic databases for the purposes of literature searching, and the tensions and issues of bibliographic database searching between intervention effectiveness and other types of systematic reviews, in Chapter 1 of this thesis. Below, I briefly revisit the salient parts of that literature in order to address the second research question of this chapter with reference to this key stage (Key Stage Five).

The position of bibliographic databases in 1994 was known to be 'inadequate' (33, 35) and researchers at the UK Cochrane Centre both acknowledged this inadequacy and sought to address it as it related to Cochrane systematic reviews, and to the benefit of other types of systematic review which focus on intervention effectiveness with RCTs as the unit of analysis (24). The work of Carol Lefebvre and others to incorporate indexing terms for studies reporting randomised and clinical trials, the

MEDLINE re-tagging project, and the creation of CENTRAL, should be positively acknowledged. It would be impossible to estimate the amount of researcher time saved due to their work and the work of the hand searchers in the MEDLINE re-tagging project.

It is not clear what led Lefebvre and colleagues to prioritise investment in bibliographic databases, since other supplementary search methods were known to be more effective (i.e. they identified a greater proportion of known relevant studies from a gold standard set) in identifying studies at this time (47, 231, 232). It would seem logical that the supposed benefits of efficiency were behind the rationale for the work set out above, specifically the ability to search a significant number of journals in one search compared to handsearching journals one-by-one. Handsearching, as I report in Chapter 5, was equally known to be inefficient and it was acknowledged as resource intensive at this time (6, 231).

The investment made in bibliographic databases appears to have been successful for the identification of trial literature. The existence of relevant controlled indexing terms for studies reporting RCTs, and improved indexing, offered the possibility of using study design literature search filters. In a 2005 comparative case study, Royle and Waugh compared a simplified search strategy for RCTs (a search of CENTRAL and a search for Random$.af. in MEDLINE and Embase) to more exhaustive search strategies undertaken in a sample of Cochrane reviews (55). Royle and Waugh concluded that: 'A brief RCT search strategy is now sufficient to locate RCTs for systematic reviews in most cases. Exhaustive searching is no longer cost-effective, because in effect it has already been done for CENTRAL' (55). This study, and its findings, seems to confirm the success of the investments made by Lefebvre and colleagues and The UK Cochrane Centre.

It is worth noting, however, that the Royle and Waugh study could have been better caveated. More specific attention should be focused on the fact that the brief RCT search filter is not quite comprehensive and it is possible to miss unpublished studies. The work of Jefferson and Doshi demonstrate how important this can be and the risks of focusing solely on published studies of RCTs (66). I have also led a review of the Royle and Waugh filter (and the Cochrane HSSS) finding that both search filters miss studies that identify by the RCT phase (i.e. 'this was a phase 3 trial') and not by study design (per CONSORT)). The work was presented at HTAi in 2016 and is currently pending submission (233).

It should be repeated that the investments made in indexing terms and 're-tagging' were undertaken exclusively focusing on studies reporting RCTs. This best serves reviews of clinical intervention effectiveness which prioritise this study design. In addressing research question two, this suggests that the conventional approach best meets the needs of this type of systematic review.

This suggests the conventional approach is not quite suitable for use in other types of systematic review since the investment in and therefore role of bibliographic databases is not as clearly established. As I set out in Chapter 1 of this thesis, no other field of study or method of research has benefited from the investment of resources and time as intervention effectiveness. Whilst researchers have called for study design indexing terms for qualitative studies (Dixon-Woods in 2001, the term followed in 2003 but it has not been retrospectively applied) (234), or clearer reporting of study design or method of evaluation (e.g. (235)), bibliographic database searching still lacks the efficiencies of searching for RCTs. It is not clear what, without a substantial investment of resource, is the solution. It would seem unlikely that the National Institute of Health Research will invest given that they have since ceased funding on bibliographic resources such as DARE and the HTA library (203) and other bibliographic databases with topic specific indexing have also closed (HEED and NHS EEDs).

There may be alternative models of resolving the issue of indexing and retrospectively indexing likely relevant studies. This would seem useful for the identification of qualitative studies, studies reporting prognostic or diagnostic data, and more broadly public health interventions. Anna Noel-Storr has led The Cochrane Crowd in the Embase alternative to the MEDLINE re-tagging project (236). This manual work has been extended into machine learning with some success (237). A limitation of the machine learning study is that it requires researchers to follow CONSORT reporting for the machine to identity likely relevant studies (47). As the results of the MEDLINE re-tagging project indicate (125,000 studies were indexed that had been missed), study authors then (1996) and in 2016 are still not following CONSORT (233) which will impact upon the success of the EMBASE crowd project. This remains a problem for studies using effectiveness designs and it will certainly be a problem for studies reporting qualitative data where there is no alternative version of CONSORT and no hierarchy of study designs. There are also informal reports of

the crowd-sourcing approach increasing the number of false-negative hits in CENTRAL, an issue which is being investigated by Cochrane (2018).

Another alternative is in exploring alternative approaches to literature searching to mediate the poor indexing for studies and study data beyond controlled trials. Different structures to literature search design (Key Stage Four), or ideas such as I explore in this thesis, for alternative approaches to literature searching, may be another solution. This would seem to shift the focus away from the resources searched and on to the person searching.

Key Stage Five: a summary of findings A key finding of this literature review is that nine different guidance documents report the same process of literature searching. All nine of the guidance documents also appear to situate bibliographic database searching as the primary search method in systematic reviews. The rationale for this is not clear but I suggest that efficiency was potentially a motivating factor with benefits for transparency and reproducibility too.

It is not clear how the primacy of bibliographic database searching evolved but it would appear best suited to systematic reviews of intervention effectiveness. It is not established, nor is it clear, that bibliographic database searching works as well in other types of systematic review. I identify this gap in knowledge and I seek to explore it a refinement to this key stage in Chapter 4 (see 4.4.10) and I evaluate the suitability of this finding in Chapter 8.

The primacy of bibliographic database would appear to be supposed rather than established in effectiveness case-studies and this should be further explored, especially in relation to supplementary search methods.

Table 6 The order of literature search methods as presented in the guidance documents

Step	The CRD Handbook	The Cochrane Handbook	Collaboration for environmental evidence	Joanna Briggs Institute reviewers manual	IQWiG Methods Resources	Systematic reviews in the social sciences: a practical guide	Eunethta[a]	Campbell Handbook	Developing NICE guidelines: the manual
1	searching electronic databases	searching bibliographic databases	searching online literature databases and catalogues	databases (development of search strategies, phase one)	bibliographic databases (1. for primary literature. 2. for SRs)	databases	bibliographic databases	bibliographic databases (1. subject databases. 2. general databases)	No list of search methods but guidance distinguishes between database searching (first) and supplementary searching (second)
2	scanning references lists of relevant studies	handsearching	searching websites of organisations and professional networks	database searching (phase two)	search in trial registries	grey literature	study registries	conference proceedings and meeting abstracts	
3	handsearching of key journals	conference abstracts or proceedings	searching the world-wide web	review reference lists	clinical practice guideline databases and providers	identifying on-going research	searching for unpublished company documents	existing review and publication reference lists	
4	searching trials registers	other reviews	searching bibliographies of key articles/ reviews	handsearching	requests to manufacturers	theses	regulatory documents	web searching	
5	contacting experts and manufactures	web-searching	contacting key individuals who work in the area		other data sources	conference proceedings	queries to authors	unpublished studies	
6	searching relevant internet resources	unpublished and on-going studies (inc. author contact)	citation searches for key papers/ included papers			citation searching	further search techniques	on-going studies	
7	citation searching					searching the web		institutional repositories	
8	using a project website to canvas for studies					contact with experts		handsearching	
9						trials registers			

Key Stage Six: Determining the process of literature searching and deciding where to search (supplementary search methods)

The guidance for Key Stage Six

Table 6 also summarises the process of literature searching which follows bibliographic database searching. This has been directly extracted as reported in the guidance documents. As Table 6 sets out, guidance that supplementary literature search methods should be used in systematic reviews recurs across documents, but the order in which these methods are used, and the extent to which their use is indicated or implied, varies. Inconsistencies were also noted in the labelling of supplementary search methods between guidance documents.

How the published studies correspond to the guidance for Key Stage Six

Rather than focus on how to use the methods, which I review in Chapter 5 of this thesis (and which has also been published (6)), I focus on the aim or purpose of supplementary search methods as reported in the guidance.

The Cochrane Handbook reported that 'efforts' to identify unpublished studies should be made (25). The implication of this guidance, specifically, is that bibliographic databases identify published studies and supplementary search methods are indicated to identify unpublished studies or studies not indexed in bibliographic databases. This is a helpful distinction since it articulates a distinction in the role and purpose of supplementary search methods. This can be nested in the guidance of four other guidance documents (3, 25, 81, 117), that searching beyond bibliographic databases is necessary since 'databases are not the only source of literature' (81).

Only the Institute for Quality and Efficiency in Healthcare (IQWIG) guidance reported on when to use supplementary methods. The IQWiG handbook reported that the use of handsearching (in their example) could be determined on a 'case-by-case basis' which implies that the use of these methods is optional rather than mandatory (117). This is in contrast to the guidance on bibliographic database searching (Key Stage 5) which is indicated as mandatory.

The key issue for supplementary search methods is similar in many ways to the issue of searching bibliographic databases: demonstrating value. The purpose and contribution of supplementary search methods in systematic reviews is increasingly acknowledged (6, 79, 80, 95, 182, 238-241) but understanding the value of these search methods to identify studies and data is unclear, not least as the studies on the

various supplementary methods have not been considered in one review before. I identified this gap in knowledge and I set out the review Chapter 5.

In Chapter 5, I review the literature on supplementary search methods looking to determine the advantages, disadvantages and resource implications of using supplementary search methods (6). I summarise the key guidance and empirical studies and seek to address the question on when to use these search methods and when not to (6). The guidance reviewed is limited in this regard and, as Table 6 demonstrates, offers conflicting advice on the order of searching, and the extent to which these search methods should be used in systematic reviews.

Determining the suitability: a critique of Key Stage Six

Supplementary search methods make contributions to, and are an important part of, systematic literature searching in systematic reviews. This finding has been consistently reported since approximately 1993 (242), with occasional studies updating this finding (e.g. (95, 111, 182)), either focusing specifically on individual supplementary search methods (e.g. handsearching (231) citation chasing (79) and contacting study authors (240)), or more generally identifying support for the use of supplementary search methods when compared to bibliographic database searching.

The most frequently cited study on the use of supplementary search methods to identify studies in systematic reviews (n=742 citations June 2018) is a 2005 case study by Greenhalgh and Peacock (111). Greenhalgh and Peacock published a case study which audited how 495 primary sources were identified for a systematic review, finding that 30% of included studies were identified by a protocol-defined search (this result is reported in the abstract as the 'headline' figure to demonstrate superiority of supplementary search methods), and 25% were identified by searches of electronic databases. The remainder of studies (45%) were identified by supplementary search methods (111).

Greenhalgh and Peacock's study is commonly referenced as a rationale for using supplementary search methods. There are a number of limitations with their study, which I consider below, and which can be used to highlight and inform the critique of the suitability of supplementary search methods in the conventional approach and systematic reviews more generally.

First, and this is not clearly stated or even acknowledged within the study, the findings reported are based on a single case study. The language used to report the

study generalises the findings of the case study in question, to apply to the process of literature searching in all future complex systematic reviews. This risks over-stating the importance of the case study. No attempt is made by the authors to situate the findings of their case study in the context of their topic or other research. This is important given the subsequent criticisms I make below;

Secondly, it is not clear to what extent their findings were influenced by confounding factors of the topic, the resources available, access to relevant or optimal resources, and the abilities of the review team to identify relevant studies. It is stated that a specialist librarian was used for bibliographic searching but no explanation is offered as to what 'a specialist librarian' means nor is an indication of the experience of the librarian given to allow researchers to determine for themselves the relevance of this claim. The specialist librarian is notably absent from the list of study authors and it is also unclear who undertook the supplementary search methods on which this study rests its findings and conclusions;

Thirdly, in an attempt to demonstrate superiority of an approach to literature searching relying on non-protocol-driven supplementary search approaches when compared to a protocol-driven literature search approaches, data on the time taken by either approach is imprecise when provided (they say 'developing and refining search strategies and adapting these to different databases, took about two weeks' (111)) or otherwise scant where it is even reported (they say 'twenty three per cent of the sources were known to us or were recommended by colleagues when we approached them by email, which took little time.' (111)). Without a clearer and explicit report of time taken on a search method-by-method basis, in a more commonly presented format, the claim that any one method or approach was superior to the other is not well supported or clear;

Fourthly, no attempt is made in the study to distinguish between studies that were identified uniquely or were available but were missed due to error from other search methods. It is not clear to what extent the studies identified by hand-searching, for example, where actually available for retrieval from the other search methods used in the protocol driven search approach but were missed. In the absence of this data, the true effect of any approach or search method is unclear. It is possible that their bibliographic database searching was deficient which led to the corresponding success of handsearching, for instance; and

Fifthly, whilst a greater number of relevant studies were identified in the non-protocol-driven approach, the value of these studies, and therefore a more accurate summary of the effectiveness of the non-protocol-driven approach, is not explored. The prevailing conclusion is that superiority is demonstrated by the approach which identified the greatest number of potentially relevant studies missed by the alternative approach and this conclusion is questionable (see Chapters 4, 6, 7, 8 for a discussion of this idea of value).

These potential limitations are important to understand given how influential this study has been since publication. Other case studies examining the role of supplementary search methods suffer from similar methodological deficiencies and, as I find in Chapter 5, this means that the evidence base available to researchers on supplementary search methods is uneven and potentially flawed. Based on the critique above, I summarise these issues below.

There is no existing review of supplementary search methods. No attempt has been made to bring together the primary studies and also review guidance. This is a gap in knowledge that I identified and that I address in Chapter 5 by undertaking the first review of supplementary search methods.

This criticism corresponds to the critique of Greenhalgh and Peacock, and the evidence-base for supplementary searchers more generally, because it reflects the fact that the evidence-base for the effectiveness of supplementary search methods is principally made up of individual case studies. These case studies do not appear to cite each other or integrate the findings of one study into the next. Scholarship requires building upon research already conducted and this does not appear to happen with the studies evaluating supplementary search methods. This prohibits comparison and knowledge building. Greenhalgh and Peacock is a good example of this. There are, however, further problems with the evidence-base for supplementary search methods.

The age of the studies
Of 12 studies that evaluate handsearching in Chapter 5, 10 were published before 2005. The only study that provided data on costs was published in 1995. For researchers to make informed decisions about handsearching as a search method, for example, data needs to be kept broadly up-to-date and reflect advances in technology. This is variously true of the other search methods reviewed too. The

studies and study data are, in most cases, out of date and it not clear how the findings have kept ahead of technological advancements.

The outcomes recorded

Researchers need to agree upon, and consistently report, key outcome measures. This would permit generalisability of outcomes and demonstrate clearer support for the use of supplementary search methods. The same is true of other search methods.

This challenge is a key objective of this thesis and it forms part of the recommendations in my review of supplementary search methods. Namely, within systematic reviews, to record the time taken on a search method-by-search-method basis (to include the time taken to search bibliographic databases) and a call to acknowledge the contribution of the studies identified in literature searching, reconciling the method of literature searching to any impact made in the synthesis of studies. This would articulate a clear understanding of the contribution of individual search methods which, if time taken to search is included in analysis, permits a preliminary form of cost-effectiveness analysis too.

The ubiquitous conclusion that supplementary search methods identified studies missed by bibliographic databases does not advance learning, it merely reflects guidance. Explaining which studies were uniquely identified by supplementary search methods, and the effect these studies had on the synthesis of studies (qualitative) or estimate of effectiveness (quantitative), more clearly demonstrates the impact of supplementary search methods. This form of proof represents a gap in knowledge which I articulate as 'value' in Chapter 6, where I set out alternatives strategies to measure the effectiveness of literature search approaches, and I aim to explore in Chapter 7 and 8. Furthermore, studies should be clearer about identifying studies uniquely and by which search methods. Greenhalgh and Peacock do not summarise this and so it remains unclear from reading their study if the studies identified by supplementary search methods were identified uniquely or were missed by their bibliographic database searches.

Naming?

This critique echoes with calls for the adjective 'supplementary' to be dropped in reference to non-database methods of literature searching in systematic reviews. The reference to 'supplementary' has arguably always been an inadequate term of

reference. Where the purpose of literature searching is for comprehensive literature searching, and studies are missed by supplementary searches, the adjective 'supplementary' can equally apply to databases. It could be time to move away from 'supplementary' to simply a database and non-database classification for search methods.

My concern with the moniker 'supplementary' is that it implies 'optional' or 'if we have time.' Guidance documents could make clearer recommendations about the expectations on researchers to engage with supplementary search methods and I identified significant variability on the guidance available between handbooks in the literature review. Authors of guidance documents, conversely, would benefit from better and up-to-date evidence in order to make guidance recommendations.

Key Stage Six: a summary of findings: The need for supplementary search methods is clear from the guidance documents and the studies identified for this review. Whilst the evidence is limited, and out of date in many cases, a case for the benefit of supplementary search methods is clear.

In terms of research question 2 (see 3.4.1), and the suitability of the conventional approach in relation to this key stage, the use of supplementary search methods would appear to be suitable. How the search methods work, when to use them, and the resource implications is less certain. The literature review I present in Chapter 5, and which has since been published (6), develops this idea. Further research is indicated to update the case study data available but great thought needs to go into what outcomes to measure and it is important to nest any new studies in the findings of previous studies.

I repeat the refinement identified in the previous key stage, the idea that the information needs should decide the approach to literature searching. It would seem to apply in this key stage too. Researchers need to develop their understanding if supplementary search methods and how best to use them to navigate literature searching beyond bibliographic databases (8).

Key Stage Seven: Managing the references

The guidance for Key Stage Seven

Five of the guidance documents provided guidance on managing references in a systematic review. Downloading, de-duplicating and managing the output of literature

searches were covered (3, 81, 115, 119). Guidance typically itemised available bibliographic management tools rather than offering guidance on how to use them specifically (3, 81, 120). The CEE handbook provided guidance on importing data where no direct export option is available (e.g. web-searching) (115).

How the published studies correspond to the guidance for Key Stage Seven

The literature on using bibliographic management tools is not large relative to the number of 'how to' videos on platforms such as YouTube (see for example (243)). These YouTube videos confirm the overall lack of 'how to' guidance identified in this study and offer useful instruction on managing references. Bramer et al. set out methods for de-duplicating data and reviewing references in Endnote (244, 245) and Gall tests the direct search function within Endnote to access databases such as PubMed, finding a number of limitations (246). Coar et al. and Ahmed et al. consider the role of the free-source tool, Zotero (247, 248).

Determining the suitability: a critique of Key Stage Seven

Five of the nine guidance documents provided guidance on managing study references generated through literature searching and through the process of a systematic review. Studies provided little further guidance on this topic but support is available on-line through YouTube videos. The importance of the practical and 'how to guides' should not be overlooked since they offer straight-forward guidance, often with videos to show researchers step-by-step how to accomplish a task.

Managing references is a key administrative function in the process of a systematic review, particularly for documenting searches that conform to PRISMA reporting guidance (249). This has been identified as a particular skill of the information professional (143). It is not clear why this should anymore be a skill of an information professional than any other researcher working on a review since the identification and citation of studies forms a key component of academic writing, it is something that all researchers should be competent at doing (250, 251).

Irrespective of who manages the references in a systematic review, the importance of this key stage should not be overlooked. It is vital to complete the PRISMA flow-chart, and so it is a critical part of undertaking and completing a systematic review (196). The competent recording of the number of studies as they progress through the process of review, represents a valid record of the systematic process of a systematic review (249). Furthermore, a transparent report of this process offers the

possibility to critically appraise and replicate the methods and process of review (249). This links to the next key stage I identified, documenting the search (Key Stage Eight).

Little attention is given to managing bibliographic references in the guidance and studies identified in this literature review (252). Managing references, both in terms of retrieving references from bibliographic databases and from supplementary search methods, and managing references as studies are simultaneously included or excluded in the process of review, is not straightforward and it can be a concerning process (253).

Approximately 25 bibliographic management tools exist (253, 254), and other products for screening and managing the process of review are common, but no one tool or product is perfect or commonly used from start to completion of a systematic review. As an example, in my last systematic review, records were de-duplicated in Endnote (255), screened in Rayyan (256), data extracted into Excel, and I undertook the meta-analysis in Rev Man (257). Every step where data is moved between resources only increases the chance of error, for studies not to export cleanly, and the possibility that some data might be lost or inadvertently altered due to user error.

An evaluative study that compared the available software or tools would be a valuable resource for the research community (258), perhaps accounting for the fact that some tools are free (259) and others cost money (253). Such a review would also help develop the guidance where the selection of bibliographic management tools could be more clearly recommended against key criteria, perhaps developed from the needs of researchers undertaking systematic reviews very specifically (249). This review would need to account for the fact that the bibliographic tools are subject to regular technological updates and hardware compatibility issues (254, 260) and that training and support for users is also indicated (261). This review would contribute to knowledge building since it would highlight that no one tool does all of the tasks needed in a systematic review.

Key Stage Seven: a summary of findings This key stage is important to managing the process of systematic review. It fits the conventional approach and the needs of any other type of systematic review. In terms of addressing research question 2 of this chapter, this key stage is broadly important to all types of systematic review.

Key Stage Eight: Documenting the search

The guidance for Key Stage Eight

The Cochrane Handbook was the only guidance document to indicate a specific reporting guideline: Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (25). Six guidance documents provided guidance on reporting the process of literature searching with specific criteria to report (3, 115, 117, 119, 120). There was consensus on reporting: the databases searched (and the host searched by), the search strategies used, and any use of limits (e.g. date, language, search filters (The CRD handbook called for these limits to be justified (3))). Three guidance documents reported that the number of studies identified should be recorded (3, 115, 117). The number of duplicates identified (115), the screening decisions (117), a comprehensive list of grey literature sources searched (and detail for other supplementary search methods) (119), and an annotation of search terms tested but not used (120) were identified as unique items in four documents.

The Cochrane Handbook was the only guidance document to note that the full search strategies for each database should be included in the appendix of the review (25).

How the published studies correspond to the guidance for Key Stage Eight

All guidance documents should ultimately deliver completed systematic reviews that fulfil the requirements of the PRISMA reporting guidelines (196). The guidance identified in this literature review broadly requires the reporting of data that corresponds with the requirements of the PRISMA reporting guidance although documents typically ask for diverse and additional items (196). In 2008, Sampson et al. observed a lack of consensus on reporting search methods in systematic reviews (262) and this remains the case in 2018, as evidenced in the guidance documents, and in spite of the publication of the PRISMA reporting guidelines in 2009 (263). It is unclear why the collective guidance does not more explicitly endorse adherence to the PRISMA reporting guidance.

Reporting of literature searching is a key area in systematic reviews since it sets out clearly what was done and how the conclusions of the review can be believed (178, 263). Despite strong endorsement in the guidance documents, specifically supported in PRISMA reporting guidance, and other related reporting standards too (such as ENTREQ for qualitative evidence synthesis, STROBE for reviews of observational

studies), authors still highlight the prevalence of poor standards of literature search reporting (157, 191, 264-270). To explore issues experienced by authors in reporting literature searches, and to look at the uptake of PRISMA reporting guidance, Rader et al. surveyed over 260 review authors and librarians to determine common problems with reporting literature searching (195). Atkinson et al. have also analysed reporting standards for literature searching, summarising recommendations and gaps for reporting search strategies (194). The work of Rader et al. and Atkinson et al. can usefully be read together since Rader et al. identify issues with reporting and Atkinson et al. present a set of focused and detailed standards for reporting the methods used, many of which highlight the issues identified by Rader et al. Common problems were reported as: lack of time, access to suitable tools and no clear understanding of the reporting requirements (195).

One area that is less well covered by the guidance, but nevertheless appears in this literature, is the quality appraisal or peer review of literature search strategies. The PRESS checklist is the most prominent and it aims to develop evidence-based guidelines to peer review of electronic search strategies (122, 183). A corresponding guideline for documentation of supplementary search methods does not yet exist and would be a useful contribution to research.

Determining the suitability: a critique of Key Stage Eight

Whilst the literature review was conceived to test the hypothesis that leading guidance documents promoted a similar methodological process to literature searching, in the key stage of documenting a literature search, there were profound differences in the guidance reported and variability between guidance documents. The problem that this finding creates becomes clearer when the guidance is contextualised with the studies, as I set out above.

The key criteria to be reported when documenting a literature search in a systematic review varied widely between the guidance documents and agreement between guidance documents was limited. Only the Cochrane Handbook reported reference to a particular reporting guideline, PRISMA (25). Reporting and documenting the approach to study identification and literature searching is key for transparency and replicability (194).

The published studies explored common problems with reporting the process of documenting literature searching. A systematic review by Sampson et al. concluded

that there was no consensus on search reporting methods in systematic reviews (262) and the study by Rader et al. discussed *passim*, surveyed over 260 review authors to determine common problems when documenting a completed literature search (195).

The studies demonstrate that the lack of guidance in the guidance documents – and in guidance checklists – is problematic. Sampson et. al undertook a systematic review in 2008 to identify validated or evaluated search reporting instruments used in systematic reviews and to compare reported and recommend searching practices (262). Their systematic review concluded that there was no consensus regarding optimal reporting of systematic review search methods (262). The finding of Sampson et al. correlates with a study by Rader et al. (195). Rader et al. analysed the results of the survey providing methodological and contextual guidance to support researchers in documenting literature searches that meet PRISMA reporting guidelines. Eighty-six percent of survey respondents (153/178) identified a need for further guidance, specifically on what aspects of the search process to report (195). Furthermore, the Sampson et al. systematic review reported, and Rader et al's survey results found, significant variability in the practice of documenting a literature search, with the survey results of Rader et al. demonstrating that some respondents were producing searches that would not meet PRISMA reporting guidelines (195, 262). Even where the guidance is consistent, and the reporting should be similar, Yoshi et al. found variability in the reporting of literature searching (270).

This is a frustrating problem and I consider below some solutions. It is an important problem to resolve since clear and transparent reporting has been linked to increased confidence in the review (262).

Differences between the application of literature search methods equates to a problem. The handbooks and studies treat search documentation as a topic with a singular solution and the reality is more complex. Documenting the search process for a bibliographic database is easier than reporting supplementary search methods. It is possible for a researcher to export the search strategy they run in MEDLINE and for another researcher to replicate this search and achieve the same result, wherever they are in the world. The standard search output from MEDLINE (Ovid interface) actually includes the key data a researcher would need to conform to PRISMA reporting guidance, namely: database name, data parameters, search strategy and

date searched. In this instance, it is even less clear why searches do not confirm the minimum guidance.

A study by Witkowski and Aldhouse reviewed 46 single technology appraisals submitted to NICE, concluding that transparency and replicability of supplementary search methods was low, especially since supplementary searches were not often reported (199). The findings of Rader et al. should perhaps be situated in the complexity of the methods used. Taking the supplementary search methods of web-searching as an example, researchers access similar versions of the internet but they will achieve different search returns due to algorithms, cookies, and regional differences in search returns (100, 200). Guidance documents and the published studies reviewed here do not appear to account for this variation, or adequately explain to researchers how the documentation required by guidance documents for a database differs for the documentation that would make sense when reporting web-searching, for example. The reality, as demonstrated in a case study by Briscoe, who examined the reporting of web-searching in HTA reports, is that the bibliographic database searching reported in a systematic review may well conform to PRISMA reporting guidelines and the needs of committee members of a technology appraisal committee, but the web-searching was often insufficiently reported and it could not be replicated (200).

Further research is required to address this problem. My suggestion is for search method-by-search method guidance, and individual reporting templates for each individual search method used, for completion by a researcher. This adds to the paperwork which researchers must complete, and it may lengthen the annexes and supporting material of a systematic review, but it would help articulate the clarion call, that a systematic review is a record of what was done and how it can be believed (178, 194). It is imperative that search documentation be provided and be adequate (191) and if this means multiple variations of search documentation on a method-by-method basis, it is a solution which is hard to ignore.

Peer review of literature search documentation is another solution. Cochrane systematic reviews are comprehensively peer reviewed to ensure the methods used, and which are reported, meet minimum methodological requirements (271). The Cochrane Collaboration already has its Methodological Expectations of Cochrane Intervention Reviews (MECIR) guidance (272) but this applies only to Cochrane reviews and to Cochrane authors (14). A broader approach is required for other

systematic reviews. An extension to the Peer Review of Electronic Search Strategies (PRESS) for systematic reviews (SRs) checklist, to include a review of search documentation, in addition to a review of the underlying search strategy and to include a full-account of the approach to searching and detail on supplementary search methods, would seem an obvious and also appropriate solution. Peer review of search strategies is, however, low in uptake, the reason for which is unclear (5). Peer review is also an imperfect solution as it relates to literature searching since peer review (typically) happens at the end of the review so that it is seen in context and it may be difficult to locate any missing or incomplete detail needed since literature searching (typically) happens at the start of a systematic review (163). The difference between start and completion of a review can be many months or more than a year. It would be more efficient to fix any issue in a search strategy early in a review process, so that the findings of the review are not compromised at a point where little can be done to amend the fault. Information professionals could be more diligent in peer review and ensure that the review of the searches is undertaken with the research team in the first instance and at a time when any errors can be corrected.

Peer review of literature searching more generally is under-recognised within journal submission. Editors rely on the broad competencies of peer reviewers for journal articles who may have little knowledge or practical experience of literature searching. A good example of this issue, and the effect it can have, was the publication of a systematic review and meta-analysis on the effectiveness of sutures versus staples for skin closure in orthopaedic surgery by Smith et al., published in the BMJ (273).

The literature search used for the review and meta-analysis is deficient in number of areas (use of indexing, direct comparison of interventions, limited use of truncation, and use of limits) which led to key RCTs being missed in the analysis (274). Open peer review was not introduced in the BMJ until 2015 (275), so it is not possible to know the identity of the peer reviewers as a mechanism for assessing their fluency in literature searching and broader competencies in systematic review and statistical meta-analysis, and it is equally not possible to know if the deficiencies in the literature searching were identified.

The review and meta-analysis has been subject to 19 on-line, predominantly negative, comments since publication but the publication has not been withdrawn. The on-line comments all question the findings of the study and the reliability of the

conclusions. Only one comment relates to the searching, stating that a search of PubMed identified three studies reporting RCTs missed by Smith et al. (274), otherwise no commentator links the inadequacy of the literature search to the broader failing of the review. In fact, the authors defend their literature searching, noting that they searched published and unpublished sources and the authors reported their literature search strategy (276). These points are not in dispute, but the search was bad in execution and limited in scope which caused it to miss studies and this should have been picked up in peer review by adequately trained researchers and peer review. The other comments raise concerns about the data and the findings of the review. There is a strong indication here for information professionals to be involved in the peer review of studies, even if only on the very discrete area of study identification. Determining the quality of a literature search can help to piece-together the impact of the wider piece of work (191).

Another solution is for training, aimed to support and educate researchers in accurately documenting a literature search, with the outcome of leading to a change in behaviour. The Rader et.al study was conceived with this aim in mind. The study provides detailed methodological and contextual guidance on what detail to report when documenting a literature search for a systematic review based on the findings of a survey conducted by the authors. One of the contributions of the Rader et al. study was the reporting of a table of summary elements to be reported when documenting a literature search (195). It was hoped that this would offer clearer guidance on what to report, leading to less confusion and fewer errors in search reporting (277).

Figure 2 Summary data to report

Table 1. Summary of elements to be recorded.
Database
Database platform (e.g., EBSCOhost, OVID, and PubMed)
Source (if nondatabase)
Search date
Timeframe of search including database coverage dates
Search strategy
Number of records retrieved

Arguably, it would have been more helpful to make the table reported in Figure 2 freely available in an open-source format. This equally, is a criticism of the PRESS reporting guidance too. The Rader et al. study has not been evaluated to asses if the guidance provided has improved search reporting or, indeed, if it has achieved its aim in supporting researchers. Whilst it is impossible to gauge a study's uptake on

the basis of citations alone, the authors met at the 2010 Cochrane Colloquium and in the context of a Cochrane Information Retrieval Methods (IRMG) Group meeting. The study was not adopted under aegis of the Cochrane IRMG, which would have given the work a somewhat higher profile. The reason for this is unclear.

In light of the fact that Sampson et al. published their systematic review in 2008, and Rader et al. published their survey in 2014, and studies still find that documentation of literature searching is poor, another solution might be required. This solution would acknowledge a point over-looked in the guidance: how to document supplementary search methods. Re-considering the findings of the Rader et. al survey in the context of this thesis, I believe that what to report and how to report are primary areas of current confusion that would warrant future investigation.

How the guidance on reporting the literature searching process corresponds to critical appraisal tools is an area for further research too. In the survey undertaken by Rader et al., 86% of survey respondents (153/178) identified a need for further guidance on what aspects of the literature search process to report (195). The PRISMA statement offers a brief summary of what to report but little practical guidance on how to report it (196). Critical appraisal tools for systematic reviews, such as AMSTAR 2 (278) and ROBIS (279), can usefully be read alongside PRISMA guidance, since they offer greater detail on how the reporting of the literature search will be appraised and, therefore, they offer a proxy on what to report. Further research in the form of a study which undertakes a comparison between PRISMA and quality appraisal checklists for systematic reviews would begin addressing the call, identified by Rader et. al, for further guidance on what to report (195, 280).

3.4.4   Limitations of this literature review

A potential limitation of this literature review is the focus on guidance produced in Europe (the UK specifically) and Australia. I justify the decision for the selection of the nine guidance documents reviewed in this chapter in section "Identifying guidance". In brief, these nine guidance documents were selected as the most relevant health care guidance that inform UK systematic reviewing practice, given that the UK occupies a prominent position in the science of health information retrieval. I acknowledge the existence of other guidance documents, such as those from North America (e.g. the Agency for Healthcare Research and Quality (AHRQ)

(281), The Institute of Medicine (282) and the guidance and resources produced by the Canadian Agency for Drugs and Technologies in Health (CADTH) (283)).

What is not clear is the extent to which the guidance documents inter-relate or provide guidance uniquely. It is likely that they influence and reflect each other. The Cochrane Handbook, first published in 1994, is notably a key source of reference in guidance and systematic reviews beyond Cochrane reviews. It is not clear to what extent broadening the sample of guidance handbooks to include North American handbooks, and guidance handbooks from other relevant countries too, would alter the findings of this literature review or develop further support for the process model. Since we cannot be clear, we raise this as a potential limitation of this literature review. On our initial review of a sample of North American, and other, guidance documents (before selecting the guidance documents considered in this review), however, we do not consider that the inclusion of these further handbooks would alter significantly the findings of this literature review.

A further limitation of this review was that the review of published studies is not a systematic review of the evidence for each key stage. It is possible that other relevant studies could help contribute to the exploration and development of the key stages identified in this review.

3.5   Chapter findings and conclusions

I summarise here the findings and conclusions of my critical literature review against the chapter's research questions. The research questions of this chapter were:

1. Can a shared model of the literature searching process be detected across guidance documents and, if so, how is this process reported and supported?

2. If a shared approach to literature searching can be identified, is it suitable to meet the needs of researchers undertaking systematic reviews?

Research question 1. Can a shared model of the literature searching process be detected across guidance documents and, if so, how is this process reported and supported?

This literature review demonstrates the existence of a shared model of the literature searching process in systematic reviews. I call this model 'the conventional approach', since it appears to be common convention in nine different guidance documents.

The findings reveal eight key stages in the process of literature searching for systematic reviews. These key stages are consistently reported in the nine guidance documents which suggests consensus on the key stages of literature searching, and therefore the process of literature searching as a whole, in systematic reviews.

In Table 6, I demonstrate consensus regarding the application of literature search methods. All guidance documents distinguish between primary and supplementary search methods. Bibliographic database searching is consistently the first method of literature searching referenced in each guidance document. Whilst the guidance uniformly supports the use of supplementary search methods, there is little evidence for a consistent process with diverse guidance across documents. This may reflect differences in the core focus across each document, linked to differences in identifying effectiveness studies or qualitative studies, for instance.

Eight of the nine guidance documents reported on the aims of literature searching. The shared understanding was that literature searching should be thorough and comprehensive in its aim and that this process should be reported transparently so that that it could be reproduced. Whilst only three documents explicitly link this understanding to minimising bias, it is clear that comprehensive literature searching is implicitly linked to 'not missing relevant studies' which is approximately the same point.

The critical literature review describes and critiques an approach to literature searching in systematic reviews. Whilst it is clear that a shared approach to literature searching exists within guidance documents, it is not clear if this approach meets the needs of systematic reviewers generally, and for differing types of systematic review specifically. As I have explored through critique, it would seem that the conventional approach was designed for, and it best meets the needs of, researchers undertaking intervention effectiveness systematic reviews. It is not clear if, or how, the conventional approach is interpreted by researchers, or if it is followed in practice, or if the process identified in this critical review is still a valid approach. These points and the conventional approach would benefit from further testing and research.

Identifying the process, and defining the key stages in this review, will help future testing and evaluation. The explicit identification of the conventional approach helps categorise the scholarship available, and it prioritises areas for development or further study. The supporting studies on preparing for literature searching (key stage three, 'preparation') were, for example, comparatively few, and yet this key stage

represents a decisive moment in literature searching for systematic reviews. It is where search strategy structure is determined, search terms are chosen or discarded, and the resources to be searched are selected (208, 284). Information professionals, librarians and researchers, are well placed to develop these and other areas within the key stages identified. It is my experience that the knowledge needed for the 'preparation' key stage is tacit within the information professional and researcher community, gained through experience of undertaking different types of review, and the shared experience of the research team in anticipating issues or opportunities (163, 208, 209). This tacit knowledge is rarely made explicit or reported in detail, making this an opaque process in the description of methods used to undertake a systematic review.

Research question 2. If a shared approach to literature searching can be identified, is it suitable to meet the needs of researchers undertaking systematic reviews?

This review calls for further research to determine the suitability of using the conventional approach. The publication dates of the guidance documents which underpin the conventional approach may raise questions as to whether the process which they each report remains valid for systematic literature searching. The nature of evidence synthesis and systematic reviews is one of evolution as methods are developed and refined. It may be appropriate to test the conventional approach, to determine if it is followed in practice and, if it is, if it suits all styles of systematic review, or if alternative methods of literature searching would be valuable. The findings of this critique suggest it is not clear if it is suitable or desirable to use the same process model of literature searching for qualitative evidence synthesis, for example, as for reviews of intervention effectiveness. It is also possible that the conventional approach, as defined in this chapter, may not now meet the needs of intervention effectiveness reviews. These ideas would benefit from testing, which I will start to do through two case studies presented later in the thesis (see Chapters 7 and 8).

Whilst at the top-level, the conventional approach would appear to meet the needs of researchers, it becomes clearer by critiquing the individual key stages that some key stages within the model would benefit from refinement or development.

Some key stages of the conventional approach meet the needs of researchers undertaking all types of systematic reviews whilst other stages are dependent on the

type of review undertaken. I have categorised each key stage identified in Figure 1 and reported above as **agreed:** meets the needs of all types of systematic reviews, **partially agreed:** meets the needs of some systematic reviews and **unclear:** unclear if it meets the needs of systematic reviews.

The top-level process model was categorised as agreed since it superficially appears to support the needs of researchers undertaking systematic reviews. Within the eight key stages in the conventional approach, I categorised:

Three key stages where the key stage appears suitable to meet the needs of researchers undertaking all types of literature searching for systematic reviews. These were; Key Stage Four: search strategy, Key Stage Seven: managing references and Key Stage Eight – reporting the search process. Therefore, I **agree**: that these key stages appear to be suitable to meet the needs of researchers undertaking systematic reviews;

Three key stages in which it was partially agreed that the key stage as I identified it are suitable to meet the needs of researchers undertaking systematic reviews. These were; Key Stage One: who should literature search; Key Stage Five: bibliographic database searching and Key Stage Six: supplementary searching. Therefore, I **partially agree**: these key stages might be suitable to meet the needs of researchers undertaking some but not all systematic reviews it is likely that refinements and/or further research is necessary and this is indicated in Table 7.

Two key stages in which it was unclear if the key stage as I identified it are suitable to meet the needs of researchers undertaking systematic reviews. These were Key Stage Two: aims and purpose of literature searching and Key Stage Three: preparation. Therefore, it is **unclear** if this key stage is suitable to meet the needs of researchers undertaking systematic reviews. It is likely that refinements and/or further research is necessary and this is indicated in Table 7.

Table 7 Findings of research question two: suitability, refinements and further work indicated

Key Stage	Suitable?	Notes	Suggested refinements or further work indicated
Overall model (Figure 1)	Agreed: meets the needs of researchers	The conventional approach represented as a process model would appear to be aligned and suitable to meet the needs of researchers undertaking systematic reviews. The model represents a process of comprehensive literature searching, which would produce a transparent and reproducible literature search.	N/A
Key Stage One: who should literature search	Partially agreed: partially meets the needs of researchers	There is an evident need for literature searching in any type of systematic review. Whilst guidance documents indicate that the information professional is the appropriate researcher to undertake the literature searching, limited evidence for this claim was identified.	Refinement: Who should literature search (see 4.4.8)
Key Stage Two: aims and purpose of literature searching	Unclear	The guidance and studies identified appear to support the need for comprehensive literature searches but it is not exactly clear what this means in an information retrieval context and if it applies equally to all types of review. Transparency and reproducibility, on the other hand, is clearly indicated as suitable for all types of systematic review.	Further work: A need is indicated to develop my understanding of comprehensive literature searches and how effectiveness of literature searches are measured. This work will also inform the comparison of the conventional approach to the tailored approach. Work reported in Chapter 6.
Key Stage Three: preparation	Unclear	The suitability of this key stage is clearest for reviews of intervention effectiveness. Elsewhere, there seems a clearer role for the development of this key stage to explore how researcher's plan for reviews, scope the evidence and plan for more complicated reviews.	Refinement: Information needs should determine the process of information retrieval (see 4.4.9)
Key Stage Four: search strategy	Agreed	Whilst PICO/PICOs was the dominate mnemonic for search strategy design, guidance and studies explored alternatives options too. This suggests that this key stage is suitable for researchers undertaking a broad range of reviews.	Refinement: Selecting search methods (see 4.4.10)
Key Stage Five: bibliographic	Partially agreed	Bibliographic database searching is an important literature search method. Its use is most clearly indicated as suitable for reviews of intervention	Refinement: Selecting search methods (see 4.4.10)

database searching		effectiveness. Beyond this, the role and status of bibliographic databases is less clear.	
Key Stage Six: supplementary searching	Partially agreed	Supplementary search methods are well supported in the guidance documents. What is less clear from the guidance documents is when and how it use these search methods, and there generally suitability.	Refinement: Selecting search methods (see 4.4.10)

Further work: is indicated to determine the suitability of these search methods. It is unclear from this review, how supplementary search methods actually work, how they are applied, and the consequent advantages, disadvantages and resource implications of each search method (see Chapter 5). |
| *Key Stage Seven: managing references* | Agreed | The management of references would seem important to all types of systematic review. | N/A |
| *Key Stage Eight: reporting the search process* | Agreed | The reporting of literature searching would seem important to all types of systematic review. | N/A |

3.6 Contributions of this chapter to the thesis

This chapter contains the literature review that supports my hypothesis that leading guidance documents articulate the same process of literature searching in systematic reviews. I illustrate this process model and I call it the conventional approach. This represents a novel contribution to knowledge since the conventional approach has not been identified or examined before.

Identifying and defining the conventional approach establishes what constitutes recommend best practice for literature searching in systematic reviews. This establishes an appropriate benchmark to test my idea of tailored literature searching in the case studies included within this thesis, although, as previously noted, the conventional approach has not yet been tested in practice.

This chapter also helps to develop the idea of tailored literature searching since I have I explored and identified refinements and further work through critique which I examine in Chapters 4 and 5.

The review of models greatly developed my understanding of the theory of information retrieval. The link between information retrieval practice, the focus of this thesis, and information retrieval theory, is not well documented in my experience. The work of Belkin, Wilson and Taylor aligned with the practical problems I had experienced and the scoping surrounding their work allowed me to explore and make firmer the links between theory and practice. This was an invaluable grounding for the exploration and development on the tailored approach, which I shall explore further in Chapter 4.

There are limitations to this work which I acknowledge and I reflect here on their impact when the work is situated in the thesis.

I acknowledge here that studies were single-screened to a loosely defined inclusion criterion. This may affect the confidence which can be placed in the findings and conclusions of the review. I seek to acknowledge and address this limitation by nesting the work in the method of a literature review. Double-screening of studies is not necessarily an acknowledged part of the process of a literature review (121).

I also acknowledge that the selection of guidance documents favours UK and European guidance. I am clear of the rationale for this pragmatic limitation and I question in the limitations section of the review if extending the literature review to include other guidance documents would ultimately change the conclusions I reach.

It is not clear to what benefit reviewing all possible guidance documents would develop further the ideas set out above. Situating the conventional approach in a review of other process models would seem a higher research priority.

I note the distinction between recommend best practice, which I have reviewed, and how researchers interpret guidance in how they literature search. There may be differences in the approaches. This is a limitation of the literature review which I acknowledge.

3.7   Implications for practice

Researchers should be aware that the conventional approach exists and that guidance broadly suggests literature searching in the same way irrespective of the type of review planned, the studies or data sought, and the time available. The existence of this process (the conventional approach) has not been acknowledged before and I believe that it makes a unique contribution to knowledge (see Table 26).

Researchers undertaking qualitative evidence synthesis, or reviews of diagnostic test or prognostic accuracy, and public health systemic reviews, should consider the implications of guidance and the conventional approach on their work. It is possible that alternative process models to the conventional approach may be better suited to the information needs of their work. In Chapters 7 and 8 I test this and I demonstrate that researchers can work with researcher funders, guideline commissioners and organisations such as the Cochrane Collaboration, to explore alternative approaches to literature searching. The review of supplementary search methods in Chapter 5 may also inform the use of alternative approaches to study identification. Since it has been published in an open-access journal, it is freely available to researchers.

It may be time to revisit some of the underlying assumptions about how studies and study data are identified. Researchers may need to familiarise themselves with 'novel' types of evidence, how to identify it in a timely way for use in systematic reviews, and also how to search within large reports to identify data. The identification and use of clinical study reports in health technology assessment, for example. The guidance on supplementary search methods was also limited and the guidance documents have not kept pace with the publication of primary studies. I attend to this gap in knowledge in Chapter 5.

Identifying the conventional approach summarises the studies and evidence available for key stages of the process of literature searching for a systematic review.

This indicates that some key stages presently lack evidence or guidance. An implication for practice is to acknowledge that there are areas of research which are ripe for development and would bring benefit to researchers. A review of which bibliographic management tools to use in a systematic review is an area of low-hanging fruit.

4 The tailored approach

4.1 Introducing the tailored approach

In Chapter 1, I set out the historical context of literature searching in systematic reviews and I started teasing out some of the challenges researchers face when literature searching in reviews which look beyond the evaluation of intervention effectiveness.

In Chapter 3, in the critical literature review, I aimed to test my hypothesis that the research community was (and is) working from a single process model of literature searching in systematic reviews. I found that nine guidance documents reported the same process of literature searching, a process seemingly designed for reviews of intervention effectiveness which included studies reporting RCTs as the primary unit of analysis.

I question if it is appropriate to use the same process of literature searching, designed for systematic reviews which have RCTs as their unit of analysis, as for other types of systematic reviews, which may require different types of studies by design or research method, or where the field of study is less advanced and/or the research is harder to identify, perhaps being largely unpublished.

I suggest that the issues which I began teasing out in Chapter 1, and which I identify through critique in the literature review presented in Chapter 3, are a result of using the same process of literature searching, now identified and illustrated as the conventional approach (see Chapter 3), where an alternative approach to literature searching might be better suited.

As I set out in Chapter 2, the aims of this thesis are to: 1) examine approaches to systematic literature searching for systematic reviews and, 2) propose and test a method of systematic literature searching for reviews that do not focus on the effectiveness of clinical interventions. In Chapter 3, I defined an approach to literature searching from leading guidance documents which I termed the conventional approach. In this chapter, I focus my attention onto tailored literature searching, which I propose as an alternative method of literature searching for reviews which do not focus on the effectiveness of clinical interventions.

4.2 How this chapter addresses the research aims of this thesis

In this chapter, I aim to introduce tailored literature searching. This chapter will, therefore, help me address the research aims of the thesis by setting out what tailored literature searching is as an alternative method of literature searching for reviews that do not focus on the effectiveness of clinical interventions.

4.3 Chapter objectives and research questions

The objectives of this chapter are:

- To discuss the original thinking behind the need for tailored literature searching.

- To theorise how tailored literature searching could work conceptually.

To address these objectives, I identified three research questions:

1. What is the origin of thinking behind tailored literature searching?

2. What is tailored literature searching?

3. How does tailored literature searching work in theory?

By addressing these research questions, I will set out how the work reported in Chapter 3 develops the work reported in this chapter, and I will take tailored literature searching forward for testing in the case studies reported in Chapter 7 and Chapter 8.

4.4 Introducing the tailored approach

The tailored approach is essentially refining and amending the conventional approach to generate a review specific and tailored approaches to literature searching in systematic reviews.

The idea is to develop new and tailored approaches to literature searching in every new systematic review, with a view that every systematic review has its own challenges in information retrieval, and so it is appropriate to tailor the approach to literature searching in order to anticipate and explore these challenges on a review-by-review basis. This is in contrast to the conventional approach which suggests approaching literature searching in the exact same way for every systematic review, irrespective of the research question, information needs and challenges of information retrieval.

4.4.1　The origin of the tailored approach

In 2009/2010, my research was focused on searching bibliographic databases, specifically UK-focused bibliographic databases, to address research questions set by UK decision makers and peculiar to the UK population.

The research question that I was then exploring was: what is the contribution of UK bibliographic databases when compared to North American databases in identifying studies for inclusion into public health reviews for the National Institute for Health and Care Excellence (NICE)?

The method I used to explore this question was a form of cohort analysis, where three separate systematic reviews constituted my cohorts. I prospectively identified i) North American databases and ii) UK-focused databases relevant to the research question and I developed and ran bibliographic database searches for the three systematic reviews on which I was working. After de-duplication, I then 'tagged' the references by database name, before handing them to the review team for screening. After screening had been completed, and studies meeting inclusion had been formally agreed, I checked the table of included studies, categorising and reporting studies by the respective database from which they had been identified.

I was able to determine the contribution of UK databases compared to North American databases in terms of the number of studies each database identified uniquely for each of the three systematic reviews. I presented the work at the joint Cochrane and Campbell Colloquium in 2010 (185) and it led to further work published elsewhere (78).

My conclusion was that UK databases, where they are searched as part of a portfolio of databases, can contribute uniquely to the process of study identification in systematic reviews by identifying studies that would have otherwise been missed. The omission of UK databases, in the portfolio of databases searched for a systematic review, may lead to the omission of relevant research, which may introduce bias into study identification and selection. I named this idea 'resource selection bias' and my work illustrated that researchers can introduce bias through their selection of bibliographic databases.

Whilst it was satisfying to demonstrate this point, it was not a satisfactory conclusion for all people associated with the research. The researchers with whom I was working questioned the need when I identified another study of low methodological

quality that met their inclusion criteria. Their concerns were situated in the work of Egger et. al (28). Namely that; on the one hand, it is important to identify all relevant studies and incorporate them into the synthesis of studies and meta-analysis (where data permits), in order to determine a reliable estimate of intervention effectiveness. On the other hand, the inclusion of a number of smaller studies of low methodological quality can actually increase bias in the review, creating greater uncertainty in the estimate of effect (28). A corollary was that, whilst these low methodological quality studies were of limited value to the synthesis of studies, they often represented research focused specifically on the UK context, and so were valued highly by NICE in decision making.

I learnt two things from this work which influenced my ideas for this thesis:

i) I demonstrated that focusing attention on searching North American databases can introduce bias into study identification. Whilst on the evidence of only one case study, I questioned if the wider concept of bibliographic database searching to the exclusion of other search techniques might be introducing bias in literature searching; and,

ii) whilst finding additional studies that meet inclusion criteria can be important in a systematic review, simply finding and reporting that a greater number of studies were found by one approach to literature search compared to another is insufficient on its own, since it does not articulate the value found in newly identified studies and, therefore, it does not show if the effort involved in study identification was worth it.

By demonstrating the idea of resource selection bias, and with questions outstanding on how researchers value the contribution or effectiveness of study identification, I began to question why researchers search bibliographic databases at all, why bibliographic databases were dominant in literature searching, and if more creative approaches to literature searching would yield different and perhaps better results if research-focused information professionals were more involved in the process of literature searching.

It was also not clear to me how a researcher would demonstrate effectiveness of literature searching between two approaches in a meaningful way, looking beyond a purely quantitative finding to demonstrate value and the rationale for the literature searching. It seemed to me, in 2010, that the use of diagnostic test accuracy terminology to report and articulate the effectiveness of literature searching was

confused in (and between) studies and confusing more generally to read. I explore and confirm this in Chapter 6.

4.4.2   What is tailored literature searching?

Tailored literature searching represents a challenge to the idea of the conventional approach. As I have demonstrated in Chapter 3, the conventional approach is the recommended best practice guidance to literature searching in systematic reviews, but in a way that does not change to reflect the type of review, the research question(s) or the studies needed to address these. The same approach to literature searching for intervention effectiveness reviews is used for qualitative evidence synthesis, for example, despite profound differences in the types of review, research questions and data needs.

In tailored literature searching, the approach to literature searching and the search methods used, are tailored to the research question and information needs of the specific systematic review in question.

I suggest that the review team, which includes the information specialist as the decision maker, invest time at the start of the review process to explore the evidence and research questions before deciding upon a process of literature searching, informed by a selection of suitable literature search methods. The order of search methods is planned relative to the studies and study data and the time available for study identification. It is possible too that not all search methods would be used in a tailored approach.

Furthermore, a new tailored approach to literature searching is developed for every new systematic review which a research team would undertake because, after all, seldom are reviews alike in the information problems and research questions that they must explore. This represents another challenge when compared to the conventional approach since the same process is simply repeated.

4.4.3   How might tailored literature searching work in theory?

On starting this thesis, I illustrated an initial process model for how a researcher would approach a tailored literature search (see Figure 3). Using the numbering in the Figure 3 below, an information specialist or researcher would approach a research question (1) with scoping, investing time in topic immersion (2) before meeting with the research team for discussion on the evidence for review (3) followed

by further scoping (4) and screening of the initial evidence found (5). The stages 1-5 would inform the selection of the search methods (6).

Figure 3 The tailored approach: first drawing (2016)

The Tailored Approach
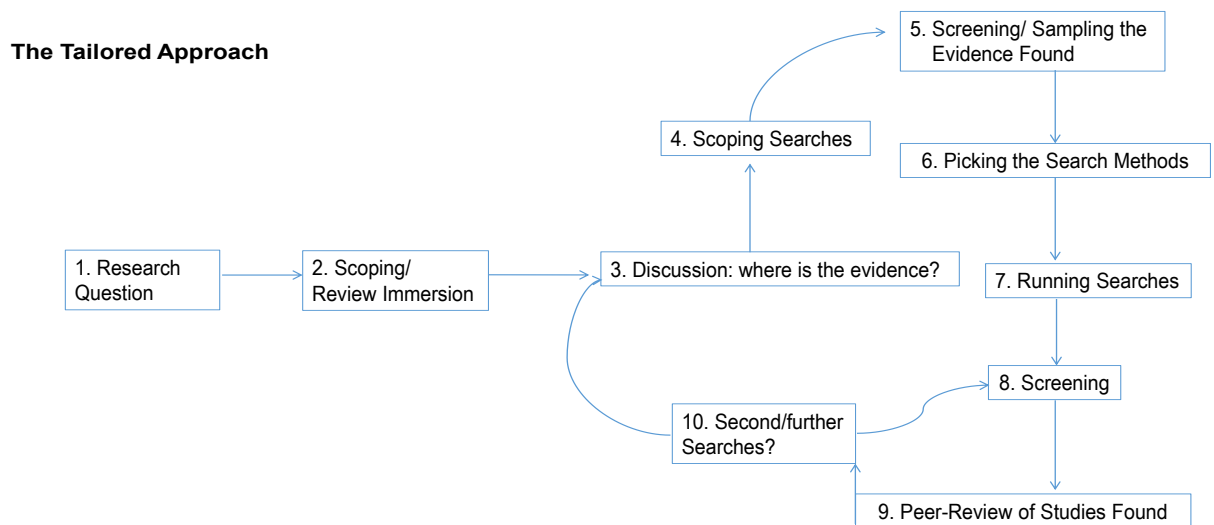
As my work developed, and in response to feedback on this tailored approach model, I have re-thought my initial drawing as set out in Figure 3. I aimed to make the process model clearer, especially around stage 6, 'picking the search methods.'

I set out two new versions of the tailored approach in Figure 4 and Figure 5 below in response to this feedback. The process illustrated is similar in design to my initial drawing in Figure 3 , although the process is arranged differently. What I aim to illustrate more clearly in Figure 4 and Figure 5 below is that the researcher selects the search methods for the review from a range of options. This is different from the conventional approach as the range of search methods is determined by guidance.

Figure 4 A re-design of the process model for tailored literature searching (2018) v1

RESEARCH QUESTION	**1**

SCOPING / REVIEW IMMERSION	**2**

DISCUSSION: WHERE IS THE EVIDENCE?	**3**

SCOPING SEARCHES	**4**

SCREENING / SAMPLING THE EVIDENCE FOUND	**5**

PICKING THE SEARCH METHODS	**6**

CONTACTING STUDY AUTHORS
HAND SEARCHING
CITATION CHASING
WEB SEARCHING
TRIALS REGISTER SEARCHING
BIBLIOGRAPHIC DATABASE SEARCHING

RUNNING SEARCHES	**7**

SCREENING	**8**

PEER REVIEW OF STUDIES FOUND	**9**

SECOND / FURTHER SEARCHES?	**10**

Figure 5 A re-design of the process model for tailored literature searching (2018) v2
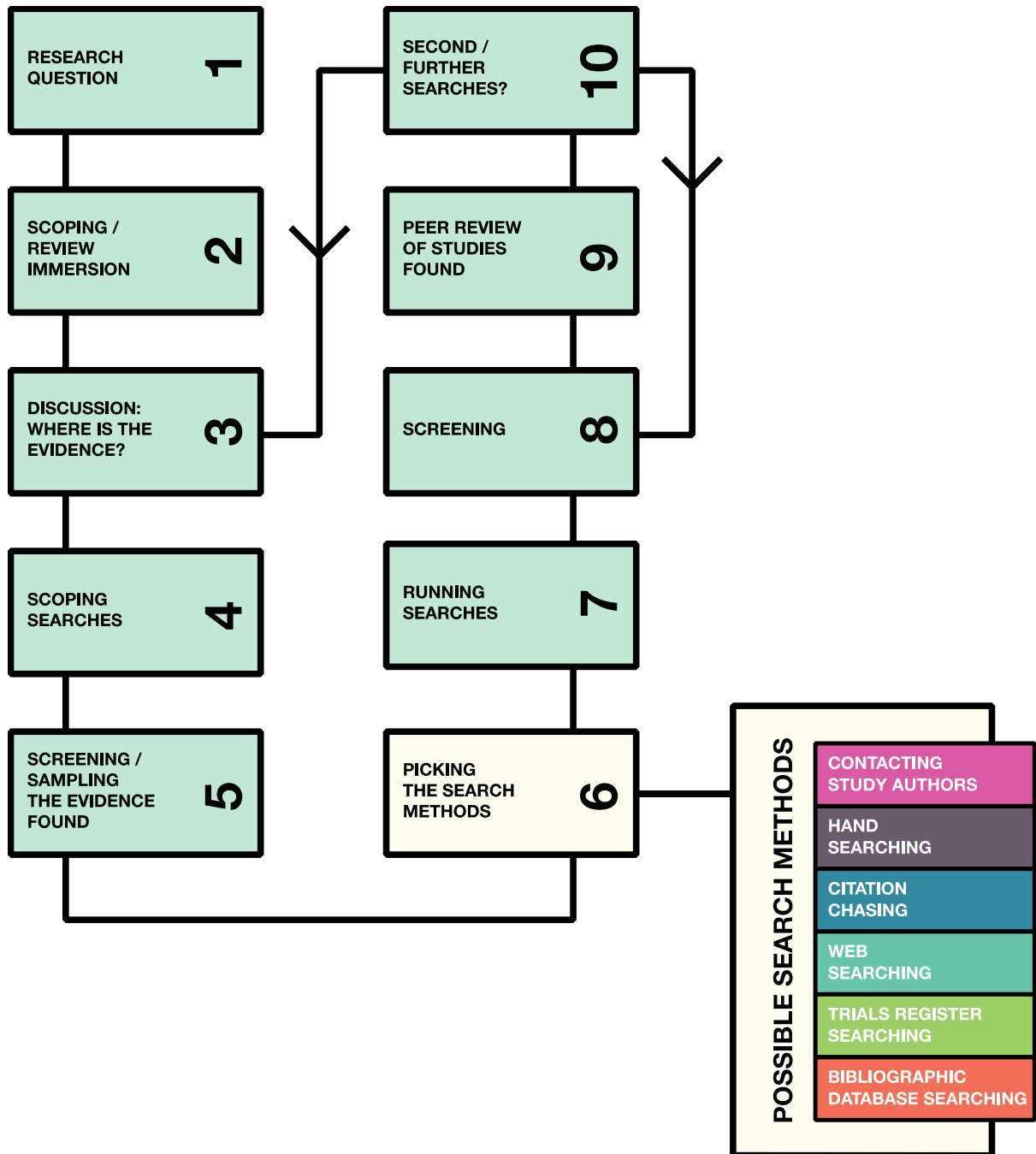
Both Figure 4 and Figure 5 are clearer, in that they set out the search methods from which a researcher can select search methods, to develop a process of study identification in a tailored approach. In thinking through the re-design of the process model for the tailored approach, it is important to articulate that, in every review in which a tailored approach is used, a new tailored approach, unique to the new review in question, should be designed in which new combinations of search methods may be used. To reflect this, I incorporate a further graphic to what I now determine is the

final version of the process model for the tailored approach (as it relates to this thesis).

This final version, presented in Figure 6, sets out the process model and it suggests some example combinations of tailored approaches in an additional graphic to the right-hand side. It should be noted too that where the order of search methods can change on a review-by-review basis, the extent to which the search methods are used may change too. It is possible in some systematic reviews, that not all available search methods would be used.

Figure 6 The final thesis version of the process model for tailored literature searching (2018)
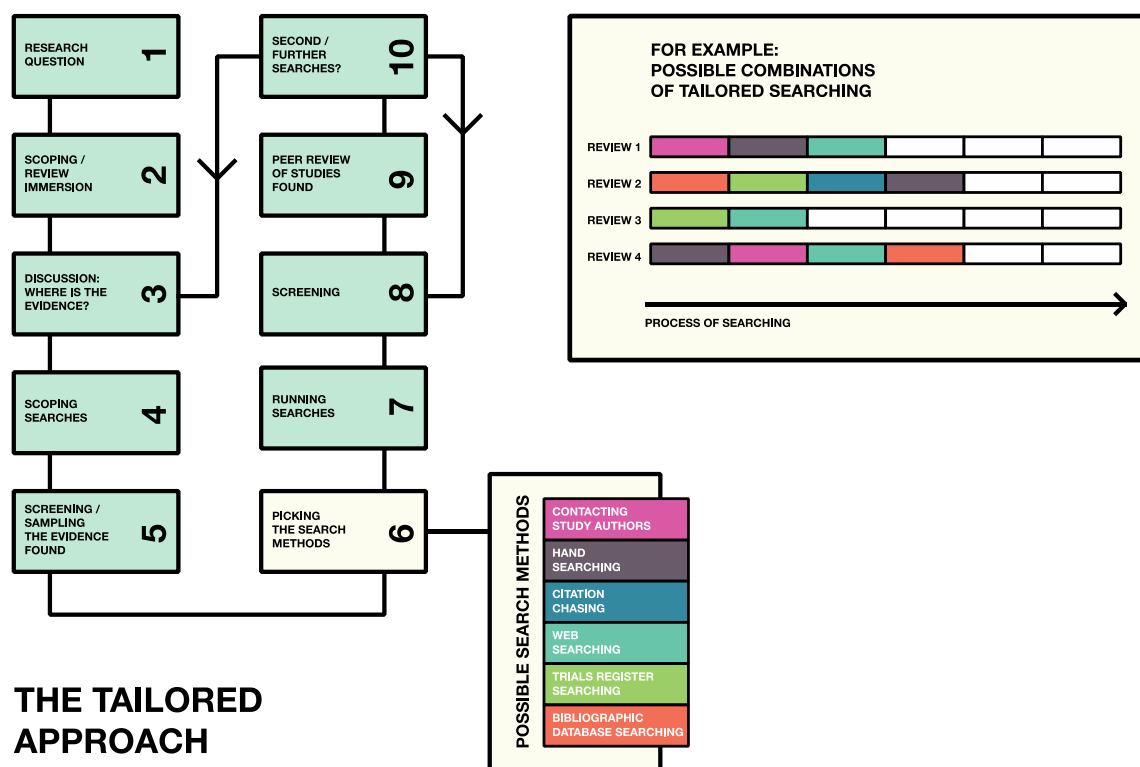
4.4.4 Developing the tailored approach: from theory into practice

Above, I have set out the concepts behind and the process to the tailored approach as it exists in theory. I now move onto how to transition a theoretical process into a practical model of information retrieval which I will then test in the case studies that I present in Chapter 7 and Chapter 8 in this thesis.

4.4.5 Revisiting the information retrieval models: is the tailored approach unique?

Whilst I determined in Chapter 3 that none of the models of information retrieval theory (Table 4) could replace the conventional approach, I wanted to explore if any could replace the need for a tailored approach and/or, to determine to what extent the idea for tailored approaches was unique. I was also conscious that, in exploring other theoretical models, I could learn more about the design and process of information retrieval models.

I reflected on the models of information retrieval identified in

Table 9. No models were identified which would replace the idea of the tailored approach. With a greater focus on picking search methods to match information needs in the tailored approach, there is a clearer need for models which focus on supplementary search methods, and/or seek to match search methods to studies. Two models promote similar search methods to those indicated for use in the tailored approach. Ellis (135), Ellis, Cox and Hall (136) (two papers presenting one model) and Bates' 'Berry picking' model (12) bore close similarities yet with a different overall focus to the tailored approach, which remains to deliver systematic reviews. The models of Ellis (135), Ellis, Cox and Hall (136), and Bates (12), aim for single-use response not a systematic and transparent retrieval of all relevant studies for use in a systematic review.

I determined, therefore, that no alternative or suitable models existed and that the tailored approach would appear to be a unique idea and process model.

4.4.6 Refinements to the conventional approach

I will now explore amendments or refinements to the conventional approach indicated in the preceding chapter's critical literature review (specifically Table 7), and I indicate any further practical work which is required to develop the idea of tailored literature searching.

The critical literature review I present in Chapter 3 allowed me to test my hypothesis that researchers were approaching study identification in systematic reviews using the same process of literature searching. By setting out the conventional approach, I established what constituted recommend best practice for study identification in systematic reviews. In practice, this also allowed to me engage with an established process of literature searching used in systematic reviews and, through critique, explore how I might develop my idea of tailored literature searching, learning from the conventional approach and my own experience. I set out my learning below.

First, I compared the tailored approach, as I set it out in Figure 6, to the conventional approach, which I set out in Figure 1. I present the results of this comparison in Table 8 below. This allowed me to identify areas of agreement between the tailored approach and the conventional approach and it highlighted areas of difference, or where there was no support for the tailored approach. This develops the work set out in Table 7. Secondly, I reflected on my critique of the critical literature review in Chapter 3, specifically the areas which I identified would benefit from further research

and needed refinement on the basis of suitability in the conventional approach. I summarise this in Table 7.

Table 8 Comparing the tailored approach to the conventional approach

The Tailored Approach (Figure 6)	The Conventional Approach (Figure 1)	Description of the tailored approach
1. Research question	Key stage 2	The research question should be discussed by the research team and broken down into one or more research questions with corresponding information needs.
2. Scoping/ review immersion	Key stage 3	Led by the information professional, the review team discuss where the evidence might be, its publication status (published or grey), and the likely study designs or methods.
3. Discussion: where is the evidence?	Key stage 3	As above.
4. Scoping searches	Key stage 3	The information professional tests the evidence using different search methods to identify studies and study data in different ways (i.e. scoping searches of bibliographic databases or web searching).
5. Screening/ sampling the evidence found	Key stage 3	The review team (including the information professional) sample screen studies and data identified. This helps refine screening criteria and informs the team's understanding what does (and what does not) merit inclusion.
6. Picking the search methods	Key stage 3/4	Search methods are selected from all possible search methods available. Search methods are prioritised as to their needs, the time needed for them to work, the likelihood of identifying data and amount of data likely to be developed.
7. Running searches	Key stage 5 and/or 6	Searches are run and de-duplicated.
8. Screening	Not represented	A review task not represented in the conventional approach.
9. Peer-review of studies found	Not represented	The list of included studies is checked and validated for omission of relevant studies.
10. Second or further searches	Not represented	Tailoring recognises that, as the review develops, the understanding of the information need develops. Further searching might be required.

Table 9 Nine conceptual models used to explore the conventional approach to searching

Model name	Author (date)	See figure
The process of asking questions & question negotiation (130).	Taylor (1962)	See appendix: Volume 2 of the thesis
Anomalous state of knowledge (ASK) (131, 132).	Belkin (1980 + 1982)	See appendix: Volume 2 of the thesis
Model of information behaviour (133).	Wilson (1981)	See appendix: Volume 2 of the thesis
Sense-Making triangle (134).	Dervin (1983)	See appendix: Volume 2 of the thesis
Ellis' model (135, 136).	Ellis (1989) and Ellis, Cox & Hall (1993)	See appendix: Volume 2 of the thesis
Berry Picking (12).	Bates (1989)	See appendix: Volume 2 of the thesis
The Information Search Process (ISP) (4).	Kuhlthau (1991)	See appendix: Volume 2 of the thesis
Information needs, seeking and use (INSU) (137).	Byström (1999)	See appendix: Volume 2 of the thesis
The information journey model (138).	Du (2014)	See appendix: Volume 2 of the thesis

This work has led me to develop a list of areas which could be refined in the conventional approach or adopted and developed in alternative process models of literature searching, such as the tailored approach. I set out this learning below, focusing on the indicated three areas of refinement and two additional pieces of work drawn from the learning of the literature review, specifically the critique of suitability (see Table 7).

4.4.7   Refinements to the conventional approach

Figure 7 represents my initial reflections on the refinements to the conventional approach which I identified in Chapter 3. I have mapped the reflections onto the conventional approach process model in Figure 7 and I explore the refinements with critical reflection in the next section.

In proposing refinements to the conventional approach, I should acknowledge that there are elements of the conventional approach which I adopt for the tailored approach without revision (see Table 7). The need for a transparent and replicable report of the process of literature searching, for instance, and the adoption of

guidance on processing and reporting the process of literature searching are well-liked by decision makers and stakeholders. There are areas of agreement between the conventional approach tailored literature searching.

Figure 7 The conventional approach with annotated refinements

The conventional approach assumes that comprehensive searching is necessary and possible (Chapter 3). Comprehensive searching may not be the gold standard of literature searching (175) and it may not be possible (188) or desirable in all types of reviews (188). The aims and purposes of literature searching (including discussion on whether 'comprehensiveness ' is possible/required) should be discussed here in the context of 'the information need' or research question.

'a great deal of exploratory searching goes on, both before and after a topic…is selected' (12). Preparation is key in a tailored search, since it is here that the approach to literature searching is defined. Team discussion is key.

Key stage one: who should literature search? → Key stage two: aims and purpose of literature searching → Key stage three: preparation → Key stage four: the search strategy → Key stage five: searching bibliographic databases → Key stage six: supple-mentary searching → Key stage seven: managing references → Key stage eight: reporting the search process

Guidance and supporting studies are clear that the act of literature searching (as in actually doing the searching) should be performed by information professionals (Chapter 3). In tailored searching, the research team (including experts and stakeholders) are key since, they can help develop the 'information need' (140) leading to a clearer sense of the aims and purpose of literature searching.

The order of searching should reflect the information need(s) and research question. This will be decided in key stage three and it is not necessarily true that bibliographic database searching is the primary method. In some cases, bibliographic database searching may not be used at all in the review.

These stages remain key and there is no obvious reason why a tailored search cannot be both transparent and replicable. Better guidance on the reporting of supplementary search methods is (however) required.

4.4.8 Refinement one: who should literature search?

The guidance documents that I draw upon in Chapter 3 were unequivocally clear: information professionals are indicated as the researchers best placed to undertake the process of literature searching in systematic reviews (3, 25, 81, 115-120).

I initially interpreted this literally and as a clear resolution of the question posed at this key stage. As I reflected, when I critiqued the conventional approach to determine the suitability of the key stages (Chapter 3), perhaps I interpreted this too literally. As I set out in Chapter 3, the evidence for the assertion is not clear.

In Chapter 3, I drew a distinction between the instrumental tasks of information retrieval, which the studies I cited indicate as appropriate tasks for the information professional and include 'doing' and managing the process the literature searching (what I term instrumental tasks in Chapter 3), and the broader conceptual tasks of literature searching. I found in this critique, a broader realisation which stemmed from my work as a researcher and reading the studies which reported the conceptual models. I believe that the information professional is the appropriate researcher to undertake the instrumental and the conceptual task of literature searching in systematic reviews but I have come to realise that they should not do this in isolation.

Reflecting on the models and the work of Dervin (134), Kuhlthau (4) and Belkin (131, 132), I suggest the concept of 'who should literature search' could be considered more broadly:

> We come from different histories and our observations today rest, at least in part, on our pasts (134);

> The criteria for making these choices [from identifying an information need to resolving the it] are influenced as much by environmental constraints, such as prior experience, knowledge, and interest, information available, requirements of the problem, and time allotted for resolution, as by the relevancy of the content of the information retrieved (4); and

> The ASK [anomalous state of knowledge] hypothesis is that an information need arises from a recognised anomaly in the user's state of knowledge concerning some topic or situation and that, in general, the user is unable to specify precisely what is needed to resolve that anomaly. Thus, for the

purposes of IR, it is more suitable to attempt to describe that ASK, than to ask the user to specify her/his need as a request to the system (132).

Using the quotes above, Dervin (134) and (separately) Kuhlthau (4), recognise that 'our' experience of the past may influence how we approach new information problems (4, 134). Belkin argues that, in information retrieval, it is 'more suitable' to describe the anomaly in knowledge rather than to define the problem faced (131).

A team has a greater collective history (viz. Dervin and Kuhlthau) and the people making up the team can work *together* to reduce the uncertainty in the anomaly of knowledge (viz. Belkin). Whilst information professionals are potentially best placed to undertake literature searching (as a practical/instrumental process), conceptualising 'who should literature search' more broadly, to mean who should make decisions on the process of literature searching and select models of information retrieval, a team-based, problem-solving approach may reduce the 'anomaly' in knowledge through the team's collective and 'different histories'. Kuhlthau anticipates this, highlighting the need for discussing possible topics and approaches in the 'initiation stage' of her model (4).

I believe that this refinement represents a gap in knowledge which I address in my second case study presented in Chapter 8. I determined (through team discussion) a need for grey literature, and so I matched this information need to search methods most likely to identify this type of evidence. The information professional knows the search methods and matches this to a process of study identification. The team may have a better understanding of where and how to access this information need and this evidence. This is a good example of a team approach.

I would situate the information professional as the 'decision maker' in this context, where they lead the discussion on shaping the information need relative to a process of, or methods for, literature searching, while drawing upon the team's collective experience of the topic. I test this idea in Chapter 7 and Chapter 8, where I played the role of 'decision maker', leading on discussions on where to identify studies, the language used by researchers to report studies, and the possible need to identify unpublished studies. I explore the idea further in Chapter 9.2, the discussion section of this thesis. This idea of the information professional as the decision maker should also be nested in the critique I set out in Key Stage Three. Here I suggest that the information professional is in transition from librarianship to an active researcher role. I set out the support that may be needed to develop this transition in Chapter 3 and

which, in my experience, is likely necessary for an information professional to be a decision maker. This support is further considered in Chapter 9 and Chapter 11.5.

The concept of the information professional as 'decision maker' extends into the second refinement of the conventional approach discussed below (Section 4.4.9).

4.4.9   Refinement two: information needs should determine the process of information retrieval

What the theoretical information retrieval models provide is some 'clarity' on quite how difficult defining the information need is. Taylor (130) summarises the problem of defining the information need as:

> One person tries to describe for another person not something he knows but rather something he does not know (130).

I suggest above (4.4.8), and I explore in the case studies (see Chapters 7 and 8), how the research team, led by the information professional as 'decision maker', may be best placed to develop a clearer understanding of the information need. I believe that understanding the information need is critical to understanding a key flaw in the conventional approach.

Belkin - specifically - Wilson and (separately) Taylor argue that the process of information retrieval (literature searching) should be matched to the information need (130-133, 284). This would suggest that different information needs may require different literature searching solutions. It is, moreover, possible that research questions in systematic reviews have multiple information needs (285), and that a clear understanding of information needs determines the approach to literature searching.

This suggests that the conventional approach should be represented as set out in Figure 8, where the process of information retrieval (i.e. search methods used) is undecided until the information need(s) is/are clearly understood by the research team.

Figure 8 The conventional approach re-drawn to reflect the information needs (2016)
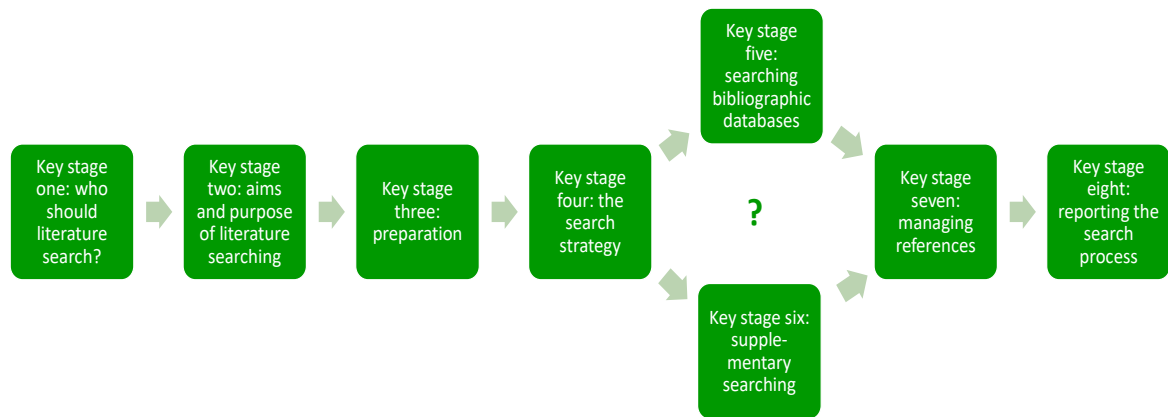
Or perhaps, more specifically and simply, a 'tool box' (101) of research methods, such as I set out below in Figure 9. This 'tool box' forms part of the final version of the tailored approach (see Figure 6).

Figure 9 The searching tool-box
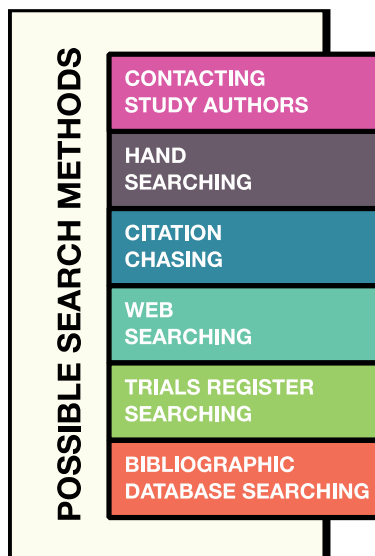
The conventional approach does not account for a link between the information need and the studies (or study data) needed. Rather, whilst aiming at comprehensive literature searching, where the aim is to avoid missing studies, the conventional approach pitches multiple search methods at the information need, in an attempt to minimise bias by not missing studies. I link this to research waste in an information retrieval context. It is my suggestion that research waste in an information retrieval context relates to over-searching and over-screening the results of literature searching in an attempt to demonstrate comprehensiveness. This I suggest stems

from the fact that the process of literature searching in systematic reviews does not change on a review-by-review basis. I explore this further in Chapter 8, where I look to prioritise the search methods most likely to identify study data, following from discussion with the review stakeholders.

4.4.10 Refinement three: selecting search methods

Whilst Belkin, Wilson, and Taylor suggest that the process of information retrieval should match information needs, they do not advocate strategies to determine how this should be achieved (130-133, 284). Information retrieval is represented as a 'black box' function in their models (see appendix, volume 2 of this thesis).

Ellis (135) and Bates (12), provide greater detail on the use of search methods in their models (see appendix, volume 2 of this thesis) but there is again no particular guidance on matching literature search methods to information needs or information retrieval tasks.

The nine guidance documents that underpin the conventional approach (see Table 3) do not advance how literature search methods should be matched to information needs either. In fact, the opposite is true, as I explore in Chapter 3 and demonstrate the effect of in Chapters 7 and 8. The conventional approach aims for comprehensive searches by using all search methods but I contend that the reliance on one search method in particular is worthy of question (185, 286).

I found in Chapter 3 that bibliographic database searching was the primary, perhaps privileged, method of study identification. The guidance documents provided far greater detail on bibliographic database searching when compared to supplementary search methods.[2] Guidance on supplementary search methods, and the evaluation of supplementary search methods, is uneven, with more evidence on some supplementary search methods than for others, and the publication dates of studies for some supplementary search methods suggest that the data available may be out of date.

As I set out above, supplementary search methods are important in tailored approaches to literature searching and I believe that the evidence base would benefit from a review to summarise the guidance and identify the studies available. Such a review would also make a unique contribution to the evidence available on

---

[2] In the CRD handbook, 266 words are used to describe bibliographic database searching whereas 607 words are used to describe all six supplementary search methods.

supplementary search methods, since no review of the literature has been undertaken before. I indicate that this is an area of development for me and a gap in research knowledge. I set out the literature review in Chapter 5.

4.5 Chapter findings

The principal findings were the presentation of the historical context of tailored literature searching, the initial and final illustration of the tailored approach represented as a process model, and a report of my work to develop an information retrieval model from 'in theory' to 'in practice';

I considered and I examined prominent models of information retrieval and behaviour in the context of systematic reviews. Combining the practical review of the conventional approach with the theory of information retrieval is a novel contribution to knowledge. These models helped articulate and nest the importance of defining the information need, the importance of a team approach, and the idea of matching individual search methods to information needs.

I further critiqued the conventional approach (defined in Chapter 3), identifying three areas of refinement and development, namely: who should literature search, the role of the team in literature searching and selecting search methods. I explore these in context of the tailored approach, suggesting that the tailored approach should adopt the information professional in the role of decision maker, that the research team should work with the information professional as decision making to make decisions about the approach to literature searching, and that information needs should match the selection of search methods and define the approach to literature searching.

4.6 Contributions of this chapter to the thesis

The contribution of this chapter to my thesis is to establish what I mean by tailored literature searching.

In this chapter, I have set out: what is the tailored approach to literature searching; the origin of my thinking behind tailored literature searching; and how I think tailored literature searching can work in theory.

I explore the refinements indicated in the conventional approach (Table 7) and I look again at the prominent models of information retrieval, finding that none specifically meet the needs of researchers undertaking systematic reviews. This work, and the

presentation of theoretical process models, make clearer the idea of tailored literature search approaches.

I also indicate two further pieces of work needed to support the development of tailored literature search approach (the literature review of supplementary search methods) and the evaluation of tailored approaches to the conventional approach (a review of how effectiveness is determined in literature searching). I set these pieces of work out in the next two chapters.

4.7 Implications for practice

The implications for practice of this chapter will become clearer in the two case studies were the suitability of the work presented here is tested and evaluated.

The immediate implication for practice is the realisation that refinements to the conventional approach are indicated. This will likely be of greatest interest to researchers undertaking systematic reviews which do not focus on intervention effectiveness. The refinements discussed above, in particular, exploring the role of the team in developing an approach to literature searching would seem of interest.

The refinements indicated above are not without interest for researchers undertaking effectiveness systematic reviews. As I briefly set out in Chapter 1, it might be time to reconsider if the conventional approach still meets the needs for intervention effectiveness systematic reviews.

5 A literature review of supplementary search methods

In the preceding chapter, I introduced the idea for tailored literature searching as an approach to literature searching for use in systematic reviews and as an alternative to the conventional approach, which I set out in Chapter 3.

In setting out how tailored literature searching might work in theory, I identified a gap in knowledge as it relates to supplementary search methods. Belkin (131), Wilson (133) and Taylor (130) suggest that the process of information retrieval (literature searching) should match information needs but they are not prescriptive about how this should be achieved or which literature search methods should be or can be used. If the information need indicates that data not commonly indexed in bibliographic databases is needed, for instance, a researcher might look to the supplementary search methods.

The guidance documents reviewed in Chapter 3 (see Table 3) provide comparatively little guidance on how to undertake supplementary search methods, compared to the guidance they offer on searching bibliographic databases. In the CRD handbook, 266 words are used to describe bibliographic database searching, whereas 607 words (roughly 101 per method) are used to describe six different supplementary search methods (3).

Published studies provide the data on how these supplementary search methods work in practice. The guidance and the studies have not been compared before. Furthermore, the guidance and studies have not previously been summarised in one review.

Undertaking the work reported in this chapter summary will develop my understanding of supplementary search methods, whilst also allowing me to address a current gap in research knowledge: namely, to undertake the first review of supplementary search methods.

5.1 How this chapter addresses the research aims of this thesis

Tailored literature search approaches potentially draw more heavily on supplementary search methods than the conventional approach. Whilst it is suggested that supplementary search methods are valuable in systematic literature searching, it is not clear what they can be used for, how they work, the advantages and disadvantages of using them, and the resource implications involved. This chapter aims to address this gap in knowledge and it will develop my understanding

of how to apply tailored literature search approaches. This work also, and specifically, addresses a research need identified through the critical literature in Chapter 3 (see Table 7).

I aimed to create an evidence-based, literature search 'tool box' through undertaking this review. When matching information needs to search methods, as is the point of a tailored literature search, it was necessary to have up-to-date knowledge to hand which I identified in Chapter 3 was lacking from guidance documents. This literature review therefore aims to bring the knowledge and understanding of supplementary search methods up to the same level of understanding as that of bibliographic databases.

5.2   Chapter objective and research question

The objective of this chapter was to review the review methodology handbooks and empirical studies on supplementary search methods, in order to determine the advantages, disadvantages and resource requirements of supplementary search methods.

I worked to the following research question: how do empirical studies of supplementary search techniques compare to the recommendations in review methodology handbooks?

I identified five supplementary search methods (contacting study authors or experts; citation chasing; handsearching; trials register searching; and web-searching) from the guidance documents.

I aim to determine:

- what the supplementary search method is used for;

- what the evidence says about the supplementary search method;

- claimed advantages of using the supplementary search method;

- claimed disadvantages of using the supplementary search method; and

- resource requirements of using the supplementary search method.

The review that follows has subsequently been published. See Table 1. I report here the final, accepted and pre-publication version of the study.

5.3 A comparison of results of empirical studies of supplementary search techniques and recommendations in review methodology handbooks: A methodological review.

5.3.1 Abstract

Background: The purpose and contribution of supplementary search methods in systematic reviews is increasingly acknowledged. Numerous studies have demonstrated their potential in identifying studies or study data that would have been missed by bibliographic database searching alone.

What is less certain is how supplementary search methods actually work, how they are applied, and the consequent advantages, disadvantages and resource implications of each search method.

The aim of this study is to compare current practice in using supplementary search methods with methodological guidance.

Methods: Four methodological handbooks in informing systematic review practice in the UK were read and audited to establish current methodological guidance.

Studies evaluating the use of supplementary search methods were identified by searching five bibliographic databases. Studies were included if they: 1) reported practical application of a supplementary search method (descriptive); or 2) examined the utility of a supplementary search method (analytical); or 3) they identified/explored factors that impact on the utility of a supplementary method, when applied in practice.

Results: Thirty-five studies were included in this review in addition to the four methodological handbooks. Studies were published between 1989 - 2016 and dates of publication of the handbooks ranged from 1994 - 2014.

Five supplementary search methods were reviewed: contacting study authors, citation chasing, handsearching, searching trials registers, and web-searching.

Conclusions: There is reasonable consistency between recommended best practice (handbooks) and current practice (methodological studies) as it relates to the application of supplementary search methods.

The methodological studies provide useful information on the effectiveness of the supplementary search methods, often seeking to evaluate aspects of the method to improve effectiveness or efficiency. In this way, the studies advance the

understanding of the supplementary search methods. Further research is required however, so that a rational choice can be made about which supplementary search strategies should be used, and when.

Keywords: Supplementary searching, systematic reviews, handsearching, citation searching, web searching, trials searching, author contact, handbooks, information science.

5.3.2 Background

The purpose and contribution of supplementary search methods in systematic reviews is increasingly acknowledged. Numerous studies have demonstrated their potential in identifying studies or study data that would have been missed by bibliographic database searching alone (79, 95, 111, 179, 182, 238, 240, 241).

It is commonly believed that the inclusion of supplementary search methods adds

value to the process of comprehensive study identification in systematic reviews. The methodological handbooks for systematic review methodology, such as The Cochrane or CRD Handbooks, provide practical (although limited) instruction on how to undertake each supplementary search method and empirical studies have evaluated the effectiveness and efficiencies of these search methods. What is perhaps less certain is how supplementary search methods actually work, and what the advantages, disadvantages and resource implications of each search method are.

Study aim: The aim of this study is to compare empirical studies of supplementary search techniques to the recommendations in methodological handbooks.

By re-considering the best practice guidance of methodological handbooks for systematic review, and reviewing how this guidance has been interpreted and evaluated within current practice by authors, this study seeks to identify claimed advantages, claimed disadvantages and resource requirements of using supplementary search methods.

The research question for this study is: how do empirical studies of supplementary search techniques compare to the recommendations in review methodology handbooks?

5.3.3 Methods

This study aims to produce a structured methodological overview of methodological handbooks on the conduct of supplementary searches in systematic reviews. In addition, we reviewed studies that report on the utility and practice of supplementary searches. In order to identify this literature, a systematic approach to study identification, study selection and data extraction was used, which is set out below. These two types of literature – handbooks and practical explorations of

applying supplementary search strategies - were then compared. The advantages, disadvantages and resource requirements of each method were evaluated.

Study identification

We selected the following methodological handbooks as the most influential handbooks in informing systematic review practice in the UK. The current editions of each handbook were read and audited to establish current methodological guidance:

- The Cochrane Handbook for Systematic Reviews of Interventions (version 5.10, March 2011) (25);

- Systematic Reviews: CRD's guidance for undertaking review in health care (2009) (3);

- The Campbell Information Retrieval Methods Group guide to information retrieval (October 2009) (119); and

- The NICE manual to developing NICE guidelines (October 2014) (120).

The following five search methods, supplementary to database searches, were identified from these handbooks:

1. contacting study authors or experts;
2. citation chasing;
3. handsearching;
4. trials register searching; and
5. web-searching.

In order to compare the existing handbook guidance to current practice, we identified studies that describe and/or evaluate how these methods are applied in practice. Studies were identified by searching five bibliographic databases: MEDLINE, EMBASE, LISTA, ASSIA and Web of Science. Forward citation chasing was applied to studies meeting inclusion at full-text and the bibliographies were appraised. Tables of included studies were examined if aggregated within systematic reviews. The search syntax for bibliographic database searching is included as a supplementary file.

Study selection

Studies were downloaded into Endnote X6 were manual de-duplication was performed. Studies were single screened by CC using the inclusion criteria below:

Inclusion criteria

For inclusion in this review, a study was required to:

1. report practical application of a supplementary search method (descriptive); or

2. examine the utility of a supplementary search method (analytical); or

3. identify/explore factors that impact on the utility of a supplementary method when applied in practice.

Exclusion criteria

The following studies were excluded:

1. studies reporting the use of supplementary search methods but not discussing the practical application of the method (such as listing their use to identify studies in a systematic review i.e. 'we handsearched the following journals');

2. studies reported as abstracts, or on-going studies; and

3. systematic reviews or reviews; in which case tables of included studies were examined to identify eligible primary studies.

Data extraction

The following data were extracted: citation details, study design, claimed advantages, claimed disadvantages, and resource requirements.

5.3.4 Results

Thirty-five studies were included in this review in addition to the four methodological handbooks. Studies were published between 1989 - 2016 and handbooks were published between 1994 - 2014. Table 10 summarises which studies cited which handbooks as their source of methodological reference. The handbooks audited for this study cited only three studies: Eysenbach et al. was cited in The Cochrane Handbook; Hetherington et al. was cited in The Cochrane Handbook and The Campbell Handbook and; Papaioannou et al. was cited in The Campbell Handbook (Table 10).

Table 10 studies citing handbooks: handbooks citing studies

Study	Cochrane (1994) (25)	CRD (2008) (3)	Campbell (2010) (119)	NICE Handbook (2013) (120)	Other Source
Adams 1994 (35)	X	X	X	X	NR
Armstrong 2005 (287)	✓	X	X	X	X
Bakkalbasi 2006 (288)	X	X	X	X	NR
Blumle 2008 (289)	✓	X	X	X	X
Bramer 2013 (290)	✓	X	X	X	X

Briscoe 2015 (200)	✓	✓	X	✓	X
Croft 1999 (291)	✓	X	X	X	X
Eysenbach 2001 (292)	X**	X	X	X	NR
Falagas 2008 (103)	X	X	X	X	NR
Gibson 2006 (293)	X	X	X	X	NR
Glanville 2012 (104)	✓	X	X	X	X
Glanville 2014 (294)	X	X	X	X	NR
Godin 2015 (295)	✓	X	X	X	X
Hay 1996 (296)	X	X	X	X	NR
Hetherington 1989 (297)	X**	X	X**	X	Study predates any handbook
Hinde 2015 (238)	✓	X	X	✓	X
Hopewell 2002 (231)	X	X	X	X	NR
Jadad 1993 (242)	X	X	X	X	Study predates any handbook
Janssens 2015 (298)	✓	X	X	X	NR
Jones 2014 (299)	✓	✓	X	X	Institute of Medicine
Langham 1999 (300)	X	X	X	X	NR
Levay 2016 (79)	✓	X	X	✓	X
Mahood 2014 (93)	✓	X	✓	X	X
Mattioli 2012 (301)	X	X	X	X	NR
McManus 1998 (240)	X	X	X	X	NR
Milne 1996 (302)	X	X	X	X	NR
Moher 1995 (303)	X	X	X	X	NR
O'Leary 2003 (304)	X	X	X	X	NR
Papaioannou 2010 (95)	✓	X	X**	X	X
Reveiz 2006 (305)	X	X	X	X	
Robinson 2014 (306)	X	X	X	X	NR
Selph 2014 (307)	✓	X	X	X	Methods Guide for Effectiveness and Comparative Effectiveness Reviews
Stansfield 2016 (308)	✓	X	✓	X	Collaboration for Environmental Evidence
Van Enst 2012 (309)	✓	X	X	X	X
Wright 2014 (310)	✓	✓	✓	X	Institute of Medicine
TOTAL	17	3	3	3	3 other sources

Key: NR, not reported

Notes: Studies in bold are ones cited by the handbooks identified by ** for informing their guidance

Table 11 Overview of results

Method	Includes	What is the method used for	What the evidence says	Implications of evidence	Claimed advantages	Claimed disadvantages	Resource requirements
Contacting study authors	6 studies (see below)	Identify: unpublished, or on-going studies, missing or incomplete data, completed but unpublished studies Expert in field to review includes	Contact original investigators through study report contact details – mainly email/telephone	Email considered effective with better responses from institutional addresses	Additional studies identified; additional study data provided	No guarantee of additional or all relevant information identified Challenging and time consuming Less successful for older studies	Additional resources needed (may need up to 3 contact attempts with authors)
Citation chasing	9 studies (see below)	Identify: further studies, clusters or networks or studies	Backwards and forwards citation chasing using 3 electronic citation databases	Effectiveness of electronic citation methods unclear and suggest using all 3 databases	Not limited by keywords or indexing as bibliographic database searching is	Reliant on the currency, accuracy and completeness of the underlying citation network	Citation chasing of 46 studies = 79hrs or 40 studies = 5 days
Hand searching	12 studies (see below)	Identify: studies or publications not routinely indexed in, or identified by, searches of bibliographic databases, including recently published studies	Manual examination of the contents of topic relevant journals, conference proceedings and abstracts	Use experts to develop list of journals to hand search	Unique study identification, increased sensitivity; identifying studies missed or not indexed in databases	Studies still missed by hand searching; time and access to resources; low precision	Range between 6 minutes and 1 hour per journal
Searching trial registers	3 studies (see below)	Identify: unpublished, recently completed or on-going trials Find adaptations to trial protocols reported study outcomes	Comprehensive list of registries to search	Should be completed as complementary and not in isolation	Unique study identification	Search interfaces lag behind major databases	None reported
Web Searching	5 studies (see below)	Identify: studies not indexed in bibliographic databases. Retrieving grey literature, study protocols and on-going studies	Relevant web-sites and using search engines	Use advanced search functions where possible	Unique study identification, hints to on-going or recently completed studies	Difficulties in transparent search, quality and quantity of searches returned	429 results in 21hrs; google searching 7.9hrs; targeted web searching 9-11hrs

The results were categorised by the supplementary search methods and reported in five domains: 1) what the method is used for; 2) what the evidence says; 3) claimed advantages; 4) claimed disadvantages, and 5) resource requirements. A summary of these results is presented in Table 11.

Contacting study authors

The handbooks focus on identifying contact details and considering how to request studies or study data (3, 25, 311). The studies evaluate the effectiveness of methods to make contact and elicit a response. Six empirical studies were included (3, 25, 240, 293, 297, 304, 305, 307).

What it is used for: Identifying unpublished or on-going studies (3), identifying missing, incomplete, discordant or unreported study data, or completed but unpublished studies (25, 293, 297, 305, 307, 311), and asking study authors (or topic experts) to review a list of studies included at full-text in a review, to see whether any studies had been inadvertently overlooked (3, 25).

What the evidence says: Two handbooks and one study provided detail on identifying contact details (3, 25, 240). The Cochrane Handbook suggests that review authors should contact the original investigators, identifying contact details from study reports, recent publications, staff listings or a search of the internet (311). Colleagues, relevant research organisations and specialist libraries can also be a valuable source of author information and contact details (3, 25). A study by McManus et al. used a questionnaire, primarily to request study data or references, but also to ask recipients to recommend the names of other authors to contact (240). A study by Hetherington et al. contacted authors and experts by letter in an attempt to identify unpublished trials (297).

Two studies reported using a multi-stage protocol to contact authors and request data: Selph et al. devised and followed a protocol that used both e-mail and telephone contact with the corresponding authors at defined stages over a period of 15 days (307). Gibson et al. devised a similar protocol, although focused on e-mail contact, targeting first the corresponding authors and finally the last author and statisticians by e-mail and then telephone (statisticians were contacted due to the specific focus of the case study) (293). Selph et al. contacted 45 authors and 28 (62%) provided study data (307) and Gibson et el. contacted 146 authors and 46 (31.5%) provided study data (293).

Two studies claimed that e-mail was considered an effective method of contact (293, 304). O'Leary reported a response rate of 73% using e-mail contact, finding that more responses were obtained from an institutional address compared to a Hotmail address (86% *v* 57%, p=0.02) (304). Conversely, Reveiz et al. achieved a 7.5% response rate from contacting 525 study authors to identify RCTs but identified 10 unpublished RCTs and links to 21 unregistered and on-going RCTs (305). Gibson et al. found that e-mail was most likely to receive a reply when compared to letter (hazard ratio [HR] = 2.5; 95 percent confidence interval [CI] = 1.3–4.0) but that a combined approach of letter and e-mail, whilst generating a higher response rate, was not statistically different from e-mail alone (73% vs. 47%, p = .36. 146 authors were contacted overall and 46 responded) (293).

Hetherington et al. sent letters to 42,000 obstetricians and paediatricians in 18 countries in an attempt to identify unpublished controlled trials in perinatal medicine (297). Responses were received from 481 individuals indicating they would provide details concerning unpublished studies and 453 questionnaires were completed and returned which identified 481 unpublished trials (297).

Chapter Seven of The Cochrane Handbook offers guidance on how to set out requests for studies or study data when contacting study authors (311). The guidance suggests considering if the request is open-ended, or seeking specific information, and whether (therefore) to include a (uncompleted or partially completed) data collection form or request specific data (i.e. individual patient data) (311). McManus et al. evaluated the use of a questionnaire to identify studies, study data and the names of relevant authors to contact for a systematic review (240). The questionnaire resulted in the identification of 1,057 references unique to the review but no unpublished data were offered (240).

Two handbooks recommend submitting a list of included studies to authors (25) or topic experts (3) to identify any potentially missing studies. The Cochrane Handbook suggests including the review's inclusion criteria as a guide to authors (25).

Claimed advantages: Five studies claimed that identifying additional published or unpublished studies, study data or references is possible by contacting study authors (240, 293, 297, 305, 307). McManus et al. identified 23 references (out of 75 included in the review overall) by contacting study authors (240); Reveiz et al. identified 10 unpublished RCTs and 21 unregistered or on-going RCTs (305); two studies stated

that they identified additional study data but did not separate their findings from contacting study authors from other methods of study identification (293, 305, 307), and Hetherington et al. identified 481 unpublished trials by contacting 42,000 obstetricians and paediatricians in 17 countries (297).

O'Leary found that more detailed study information was provided as a result of contacting study authors (304).

Claimed disadvantages: The CRD handbook claims that contacting authors/experts offers no guarantee of obtaining relevant information (3). Selph et al. found that, whilst identifying additional studies or study data is possible, contacting study authors is challenging and, despite extensive effort, missing data remains likely (307).

Hetherington et al. claimed that methodologically sound trials were not reported through author contact, even by the investigators responsible for them. This was attributed, anecdotally, to the possibility that the trials yielded results that the investigators found disappointing (297).

Reveiz et al. reported low response rates. Of 525 study authors contacted only 40 (7.5%) replied (305).

Two studies and one handbook claimed that contacting authors/experts is time consuming for researchers (3, 293, 307). Selph et al. noted that this method is time consuming for the study authors too, who must identify the data requested (3, 293, 307).

Gibson et al. claimed that contacting authors/experts may be less successful for older studies, given the increased possibility that authors' contact details are out of date (293). Gibson et al. reported a 78 percent (CI=0.107–0.479) reduction in the odds of response if the article was 10 years old or older (293).

Resource requirements: Gibson et al. claimed that additional resources were required to undertake author contact (293). No specific details of the costs or time implications were recorded.

Gibson et al. recorded the duration between the information request and response (293). This averaged 14±22 days (median = 6 days) and was shortest for e-mail (3 ±

3 days; median = 1 day) compared to e-mail plus letter (13±12 days; median=9 days) and letter only (27± 30 days; median=10 days) (293).

Selph et al. reported that all authors who provided data did so by the third attempt, suggesting that repeated attempts to elicit studies or study data may be ineffective (307).

Citation chasing

The handbooks provide a brief overview of the method and list some of the tools commonly used (3, 25). The studies typically evaluate the effectiveness of the tools used to undertake the search methods. Nine studies assessing the use of citation chasing were included (51, 79, 95, 103, 238, 288, 290, 298, 306, 310).

What it is used for: Identifying further studies, and clusters or networks of studies, that cite or are cited by a primary study (3).

What the evidence says: Two studies provided detail on the application of the search method (79, 95). The studies noted that backwards citation searching is undertaken by reviewing bibliographies of relevant or included studies and forwards citation chasing is undertaken by checking if a study, already known to be relevant, has since been cited by another study (79, 95).

Three tools for electronic citation searching dominate the studies: Web of Science, Scopus and Google Scholar. The first two are subscription databases and Google Scholar is presently free (288).

Claimed advantages: Four studies claimed that an advantage of citation chasing is that it is not limited by keywords or indexing as is bibliographic database searching (95, 238, 298, 306). Accordingly, four studies claimed the following advantages: Robinson et al. claimed that a small initial number of studies can create a network (306); Hinde at al. claimed that citation searching can help inform researchers of parallel topics that may be missed by the focus of bibliographic database searches (238); Janssens et al. claimed that citation searching may be valuable in topic areas where there is no consistent terminology, so searches focus on links between studies rather than keywords (298); and Papaioannou et al. reported that citation searching facilitated 'serendipitous study identification' due to the unstructured nature of citations (95).

One study appraised the quality of the studies identified through citation searching (and by other search methods) (95). Papaioannou et al. reported that citation searching identified high quality studies in their case study, although they do not define which quality appraisal tool was used to appraise study quality, so it is not clear if this observation is empirically derived (95).

Claimed disadvantages: Three studies stated that citation searching is reliant on the currency, accuracy and completeness of the underlying citation network (51, 79, 306). Levay et al. identified 'linking lag,' namely the delay between a study being cited and the citation being recorded in a citation database, which impacts on the currency of results (79); Janssens et al. stated that the accuracy and efficiency of citation searching depends on study authors citing studies, which means that selective citation of studies could cause relevant studies to be missed in citation searching (298); Robinson et al. reported limited returns from citation searching where 'broken citation links' created 'island' studies which makes for incomplete citation networks and study identification (306).

Two studies questioned the efficiency of citation searching (238, 310). Wright et al. screened 4,161 studies to identify one study (yield rate of 0.0002) (310) and Hinde et al. screened 4,529 citations to identify 76 relevant studies (yield rate of 0.0168) (238). Wright et al. specifically recorded the time to undertake citation chasing in their study (discussed below in resource use), (310) whereas Hinde et al. did not report the time taken to search but state that the search was 'very time consuming' (238).

Two studies claimed that replicability of citation searching strategies could be affected by the choice of the tools used (79, 290). Levay et al. questioned the replicability of Google Scholar, since search returns are controlled by Google's algorithm, meaning that the results returned will change over time and cannot be replicated (79). Bramer et al. found reproducibility of citation searching to be low, due to inaccurate or incomplete reporting of citation search strategies by study authors (290).

Resource requirements: Two studies recorded the time taken to citation search and one study commented on the time needed (79, 95, 310). Levay et al. reported that citation searching the same 46 studies in Web of Science and Google Scholar took 79 hours (Web of Science = 4 hours and Google Scholar 75 hours) to identify and de-duplicate 783 studies (Web of Science = 46 studies and Google Scholar = 737

studies) (79). Wright et al. reported that citation chasing the same 40 studies in Web of Science, Medline, Google Scholar and Scopus took 5 days in total (two days to download 1,680 results from Google Scholar, one day to download 2,481 results from Web of Science, Scopus and Medline, and two days to screen all the studies) (310). Both studies commented on the administrative burden of exporting studies from Google Scholar which accounted for the majority of time searching in both cases (79, 310). Conversely, Papaioannou et al. claimed reference tracking and citation searching to be minimally time intensive, yielding unique and high-quality studies. The number of studies citation chased, the time taken to search, and the tool used to appraise study quality, were not reported (95).

One study provided data on the costs involved in citation chasing (79). Levay et al. reported that the staff time to search Web of Science for four hours cost between £88-£136 and the 75 hours to search Google Scholar cost between £1,650-£2,550, based on staff grades ranging from £22-£34 per hour (all UK Sterling: 2012) (79).

Handsearching

The handbooks focus on where to handsearch (3, 25) and they provide guidance on who should do this (25). The studies have a similar focus but they have sought to evaluate effectiveness compared with other search methods (35, 231, 242, 296) as well as to evaluate the effectiveness and/or the efficiency of handsearchers in identifying studies (300, 303). Twelve studies were included (35, 104, 231, 242, 287, 289, 291, 296, 300-303).

What it is used for: Ensuring the complete identification of studies or publication types that are not routinely indexed in, or identified by, searches of bibliographic databases, including recently published studies (3).

What the evidence says: Handsearching involves a manual, page-by-page, examination of the entire contents of relevant journals, conference proceedings and abstracts (3, 25, 231, 287).

Two handbooks and six studies provide detail on selecting journals to handsearch (3, 25, 35, 231, 287, 289, 291, 300) . Three strategies were identified, as set out below.

Using databases (or database search results) to identify journals to handsearch

The handbooks suggest that bibliographic databases can be used to identify which journals to handsearch (3, 25). The Cochrane Handbook, with its focus on identifying

studies reporting Randomised Controlled Trails (RCTs), suggests that searches of The Cochrane CENTRAL database, MEDLINE and Embase can be used to identify journals that return the greatest number of studies by study design in the relevant topic area of research (25). Variations of this approach to selecting journals to handsearch were utilised in three studies (35, 287, 300). The CRD Handbook suggests analysing the relevant results of the review's bibliographic database searches in order to identify journals that contain the largest number of relevant studies (3).

Handsearching journals not indexed in bibliographic databases

The Cochrane Handbook suggests that journals not indexed in MEDLINE or EMBASE should be considered for handsearching (25). A study by Blümle et al. considered this strategy necessary to obtain a complete search (289).

Contacting experts to identify journals to handsearch

Two studies contacted experts to develop a list of journals to handsearch (287, 300). Armstrong et al. contacted organisations to develop a list of non-indexed journals to handsearch (in addition to database searching) and Langham et al. used a combination of database searches, contacting organisations and searches of library shelves to identity relevant journals (in addition to database searching) (287, 300). A list of possible journals to handsearch was provided to professional contacts to appraise and identify any missing journals (300). Neither study specifically reports the number of journals identified by experts to handsearch, when compared to the number of journals to handsearch identified by database searching, and there is no discussion of the effectiveness of either method in identifying journals to handsearch.

Five studies explored specifically where or which sections of a journal to handsearch (35, 231, 242, 287, 291). A study by Hopewell et al. handsearched full reports, short reports, editorials, correspondence sections, meeting abstracts and supplements (231). Hopewell et al. found that, of the 369 reports uniquely identified by handsearching, 92% were abstracts and/or published in the supplement of journals (231); two studies reported greatest value in searching supplement editions of journals (242, 287), since these are not routinely indexed in databases (242). Armstrong et al. identified three studies (out of 131) through searching supplement editions of journals (287) and Jadad et al. identified 162 eligible RCTs from a total of 2,889 abstracts reported in four journals (242); Croft et al. claimed value in searching

the correspondence section of journals but they did not record the effect of handsearching this section in terms of identification of studies (291); and Adams et al. reported handsearching book reviews and identifying one study (35).

Claimed advantages: Table 12 summarises a claimed advantage of handsearching, since the studies demonstrate that handsearching identifies studies missed through database searching. Where the studies reported the reason that the studies were missed by database searching (the advantage of handsearching) these are summarised in Table 12.

Claimed disadvantages: Table 12 also summarises a claimed disadvantage of handsearching since, even though this method is often defined as a 'gold standard', the studies demonstrate that database searching can identify studies missed by handsearching. Where the studies reported the reason that the studies were missed by handsearching (the disadvantage over database searching) these are summarised in Table 12.

Two studies claimed that the precision of handsearching was low when compared to the precision found in database searching (35, 242). Table 12 records the relative precision between handsearching and MEDLINE searching. Two studies claimed that the time needed to handsearch, and access to resources (including handsearchers), was a disadvantage of handsearching (287, 302).

Resources requirements: Seven studies reported detail on the time taken to handsearch (35, 104, 242, 287, 291, 302, 303). There was no agreement between the studies on how long handsearching takes. The range was between six minutes (302) and one hour (303) per journal handsearched. It is not possible to calculate an average, since not all studies reported their handsearching as time per journal handsearched. One study reported handsearching in 'two hour bursts' across three months in order to focus concentration but the detail of how often these 'bursts' occurred and the effectiveness relative to 'non-burst' handsearching is not reported (291).

Jadad et al. reported the time taken specifically to handsearch the supplement editions (242). Two thousand, eight hundred and eighty-nine abstracts were handsearched in 172 minutes with an average of 1.1 minute per eligible study identified (242).

The use of volunteers (300, 303) or experienced handsearchers (231, 287) varied in studies. Due to the varied outcome measures used between the studies, it is not possible to aggregate the effectiveness of experienced handsearchers against volunteers. Moher et al., however, specifically sought to test the effectiveness of volunteers in identifying RCTs, finding that volunteers with minimal training can contribute to handsearching (303). Conversely, a study by Langham et al. discussed a possible explanation of their volunteer handsearchers missing studies was a lack of specific knowledge to identify RCTs (300), which suggests experience or training is necessary. Milne et al. suggested that handsearching may need to be undertaken by more than one person (302).

Five studies provided data on training given to handsearchers (35, 104, 231, 300, 303). This included specific training on RCTs (231, 303), a two-hour training session (104, 303) and an information pack including guidelines to handsearching, developed by experienced handsearchers, and a thesaurus of terms to identify RCTs (300).This data was reported narratively and supporting information, such as the information pack reported in the study by Langham et al., was not provided in the studies (300).

Two studies provided guidance on approaches to handsearching if resources were limited (231, 242). Hopewell et al. claimed that, where resources are limited (and it was accepted that studies would be missed), and the aim of searching is the comprehensive identification of studies reporting RCTs, handsearching is best targeted on journals not indexed in MEDLINE and journals published before 1991 (the year the publication type indexing term for RCTs was introduced into MEDLINE (14)) (231). Jadad et al, in a study focused on identifying RCTs, claimed that a combination of MEDLINE searches with selective handsearching of abstracts of letters may be a good alternative to comprehensive handsearching (242).

Armstrong et al. claimed that researchers handsearching for non-randomised study designs may need more time to handsearch. No guidance on speculative timing was given (287).

Moher et al. provided data on costs. Moher et al. recorded costs for photocopying (10-15 Cents Canadian per page) and car parking (10 Dollars Canadian) in their 1995 study assessing the use of volunteers to handsearch (303).

Table 12 Handsearching results

Studies	Advantages		Disadvantages	
	No. identified by handsearching but missed by MEDLINE	Why studies missed by MEDLINE (claimed advantages of handsearching)	No. identified by MEDLINE but not missed by handsearching	Why studies missed by handsearching (claimed advantages of database searching)
Adams 1994 (35)	9% (67 out of 698) of RCTs (CI 7-11%; Sensitivity 94% (CI 93-95%); Precision 7% (CI 6-8%).	conference abstracts and letters not indexed in databases; RCTs not indexed, or no methodological data available to identify studies; methodological descriptors (i.e. 'random' for allocation) were overlooked by database indexers.	*Standard MEDILNE search*: sensitivity 18% (CI 15-21%) Precision 40% (CI 35-45%) *Skilled MEDLINE searched*: Sensitive: 52% (CI 48-56%) Precision 59% (CI 55-65%)	studies missed by searcher error/fatigue; methodological data being 'hidden' in article
Armstrong 2005 (287)	6 out of 131 (4.6%) RCTs/CCTs	trials made no reference in abstract, title or subject headings to random allocation; trials used terms for random allocation it the title, abstract or MeSH but were not correctly indexed by publication type; trials were abstracts; studies were identified in supplement editions of journals not indexed in MEDLINE; and not found in MEDLINE as issue appeared missing in MEDLINE.	125 (of 131) studies *would* have been identified by a MEDLINE using PICO search. 118 (of 131) *would* have been identified by a PICOs search	Not reported
Blümle 2008 (289)	10, 165 RCTs/CCTs out of 18,491(55%)	incorrect indexing and incomplete compilation of health care journals in electronic databases impair result of systematic literature search.	Not reported in abstract	Not reported in abstract
Croft 1999 (291)	7 out of 10 (70%)	two RCTs identified through letter to editors not picked up in MEDLINE search	3 studies identified in MEDLINE (30%)	Not reported
Glanville 2012 (104)	7 out of 25, although none of these studies met the review's inclusion criteria.	Not reported	Electronic searching (including reference checking), by comparison, yielded 30 included papers.	Not reported
Hay 1996 (296)	5 of 40 studies identified (compared to EMBASE) or 13 of 40 (compared to PsycLIT)	Not reported	EMBASE n=35 (out of 40) RCTs (88%) and Precision 9%. PsycLIT n=27 (out of 40) and Precision 9%.	EMBASE: n=3 journal years not indexed n=2 reason unclear. PsycLIT: n=13 gap in indexing and current material being loaded.

Hopewell 2002 (231)	369 out of 714 (52%) RCTs	252/369 (68%) no MEDLINE record. 232/252 (92%) abstracts and/or published in supplements.	32 of 714 (4%)	
Jadad 1993 (242)	*handsearching vs MEDLINE* 25 out of 151 (16.5%) precision 2.7% *handsearching vs supplement editions and MEDLINE.* 150 out of 162 eligible (precision 5.6%).	non-indexed abstract (n=7); non-indexed letter (n=1); search term random not in MeSH or abstract summary (n=9); key search term not in MeSH or in abstract summary (n=7); and no apparent reason (n=1).	*handsearching vs MEDLINE* 2 of 245 (0.8%) *handsearching vs supplement editions and MEDLINE.* 1 out of 13 (7.6%)	Why studies were missed by handsearching is not reported or explored
Langham 1999 (300)	227 out of 710 (32%)	Not reported	MEDLINE identified 118 (16.6%) of studies missed by Handsearching.	Not reported
Mattioli 2012 (301)	0 out of 25 (0), (all identified by handsearching)	Not reported	*Specific PubMed Search* 16 out of 25 (64%) *Sensitive PubMed Search* 9 out of 25 (36%)	Not reported
Milne 1996 (302)	34 out of 82 (41.5%)	Not reported	Capture/recapture used to test. estimated n=3 missed by Handsearching.	inadequate indexing or trials not indexed on MEDLINE prohibits are not located by computerized searches

Searching trials registers

The handbooks focus on the benefit of searching registers (3), with The Cochrane Handbook providing specific guidance on where to search (25). The studies focused on the searching of the registers (294) and the advantages and disadvantages of doing so (299, 309). Three studies were included (294, 299, 309).

What it is used for: Identifying unpublished, recently completed or on-going trials (3, 25, 299, 309) and keeping a track of any adaptations to trial protocols and reported study outcomes (299, 309). Trials that have been stopped, or were unable to reach optimal recruitment, can also be identified.

What the evidence says: The Cochrane Handbook includes a comprehensive list of trials registers to search (25). Distinctions are made between national and international trials registers (which hold trials of any population or intervention), subject (i.e. population) specific registers, and pharmaceutical/industry trial registers (25). There is a further distinction between on-going, completed trials registers, and results registers. Glanville et al. also drew a distinction between trials registers (e.g. ClinicalTrials.gov) and portals to trials registers (e.g. WHO) (294).

Glanville et al. explored the need to search trials registers as a complementary search method to comprehensive searches of bibliographic databases (294). Glanville et al. reported that, in both ClinicalTrials.gov and WHO International Clinical Trials Registry Platform (ICTRP), their 'highly sensitive single concept search' of the basic interface offered the greatest reliability in identifying known records. The methods of searching are explored in greater detail in this study (294).

Claimed advantages: Two studies claimed that searching trials registers will identify unique studies or study data (299, 309). Van est et al. reported that, in four out of 80 Cochrane reviews included in their study, primary studies were identified and included from a prospective search of a trials register search (309). Jones et al. reported that, of 29 studies to record registry search results in their study, 15 found at least one relevant study through searching a register (299).

Two studies claimed that searching of trials registers facilitates checking of *a priori* outcome measures against reported final outcome measures (299, 309). Jones et al. suggested that the comparison of registered trials (and trial data) against published trials (and data) will aid the understanding of any potential bias in the trials (299).

Jones et al. noted that an advantage of trials registers is that they often include contact details for trial investigators, thereby facilitating author contact (299).

Claimed disadvantages: Two studies concluded that trials registers must be searched in combination with other bibliographic resources (294, 309). Glanville et al. concluded that trials registers lag behind major bibliographic databases in terms of their search interfaces (294).

Resource requirements: none were reported.

Web Searching

The handbooks report limited guidance for web-searching. The CRD Handbook suggests that web-searching may be a useful means of identifying grey literature (3) and The Campbell Handbook provides some guidance on how to undertake web-searches, including a list of grey literature websites (119). The studies explored the role of web-searching in systematic reviews. Five studies were included (93, 200, 292, 295, 308).

What is it used for: Identifying published or unpublished studies not indexed or included in bibliographic databases, or studies missed by database (or other) search methods, identifying and retrieving grey literature, and identifying study protocols and on-going studies (3, 119, 292, 308).

What the evidence says: The CRD Handbook makes a separation between a search of the internet through a 'search engine' and searches of specific and relevant web-sites (3). It considers the latter to be more practical than a general search of the world wide web in systematic reviews (3).

The Campbell Handbook provides guidance on searching using a search engine (119) and Eysenbach et al. reported the results of a pilot study to assess the search features of 11 search engines for use in searching for systematic reviews (292)[3]. The Campbell Handbook suggests that, when using search engines, researchers should use the advanced search function. In some cases, this allows searchers to use Boolean logic and employ strategies to limit searches, such as precise phrases like "control group" (119).

---

[3] Eysenbach recommend Alta Vista but this search engine no longer exists.

Godin et al. reported the development and use of a web-searching protocol to identify grey literature as a part of study identification in a systematic review (295). Godin et al. broke their web-searching into three parts: first, searches using Google for documents published on the internet; secondly, searches using custom Google search engines; and thirdly, browsing targeted web-sites of relevant organisations and agencies (295).

Claimed advantages: Two studies identified studies uniquely by web-searching (292, 295). Eysenbach et al. identified 14 unpublished, ongoing or recently finished trials, and at least nine were considered relevant for four systematic reviews (292). Godin et al. identified 302 potentially relevant reports of which 15 were included in their systematic review (295).

Three studies commented on the types of study or study data identified (292, 295, 308). Eysenbach et al. claimed that internet searches may identify 'hints' to on-going or recently completed studies via grey literature (292); Godin et al. uniquely identified report literature (295); and Stansfield et al. suggested that web-searching may identify studies not identified from 'traditional' database searches (308).

Claimed disadvantages: Five studies discussed the disadvantages of web-searching (93, 200, 292, 295, 308). The studies drew illustrative comparisons between database searching and web-searching in order to highlight the disadvantages of web-searching:

Three studies commented on searching using a web search engine: Eysenbach et al. reported that current search engines are limited by functionality and that they cover only a fraction of the visible web (292); Mahood et al. claimed that their chosen search engines could not accommodate either full or modified search strategies, nor did they support controlled indexing (93); and Godin et al. claimed that, in contrast to systematic searches of bibliographic databases, where one search strategy combining all search terms would be used, Google searches may require several search enquiries containing multiple combinations of search terms (295).

Three studies commented on the number of studies returned though web-searching (93, 292, 295). Godin et al. claimed that searching Google can be overwhelming due to the amount of information and lack of consistent organisation of web-sites (295); Mahood et al. had to limit their web-searches to title only in order to control search returns (93), and Eysenbach et al. recorded recall of between 0% and 43.6%, finding

references to published studies and precision for hints to published or unpublished studies ranged between 0% and 20.2% (292).

Three studies commented on the search returns (93, 292, 295). Eysenbach et al and Stansfield et al. commented on the lack of abstracts when web-searching, which impacts on the precision of web-searching and volume of studies identified (292, 308), and Godin et al. claimed that it was impossible to screen all results from a Google search, so researchers were reliant on page ranking (295).

Three studies claimed potential issues with the reliability of items identified through web-searching (200, 292, 295). Godin et al. discussed the possibility of bias created in web-searching, where search results are presented depending on geographic location or previous search history (295); Briscoe reported that algorithms used by search engines change over time and according to the user, which will influence the identification of studies and impact the transparency and replicability of search reporting (200); and Eysenbach et al. reported identifying a study published on-line that differed in reporting to the copy published in the peer-reviewed journal, where adverse event data was omitted in the on-line version (292).

Stansfield et al. claimed that the lack of functionality to export search results presented a challenge to web-searchers (308). Three studies claimed that web-searching presented difficulties in transparent search reporting (93, 200, 295).

Resources requirements: Two studies discussed time taken to web-search (292, 295). Eysenbach et al. reported searching 429 returned search result pages in 21 hours (292) and Godin et al. reports custom Google searching taking 7.9 hours and targeted web-searches taking 9-11 hours, both timings being specific to the case studies in question (295).

Stansfield et al. discussed planning when to undertake web-searching (308). Stansfield et al. linked planning a web-search to the time-frame and resources available in order to inform where to search (308).

Mahood et al. claimed that large yields of studies can be difficult and time consuming to explore, sort, manage and process for inclusion (93). Mahood et al. initially had to limit their web-searching to title only (as a method to control volume) before eventually rejecting their web-searching due to concerns about reproducibility and ability to manage search returns (93).

No studies reported any data relating to the costs involved in web-searching.

5.3.5 Discussion

The discussion will focus on two elements inherent in the research question of this study: how does current supplementary search practice compare with recommended best practice and what are the implications of the evidence for searching using these supplementary methods.

Contacting study authors

The advent of e-mail (and more specifically the standardised reporting of e-mail addresses for corresponding study authors) would appear to have improved the efficiency of contacting study authors (3, 119), although it is possible that it hasn't altered the effectiveness (312). Identifying additional studies or data (the effectiveness) is conditional upon a reply, whatever the method of contact. The guidance of the handbooks, to consider how best to set out requests for studies or study data, is well made but seldom explored in the studies themselves. Whilst making contact is important, which the studies evaluate, exploring techniques to improve the rate of reply would be a valuable contribution to improve the efficiency and effectiveness of identifying studies or study data through author contact.

When to contact study authors is worthy of consideration, since the studies included in this review reported a delay between asking for studies or study data and a response. Sufficient time should be allowed between identifying the need for author contact, making contact, a response being provided, and the study or data being integrated into the review (with all the methodological implications considered). A recognition for the need of this method, combined with the realisation that this method takes time to yield results, is important. It is perhaps for this reason that, whilst contacting authors is common in systematic reviews, it is not a method of study identification that is undertaken as a matter of course (313).

The concept of contacting authors could also be understood more broadly than simply contacting with a view to requesting known studies or data. Whilst in contact with authors, requests for unpublished, linked or forthcoming studies are not unreasonable requests, and authors can assist with the interpretation of specific elements of studies or topics, in order to aid the process of critical appraisal. Furthermore, Ogilvie et al. found the value in contacting experts was the link to better reports of studies already identified (94, 314). This highlights the potential flexibility of

the search method: it is not only the chance to identify known studies or study data but also it offers the opportunity to <u>speak</u> with experts.

Citation chasing

The advantages and disadvantages (and resource requirements) were most clearly stated for this supplementary search method. The handbooks, and some studies, suggested and found advantages and disadvantages in the methods and tools.

The Cochrane Handbook suggested that there is little evidence to support the methodology of citation searching, since the citation of studies 'is far from objective' (315). The studies included in this review suggested that the reasons for 'non-citation' are unclear and could range from selective citation (i.e. selective reporting) to pragmatic reasons, such as a review of trials being cited instead of each individual trial reviewed (306). Furthermore, a high number of citations for a study should not necessarily be confused as an indicator of study quality (316, 317) or a complete citation network. Non-citation of studies, or 'linking lag' (79), forces a break in citational networks (79, 238, 306), meaning it becomes unclear when (or if all) studies have cited a primary study (298). There is presently no method to assess the completeness of citational networks and no certainty as to the comprehension of any citation chasing.

There is little common agreement between the studies as to which tool (or combination of tools) is superior in citation chasing, since the relative merits of each resource depend greatly upon the topic of review, the data range of the resource, and the currency of the results (c.f. (79, 103, 105, 290, 318, 319)). A study that evaluated the tools (Web of Science, SCOPUS and Google Scholar), how the tools are best searched, how the platform hosts select data for inclusion, and the advantages and disadvantages of use, would make clearer statements on when (or if) to use which tools.

There are, undoubtedly, advantages to citation searching. The citational link is neutral, in the sense that it only links the studies but it does not explain the nature of the link. This is important, since a citation search will identify **<u>any</u>** study linked to the primary study, including erratum studies, studies that dispute or disagree with the primary study, and it should also link different publication types, such as editorial content, reviews or grey literature. This could not only aid interpretation of studies but also it could help researchers explore the idea of study impact. Furthermore, as

reported in the results section, a citation search links by citation and it is not beholden to the use of 'the correct' search terms or database indexing. It may, therefore, as Papaioannou et al. reported, facilitate serendipitous study identification (95), suggesting that citation chasing is valuable in scoping review topics, to aid development of searches, and review searches, in order to ensure all studies have been identified.

The nature of bi-directional citation chasing suggests that, given the relative specificity, this method could possibly be used to efficiently update systematic reviews using known includes as the citations to chase (298). Researchers have had positive, although incomplete, success trialling this method and studies suggest that citation chasing alone is not a substitute for standard update searches (102, 320).

Handsearching

The evidence on handsearching can be summarised as: 1) selecting where to handsearch; 2) what to handsearch and 3) who does the handsearching. In relation to 1, the handbooks advocate selecting journals to handsearch on the basis of the number of relevant studies included from journals identified in database searching. This approach means handsearching is a supplementary method to database searching, since to undertake handsearching – following this method – database searches define the list of journals to handsearch.

Studies included in this review provided empirical evidence that handsearching journals identified by database searching was effective in identifying studies missed by poor indexing, lack of study design or omission of key search terms, or where sections of journals are not indexed on databases. In this way, this approach to selecting journals to handsearch could be categorised as a 'safety net search,' since it aims to identify studies missed by deficiencies in literature searching and database indexing. This approach to selecting journals to handsearch, even though it is effective, could be argued to be a duplication of effort, since the journals being handsearched have already been 'searched' through the bibliographic databases. This is likely why the studies recorded low precision (compared to database searches) and why handsearching takes longer (242).

The Cochrane Handbook and three studies suggested alternative ways to identify journals to handsearch: namely, selecting journals not indexed on MEDLINE or EMBASE (25, 289) – a suggestion that is easily changed to read 'primary databases'

relevant to the field of study (i.e. ERIC for reviews of educational topics) - and contacting experts, contacting organisations, and searches of library shelves (287, 300). Neither the study by Armstrong et al. nor the study by Langham et al. listed the journals identified by method of identification, so it is not clear if there were differences between the list of journals provided by experts when compared to those provided by databases (287, 300). This review did not identify any studies that compared the use of databases to identify journals to handsearch as against these alternative methods but such a study may be of value if efficiencies could be found in practice.

It may be that, in reviews in which a comprehensive identification of studies is required, identifying journals to handsearch should be done both by using databases and contacting experts or organisations. The former being to cover any deficiencies in the database searching and the latter to capture any unique journals or conferences known to experts but not indexed in databases.

Selecting what to handsearch and who should handsearch was another notable difference between the handbooks and studies. The studies included in this review identified studies uniquely from handsearching various sections of journals (from abstracts through to book reviews), and the studies used volunteers, provided training to handsearchers, and used experienced handsearchers to handsearch, with varying degrees of success and failure since handsearching relates to effectively identifying studies when compared to database searching. The Cochrane Collaboration arguably has one of the longest track-records of handsearching projects (c.f. (14)) and it is their recommendation that handsearching is the page-by-page examination of the entire contents of a journal (3, 25) by a well-trained handsearcher (25). Handsearching is commonly referred to and used as a 'gold standard' comparator to establish effectiveness of other search methods. Given that every study included in this review uniquely identified studies by handsearching but also missed studies by handsearching too, a reminder of what constitutes handsearching is likely warranted.

Trials registers

The handbooks provide guidance on where to search and the studies focused on the effectiveness of study identification in selected registers and/or the practicalities of searching registers. In this way, the studies advance the guidance of the handbooks, since they provide empirically derived case-studies of searching the registers. The

implications for searching, however, are clear: searching trials registers should still be undertaken in combination with bibliographic database searching (294, 321). Even despite the aims of the International Committee of Medical Journal Editors (322), comprehensive and prospective registration of trials – and keeping the trial data up to date – is still not common place. It is unclear what pressure (if any) is put upon trial managers who do not prospectively register their trials and, in fact, if there is any active penalty if trial managers do not do so. Until this issue is resolved, the comprehension of registers will remain uncertain and a combination of bibliographic database searching (to identify published trials) and searches of trials registers (to identify recruiting, on-going or completed trials) is required.

The advantages of searching trials registers are worthy of discussion. Registered trials include an e-mail address for trial managers, which can facilitate author contact, and the studies concluded that more consistent searching of trials registers may improve identification of publication and outcome reporting bias (299, 323). If trial managers were using the portals correctly, it would also be a practical method of reporting results and sharing study data, perhaps akin to a 'project website,' as recommend in the Cochrane Handbook (25). The variability of the search interfaces is notably a disadvantage and something upon which could be improved. Glanville et al. observed that the search interfaces lag behind major bibliographic databases (294). If the registers themselves are hard to search (and in some cases impossible to export data from), they are less likely to be searched. Trial managers and information specialists/researchers could usefully work together with the registers to develop the interfaces in order to meet the needs of all who use them. The use of trials registers may be broader than only researchers (324).

Web-searching

In their 2001 study, Eysenbach et al. stated that the role of the internet for identifying studies for systematic reviews is less clear when compared to other methods of study identification (292). The handbooks do not update this view and very few studies were identified in this review which improve upon Eysenbach et al's claim. The studies have attempted to take on Eysenbach et al's suggestion that a systematic investigation to evaluate the usefulness of the internet for locating evidence is needed. Mahood et al., however, had to abandon their attempts to web-search (93) but Godin et al. took this work a little further in their case study with reference to identifying grey literature (295).

The comparative lack of guidance in the handbooks could stem either from a lack of certain knowledge of how to web-search or perhaps a lack of certainty of how to do this systematically, such that web-searching could be replicable, and therefore, be included as a method to identify studies without introducing bias. Researchers are exploring the idea of how far web-searching can meet the need to be replicable and transparent but still functional (200). Further guidance is undoubtedly needed on this supplementary search method.

Limitations

The date range and age of the handbooks and studies included in this review could be considered a limitation of this study.

Comparative and non-comparative case studies form the evidence base for this study. The studies included in this review have been taken at face-value and no formal quality appraisal has been undertaken since no suitable tool exists. Furthermore, supplementary search methods are typically evaluated in the context of effectiveness, which is potentially a limited test of the contribution they may offer in the process of study identification. Different thresholds of effectiveness and efficiency may apply in the use of supplementary search methods in systematic reviews of qualitative studies when compared to reviews of RCTs, for example.

The studies themselves do not necessarily correlate to the concepts of claimed advantages and disadvantages. In most cases proposed advantages and disadvantages have not been tested in practice.

Whilst have aimed to comprehensively identify and review studies for inclusion, the use of supplementary search methods is a broad field of study and it is possible that some completed studies may have been inadvertently missed or over-looked. It is possible that standard systematic review techniques, such as double-screening, would have minimised this risk, but we are confident that, whilst a more systematic approach may have improved the rigour of the study, it is unlikely to alter the conclusions below.

5.3.6 Conclusions

Current supplementary search practice aligns methodologically with recommended best practice. The search methods as recommended in the handbooks are perceptibly the same methods as used in the studies identified in this review. The

difference between the handbooks and the studies is of purpose: the studies sought to test the search methods or tools used to undertake the search methods.

The causal inference between methods (as presented in the handbooks) and results (as found in the studies) could be usefully tested to develop our understanding of these supplementary search methods. Further research is needed to better understand these search methods. Specifically, consistency in measuring outcomes, so the results can be generalised and trends identified, which would provide a link not only to better effectiveness data but also to efficiency data, offering researchers a better understanding of the value of using these search methods, or not.

Time

All of the studies discussed in this review claimed to identify additional includable material for their reviews using supplementary search methods that would have been missed using database searches alone. Few of the studies, however, reported the resources required to identify these unique studies. Further, none of the studies used a common framework, or provided information that allows a common metric to be calculated. It is not, therefore, possible to compare the resources required to identify any extra study with each search method. This, alongside the use of comparative and non-comparative case studies as the primary study design to test effectiveness, limits our ability to generalise the results of the studies and so reliably interpret the broader efficiency of these search methods. Researchers could usefully consider reporting the amount of time taken to undertake each search method in their search reporting (242, 325).

Value versus Impact?

Identifying unique studies is commonly interpreted as adding value to the review and the process of searching in and of itself. Only three studies sought to extend this, appraising either the quality of the studies identified, or the contribution of the studies to the synthesis as a way of considering the value of the additional studies (95, 240, 292). In reviews of effectiveness, where all studies should be identified so as to generate a reliable estimate of effect, study value might be a moot point but, in resource limited situations, or for reviews where a comprehensive identification of studies is less important, study value is an important metric in understanding the contribution of supplementary search methods and the extent to which researchers invest time in undertaking them.

Time + Value

Comparing the time taken to search, with a summary estimate of the contribution or value of the studies identified uniquely, against the total number of studies identified, could alter how researchers value supplementary searches. It would permit some basic form of retrospective cost effectiveness analysis, which would ultimately move literature searching beyond simply claiming that more studies were identified to explaining what studies were identified, at what cost, and to what value.

5.4   Chapter findings

Having reproduced the journal article above, I now situate the findings in the context of this thesis.

I categorised the advantages, disadvantages and resource requirements of five supplementary search methods: contacting study authors; citation chasing; handsearching; trials register searching; and web searching. The contribution of supplementary search methods is increasingly acknowledged but, what is less certain, is how supplementary search methods actually work and the resource requirements of using them. This chapter and associated publication, therefore, addresses a gap in the current literature and in current guidance, as I identified in the Critical literature review reported in Chapter 3.

Understanding the advantages, disadvantages and resource requirements of supplementary search methods can inform the selection of search methods in systematic reviews. This makes the findings of this chapter (and publication) a useful resource and contribution to knowledge for information professionals and researchers.

The findings, as it relates to this thesis, aim to support the selection of search methods in tailored literature searching approaches. For example, this informed the selection, and order, of search methods used in the tailored approach to literature searching in Chapter 8 (77).

5.5   Contributions of this chapter to the thesis

In the absence of guidance (identified as an issue in the critical literature review (Chapter 3), and explored as a refinement to the conventional approach in Chapter 4), and without an existing published review of supplementary search methods, information professionals and researchers are potentially making decisions on using

supplementary search methods with no clear guidance or evidence on their purpose, their effectiveness and resource requirements.

In tailored literature searching, it is possible that supplementary search methods play a greater role in the process study identification (something that I explore and demonstrate in my case studies reported in Section Two). The review of supplementary search methods reported above has the potential to represent a 'searching tool kit.' A resource which I (and other information professionals and researchers) could consult to examine what supplementary search methods are used for (i.e. what types of study or data could I identify using a particular supplementary search method), their potential advantages and disadvantages (i.e. when to use them), and the resource requirements involved (i.e. how long it takes to use or how much using them might cost). For tailored literature searching, this could inform the selection of supplementary search methods and how/when they are prioritised for use in a systematic review. This is how I used the work reported in this chapter in the second case study, reported in Chapter 8.

In situating this review in the thesis, and reviewing the contributions of this work to the broader thesis, I identify some limitations which represent learnings from the work.

1) The title of the published work.

Whilst the title accurately reflects the research question of the study, it is not an easy handle with which to locate the work it presents, and the study is not situated in a recognised research method.

On reflection, I would alter both aspects now. I would re-title the work 'a literature review of supplementary search methods.' This title accurately reflects the work undertaken and it situates it in the context of an acknowledged method of research and review. This would, I think, further increase access to and uptake of this study;

2) Literature review or systematic review

This work was undertaken as a literature review and not a systematic review for pragmatic reasons within the constraints of a Ph.D. This meant that I conducted the research alone.

On reflection, I acknowledge that a second reviewer may have strengthened the study, particularly in the interpretation required to categorise the advantages and disadvantages of the search methods. The categorising I present in the study is

based on one researcher's (my) informed opinion. It may be preferable that a second reviewer was involved but only if it is assumed that a systematic review was necessary for this review of supplementary search methods which may not be the case.

Wagner et al. explore the idea that decision makers are sometimes willing to exchange certainty for speed (181). Whilst not strictly relevant to this problem, since I am exploring a methodological problem and not a decision relating to a guideline or reimbursement, Wagner et al's work finds and it suggests that absolute certainty is not a necessity in decision making if decision makers are willing to accept a review of lower rigour (181). I suggest that this review represents a similar case. The purpose of this review was to combine the guidance and studies relating to supplementary search methods into one review and this has been done for the first time. It is intended to, and it will, serve as a summary for future work. I am not confident that a second reviewer and change in review methodology would improve uptake of the work or alter the findings and conclusions which I found but I cannot be sure; and

3) Critical appraisal

I did not undertake critical appraisal of the studies included in the review. This is defensible since the study as I present it is a literature review and critical appraisal is not necessarily a recognised methodological component of a literature review (121).

Critical appraisal is central to informed decision making (326) since assessment of study quality gives an indication of the strength of evidence or reliability of the findings (3). Undertaking critical appraisal of the studies identified and included in this review would have contributed to knowledge by determining the quality of the studies, offering a quality score on a study-by-study basis. The advantage of this is clear. It would have allowed me to draw conclusions about the quality of the studies and the underlying evidence-base, and comment on the strength by which a researcher could be certain that the effect detailed in the study (for example) was reliable. Furthermore, I would have been able to ground future research recommendations more strongly in the context of lessons learned through critical appraisal.

In the study, I cite the absence of a suitable critical appraisal tool as the reason for not undertaking critical appraisal (6). This is valid since the absence of a suitable tool would have meant devising a tool for the purposes of this study and potentially without validation. This creates its own limitations but I acknowledge the possible

value of determining study quality more clearly now. It would have been a valuable extension to this work.

5.6   Implications for practice

This chapter (and publication) draws together the primary studies that evaluate supplementary search methods into one review for the first time and it correlates these studies with guidance. This highlights that the guidance documents seldom cite empirical studies, which suggests that researchers need to read guidance to understand when to use supplementary search methods and to then read studies to know how to use them. The chapter seeks to address this issue by addressing guidance and studies in one review for the first time.

The review reported in this chapter would also seem to demonstrate that guidance has not kept pace with the developments of supplementary search methods and evaluation studies. On-line resources such as SuRe Info (summarised research in information retrieval for HTA) may prove more useful to researchers since they are better able to incorporate new studies and evidence than published guidance handbooks. The publications reported in this thesis have already been incorporated into SuRe Info by researchers.

The principle implication for practice is that, by publishing in an open-access journal, I report the first review of supplementary literature search methods. By setting out advantages, disadvantages, and resource requirements, I set out the evidence available on supplementary search methods which will help researchers make informed decisions when selecting supplementary search methods for use in systematic reviews and other types of evidence synthesis.

The review highlights how little researchers know about these supplementary search methods and how variable the findings of the primary studies are. A study by Levay et al. and a study by Wright et al., for instance, provides conflicting results on the use of citation searching suggesting that further studies may be required (79, 310).

This review also highlights that the age of the studies and the data they report would benefit from updating so that researchers can make decisions on more recent data. The only study to provide data on the costs involved with handsearching was a 1995 study by Moher et al. which set out photocopying charges of 10-15 cents and parking charges of 10 dollars Canadian. Whilst it would not be difficult to identify up-to-date costs, it demonstrates that the studies researchers are using to make decisions

about supplementary search methods are potentially out of date (303). This is potentially true for the effectiveness findings too, where the effect of technology may have altered the underlying efficiency of some studies. I indicate that this is an area for further research (see Chapter 11.3).

6 How do researchers measure effectiveness in literature searching and how I will compare the conventional approach to the tailored approach in the case studies that follow

At the start of work on the thesis, I was aware of the body of work invested into the evaluation of study design literature search filters. The prevailing metrics to evaluate effectiveness of these literature search filters appeared to borrow heavily from the evaluation of diagnostic tests, treating a handsearched data set as the reference standard and the search filter as the index test. A measurement of effectiveness was derived using similar language and methods.

At the outset of this thesis, this felt an unsatisfactory way to determine the effectiveness of literature searching, particularly in the case of this thesis, which seeks to compare two approaches to literature searching and not one method to another, as historically has been the case. This approach felt unsatisfactory because:

first; a researcher could only determine 'effectiveness' at the end of a process of literature searching, and arguably when it was too late to improve upon the outcomes of the literature search process;

secondly; the evaluation of literature searching in the context of sensitivity and specificity would appear to provide absolute summaries of effectiveness to a reader which is inconsistent with what it is possible to know (189, 327). A researcher cannot possibly know the number of all relevant studies that exist so any attempt to calculate 'effectiveness' which includes true negatives is defunct at the outset (189) and;

thirdly, it felt as if there was confusion on the terminology and methods used, such that it was not clear if researchers completely understood the purpose and results of these evaluations (328).

6.1 How this chapter addresses the research aims of this thesis

In this chapter, I seek to systematically explore the literature on the evaluation of literature search effectiveness. By the end of this chapter, I aim to determine how I will compare the differences between the conventional approach, defined in the critical literature review (see Chapter 3), and the tailored approach, set out in Chapter 4. Determining how I compare and evaluate the differences between the two approaches to literature searching will inform how I address the over-arching research aims of this thesis.

6.2 Chapter objectives and research questions

The objectives of this chapter are:

- To conduct a systematic review to identify all metrics or methods currently used by researchers to demonstrate literature search effectiveness.

- To use the findings from this systematic review to identify which methods are formative and summative and how study authors define what effectiveness means in literature searching.

To address these objectives, I identified three research questions:

1. What metrics or methods are used to calculate literature search effectiveness?

2. Which metrics or methods are used formatively or summatively?

3. How is effectiveness defined by study authors?

To address these research questions, I undertook a systematic review of the published literature. The systematic review has since been published (5). I report here the final accepted, pre-publication version of this systematic review.

6.3 Systematic review identifies six metrics and one method for assessing literature search effectiveness but no consensus on appropriate use

6.3.1 Abstract

Objectives:
To identify the metrics or methods used by researchers to determine the effectiveness of literature searching where supplementary search methods are compared to bibliographic database searching. We also aimed to determine which metrics or methods are summative or formative and how researchers defined effectiveness in their studies.

Study design and setting:
Systematic review. We searched MEDLINE and Embase to identify published studies evaluating literature search effectiveness in health or allied topics.

Results:
Fifty studies met full-text inclusion criteria. Six metrics (sensitivity, specificity, precision, accuracy, number needed to read, and yield) and one method (capture recapture) were identified.

Conclusion:

Studies evaluating effectiveness need to identify clearly the threshold at which they will define effectiveness and how the evaluation they report relates to this threshold. Studies that attempt to investigate literature search effectiveness should be informed by the reporting of confidence intervals, which aids interpretation of uncertainty around the result, and the search methods used to derive effectiveness estimates should be clearly reported and validated in studies.

6.3.2 Background

Various metrics or methods are used to calculate the effectiveness of literature searching. In the absence of definitive guidance, the decision on which metrics or methods can be used to evaluate literature search effectiveness is unclear. It is also unclear why researchers select the metrics they use to undertake effectiveness evaluations (329). Determining the effectiveness of literature searching can demonstrate the 'effect' of a process of literature searching, demonstrating the efficiency of a search filter, the reduction in studies to screen without missing relevant studies (time saving), and the benefits of one search approach over another.

In this systematic review, we seek to identify the metrics or methods used to calculate the effectiveness of literature searching in health and allied topics. We also seek to explore if the metrics or methods are used formatively or summatively (that is, do they seek to predict or to evaluate effectiveness (see Figure 10)). This study extends beyond simply documenting how the effectiveness of literature searching has been calculated to conducting a broader examination of what effectiveness means and how it might be defined.

Figure 10 Key terminology defined

> **Reference standard (s):** The reference standard is usually the best test currently available and it is the standard against which the index test is compared*.
>
> **Index test:** The test which is being evaluated*.
>
> **Formative:** A formative method or metric provides researchers with a potential estimate of literature search effectiveness whilst the process of literature searching is on-going. An example would be estimating the likely number of potentially relevant studies that a literature search might identify.
>
> **Summative:** A summative method or metric provides the researcher with data on the performance of a completed literature search. This helps to determine the effectiveness of a completed literature searching since values can only be determined when searching is completed. An example would be calculating the Number Needed to Read. This shows how many studies a researcher read to identify an includable study.

* source: Centre for Reviews and Dissemination. Systematic reviews – CRD's guidance for undertaking reviews in healthcare. York: Centre for Reviews and Dissemination, University of York; 2009 (3).

6.3.3 Methods

We followed a systematic approach to identify studies in which the calculation of literature search effectiveness was the primary objective of the study.

Research questions:

1. What metrics or methods are used to calculate literature search effectiveness?

2. Which metrics or methods are used formatively or summatively?

3. How is effectiveness defined in the studies?

Identifying studies and study data

Searching bibliographic databases

A literature search strategy was developed taking the following form: ((search terms for metrics or methods) OR (search terms for evaluation of literature searches)). This was applied to the title search field in two health-focused bibliographic databases: MEDLINE (OVID interface) and EMBASE (OVID interface). The title field was searched to identify studies in which the calculation of literature search effectiveness was the primary purpose of the study. The high prevalence of studies describing methods for literature searching, and the consequent risk of prohibitive numbers of "false hits," necessitated a strategy that placed an emphasis on search evaluation, to control the number of studies returned within resource limits for this study. Study identification was not limited by language or publication date and searches were run

from database inception (MEDLINE 1946 and Embase 1974) to February 23rd 2017. The search strategies are recorded in supplementary file one.

Study selection

After visual inspection for de-duplication in Endnote X7, all studies were independently screened at title and abstract and again at full-text by two reviewers (CC and JVC).

The following inclusion criteria were applied hierarchically:

An original study published in the peer-reviewed literature that:

1. calculated literature search effectiveness;

2. provided sufficient information to replicate the calculation; and

3. calculated effectiveness between a supplementary search method (e.g. handsearching, citation chasing, web searching, contacting study authors or trials register searching) and bibliographic database searching.

The following studies were excluded:

- studies which did not compare the effectiveness of a supplementary search method against bibliographic database searching;

- studies evaluating effectiveness of teaching literature searching (i.e. trained vs. novice literature searchers);

- studies evaluating only search filters (i.e. 'search filter (a)' was compared to 'search filter (b)');

- studies evaluating the effectiveness of tools (i.e. Google Scholar vs. Web of Science); and

- abstracts, non-English language papers, letters, reviews and incomplete studies (i.e. those which do not report effectiveness outcomes).

Data extraction

Data was extracted independently into a bespoke data extraction form by CC and checked by JVC.

The following data were extracted: study citation, reference standard index test metric(s) or method(s) to calculate effectiveness, definition of effectiveness reported in the study (i.e. threshold), and claimed advantages and disadvantages relating to

the calculation of effectiveness. Data were also extracted if search strategies for a reference or index test were reported and if methods to validate or quality appraise the reference standard or index test were reported. Furthermore, we determined if the evaluation was derived formatively (the purpose of the evaluation was to estimate) or summatively (the purpose of the evaluation was to calculate). The following terms are defined in Figure 10: reference standard, index test, summative and formative.

Figure 11 PRISMA diagram
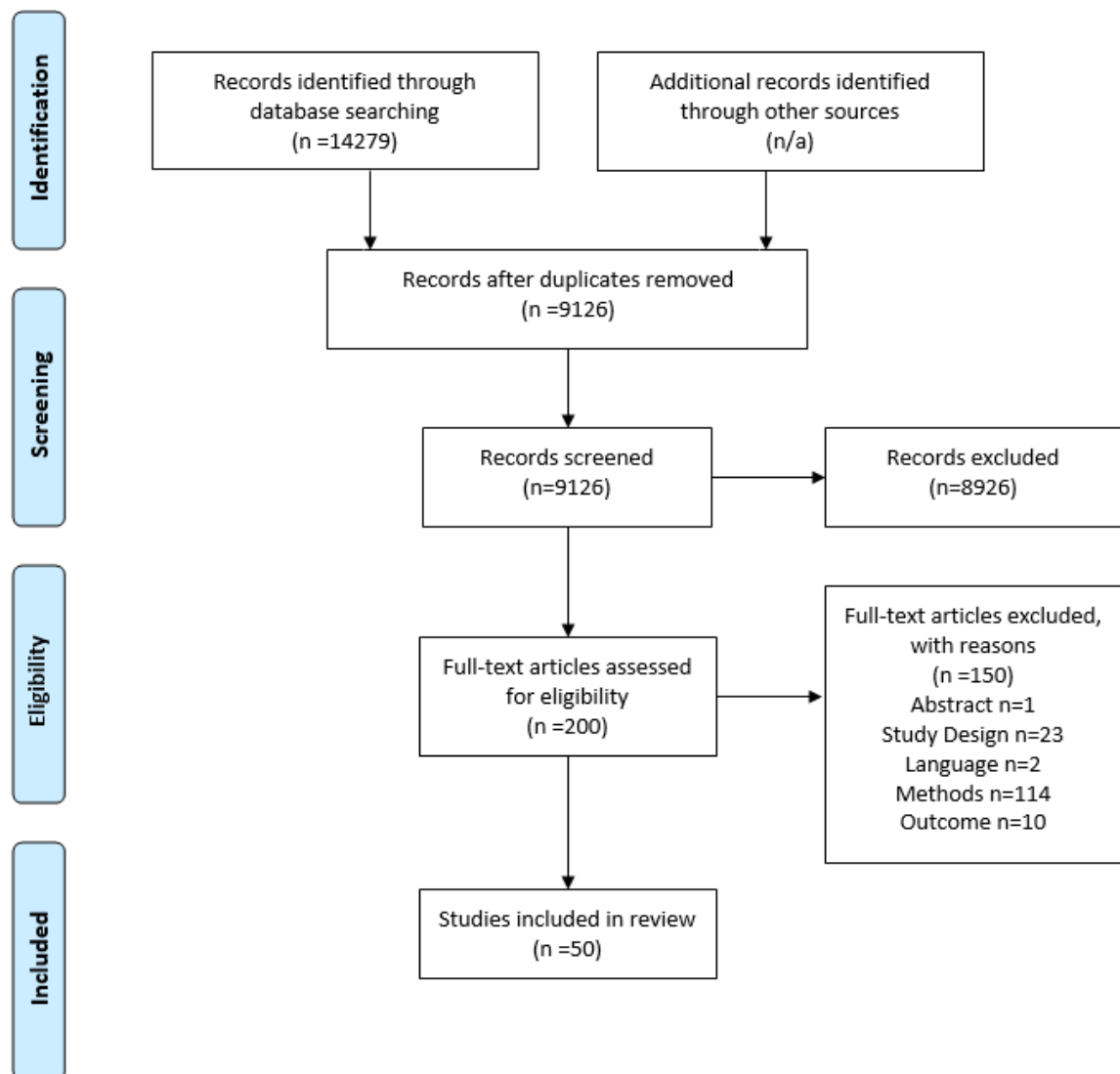
Quality assessment

The quality of studies was not appraised, since no appropriate quality appraisal tool exists, and this study focuses on mapping measures used and not on evaluating the studies in which they are reported.

Data synthesis

Data were synthesised narratively and summarised in tables to report the calculations for each method identified. The narrative synthesis of results was performed as follows: for each metric or method, the studies meeting full-text inclusion were read to identify the definition of the metric or method as reported by study authors. These definitions were extracted into Microsoft excel (2013) and read repeatedly to identify commonalties or differences between definitions in the studies. A meta-definition was drafted following this exercise which was then read ('tested') against each extracted definition to ensure all the relevant aspects of definitions from the relevant studies had been captured.

6.3.4 Results

Database searching identified 9,126 studies for title/abstract screening after de-duplication. 200 studies were screened at full-text and 50 studies met the inclusion criteria. The Preferred Reporting in Systematic Reviews and Meta-Analysis (PRISMA) flow diagram is recorded in Figure 11 (196) and studies excluded at full-text are identified in supplementary material.

Study characteristics

Of the 50 included studies (Table 13), 46 (92%) used handsearching as the reference standard. The remaining four studies used another review (n=1) or a specific combination of database searching (n=3). Validating the method or searches used to develop the reference standard was reported in 26 of 50 studies (52%) and to develop the index test in three of 50 studies (3%). Identifying a threshold to test effectiveness against was reported in 17 of 50 studies (34%). Confidence intervals were reported in 52% (26 of 50) of studies

Research Question 1 and 2: what metrics and methods are used to measure literature search effectiveness and which metrics or methods are formative or summative?

The metrics and methods used to calculate effectiveness (including specific equations) are reported in Figure 12. Six metrics and one method used to calculate and evaluate literature search effectiveness were identified and had been used either individually or in combination. These metrics and methods are summarised narratively below and the calculations are reported in Table 13.

Table 13 Included studies

		Reference Standard	Index	Metric or Method						
				Sensitivity	Specificity	Precision	Accuracy	NNR	Yield	Other
1	Adams 1994 (35)	Hand searching	Identify RCTs in MEDLINE	X		X				
2	Astin 2008 (330)	Hand searching	Identify diagnostic studies in MEDLINE	X	X	X				
3	Bachmann 2002 (331)	Hand searching	Identify diagnostic studies in MEDLINE	X		X		X		
4	Blanc 2015 (332)	PubMed searches	Web-searching						X	
5	Cathey 2006 (333)	Hand searching	Searching EMBASE	X						
6	Dumbridge 2000 (334)	Hand searching	Identify RCTs in MEDLINE	X	X	X				
7	Geersing 2012 (83)	Hand searching	Identify prognostic and diagnostic studies in MEDLINE	X	X			X		
8	Gehanno 2009 (335)	Hand searching and PubMed searches	Identify studies in MEDLINE	X		X		X		
9	Glanville 2012 (104)	Hand searching	Electronic searching for diagnostic studies						X	Time
10	Glanville 2014 (294)	Searching CDSR/ MEDLINE	Searching trials registries	X		X			X	
11	Haynes 1994 (52)	Hand searching	Identify studies in MEDLINE	X	X	X	X			
12	Haynes 2004 (336)	Hand searching	Identify studies in MEDLINE	X	X	X	X			
13	Haynes 2005[a] (337)	Hand searching	Identify studies in MEDLINE	X	X	X	X			
14	Haynes 2005[b] (338)	Hand searching	Identify studies in EMBASE	X	X	X	X			
15	Hilderbrand 2014 (339)	Hand searching	Identify studies in PubMed, MEDLINE and EMBASE	X	X	X	X			

		Reference Standard	Index	Metric or Method						
				Sensitivity	Specificity	Precision	Accuracy	NNR	Yield	Other
16	Holland 2005 (340)	Hand searching	Identify studies in EMBASE	X	X	X	X			
17	Hopewell 2002 (231)	Hand searching	Identify RCTs in MEDLINE						X	
18	Ingui 2001 (341)	Hand searching and other reviews	Identify studies in MEDLINE	X	X	x				
19	Jenuwine 2004 (342)	Hand searching	Identify studies in PubMed	X	X					
20	Kassai 2006 (343)	Hand searching	Identify diagnostic studies	X						Capture-recapture
21	Layton 2016 (344)	Hand searching	Identify studies in MEDLINE	X		X		X		
22	Linder 2015 (345)	Database searching	Citation chasing	X		X				
23	Marson 1996 (232)	Hand searching	Identify RCTs in MEDLINE	X		X				
24	Mattoli 2012 (301)	Hand searching	Identify studies in PubMed					X		
25	McKibbon 2006 (346)	Hand searching	Identify studies in PsycINFO	X	X	X	X			
26	McKibbon 2009 (347)	Hand searching	38 RCT search filters	X	X	X	X			
27	McKinlay 2006 (348)	Hand searching	Identify cost and economic studies in EMBASE	X	X	X	X			
28	Montori 2004 (349)	Hand searching	Identify SRs in MEDLINE	X	X	X				
29	Nasser 2006 (350)	Hand searching	Identify RCTs	X						
30	Rogerson 2015 (351)	Hand searching	Identify diagnostic studies	X	X	X		X		
31	Spoor 1996 (352)	Hand searching	Identify RCTs in MEDLINE							Capture recapture
32	Taljaard 2010 (353)	Hand searching	Identify RCTs in MEDLINE	X	X	X				

		Reference Standard	Index	Metric or Method						
				Sensitivity	Specificity	Precision	Accuracy	NNR	Yield	Other
33	Ugolini 2010 (354)	Hand searching	Identify studies in MEDLINE	X	X	X	X			
34	van de Glind 2012 (355)	Hand searching	Identify studies in MEDLINE	X	X	X	X	X		
35	Walters 2005 (356)	Hand searching	Identification of qualitative studies in EMBASE	X	X	X	X			
36	Watson 1999 (53)	Hand searching	Identify studies in MEDLINE and PsycINFO	X		X				
37	Wilczynski 1993 (357)	Hand searching	Identify studies in MEDLINE	X	X	X				
38	Wilczynski 1994 (358)	Hand searching	Identify studies in MEDLINE	X	X	X				
39	Wilczynski 2003 (359)	Hand searching	Identify studies in MEDLINE	X	X	X	X			
40	Wilczynski 2004 (360)	Hand searching	Identify prognostic studies in MEDLINE	X	X	X	X			
41	Wilczynski-2004 (361)	Hand searching	Identify studies in MEDLINE	X	X	X				
42	Wilczynski 2005 (362)	Hand searching	Identify studies in EMBASE	X	X	X	X			
43	Wilczynski 2006 (363)	Hand searching	Identify studies in MEDLINE	X	X	X	X			
44	Wilczynski 2007 (364)	Hand searching	Identify SRs in EMBASE	X	X	X	X			
45	Wilczynski 2010 (365)	Hand searching	Identify studies in MEDLINE	X	X	X	X			
46	Wilczynski 2011 (366)	Hand searching	Effect of NOT'ing content from search strategies	X	X	X		X		
47	Wong 2003 (367)	Hand searching	Identify studies in MEDLINE	X	X	X	X			
48	Wong 2004 (368)	Hand searching	Identify studies in MEDLINE	X	X	X	X			
49	Wong 2006 (369)	Hand searching	Identify studies in CINAHL	X	X	X	X			

		Reference Standard	Index	Metric or Method						
				Sensitivity	Specificity	Precision	Accuracy	NNR	Yield	Other
50	Wong 2006 (370)	Hand searching	Identify studies in EMBASE	X	X	X	X			
	TOTAL			45	34	40	22	8	4	3

Figure 12 Schematic of key metric and methods to evaluate literature search effectiveness and their respective calculations

Index Test (e.g. database search)	Reference Standard (e.g. handsearch)	
	Article meets criteria (relevant)	Article does not meet criteria (not relevant)
Articles identified	a (true positives)	b (false positives)
Articles not identified	c (false negatives)	d (true negatives)

Sensitivity (or recall)
The proportion of studies correctly identified as relevant, relative to the total number of relevant studies that may exist

$$\frac{\text{eligible articles retrieved}}{\text{total number of eligible articles}} \times 100$$

$a / (a + c)$

Specificity
The number of irrelevant studies excluded or not identified by the literature search strategy

$$\frac{\text{ineligible articles correctly not retrieved}}{\text{total number of ineligible articles}} \times 100$$

$d / (b + d)$

Precision
The proportion of retrieved articles that are eligible

$$\frac{\text{eligible articles retrieved}}{\text{total number of articles retrieved}} \times 100$$

$a / (a + b)$

Accuracy
The proportion of all studies correctly identified and correctly not identified compared to all articles

$$\frac{\text{total number of articles correctly retrieved or correctly not retrieved}}{\text{all articles}} \times 100$$

$(a + d) / (a + b + c + d)$

Number Needed to Read (NNR)
The number of studies a researcher must read to identify a relevant study

$$\frac{1}{\text{precision}}$$

$(a + b) / a$

Yield
The number of studies identified by a literature search method

total number of articles retrieved

$a + b$

Capture recapture/ Population Estimate
Provides an estimate of the 'population' of potentially relevant studies that might meet inclusion criteria

$$\text{number of articles by search method A} \times \frac{\text{number of articles by search method B}}{\text{articles appearing in both search method A and B}}$$

$M (m/n)^*$

* use $[((M + 1)(n + 1))/(m + 1) – 1]$ for small samples

Six Metrics: summative

Sensitivity: 45/50 (90%) studies identified (35, 52, 53, 83, 232, 294, 330, 331, 333-351, 353-370).

Sensitivity refers to the proportion of studies correctly identified as relevant, relative to the total number of relevant studies that may exist. All 45 studies evaluating sensitivity used the same calculation to determine a value, although the calculations are reported differently according to the type of study in which they are used (Table 13). Sensitivity is also referred to as: Recall (335, 345, 367) or relative recall .

Specificity: 34/50 (68%) studies identified (52, 83, 330, 334, 336-342, 346-349, 351, 353-370).

Specificity refers to the number of irrelevant studies excluded or not identified by the literature search strategy. All 34 studies evaluating specificity used the same metric to determine a value (Table 13).

Precision: 40/50 (80%) studies identified (35, 52, 53, 232, 294, 330, 331, 334-341, 344-349, 351, 353-370).

Precision refers to the number of relevant studies identified by a literature search. All 40 studies used the same metric to determine a value (Table 13). Precision was also referred to as: Positive predictive value (or PPV (330, 341)).

Accuracy: 22/50 (44%) studies identified (52, 336-340, 346-348, 354-356, 359, 360, 362, 364, 365, 367-370).

Accuracy refers to the proportion of all studies correctly identified compared to the number of non-relevant studies. All 22 studies used the same metric to determine a value (Table 13).

Number Needed to Read (NNR): 8/50 (16%) studies identified (83, 301, 331, 335, 344, 351, 355, 366).

 NNR is defined as the number of studies a researcher has to read to identify a relevant study. All 7 studies used the same metric to determine a value (Table 13). NNR was also referred to as: Number Needed to Search (351).

Yield (summative): 4/50 (8%) studies identified (104, 231, 294, 332).

Yield refers to the number of studies identified by a literature search method. All 4 studies interpreted yield in the same way (Table 13).

Yield was often not stipulated as a metric to evaluate effectiveness but rather the yield of results from one search was directly compared with another and an assessment of effectiveness was therefore presented.

One Method: formative

Capture-Recapture (Population Estimate): 2/50 (4%) studies identified (343, 352).

Capture-Recapture (or capture mark recapture) is a formative method which provides an estimate of the 'population' of potentially relevant studies that might meet inclusion criteria (Figure 12).

Combinations of the above methods were commonly used. These combinations are summarised in Table 13.

Research Question 3: how is effectiveness defined in the studies?

None of the studies included in this review explicitly defined effectiveness or clearly reported what the threshold (or cut-off) was for an "effective" result in the context of their evaluation. The use of thresholds to define effectiveness were reported in 34% (17 of 50) of the studies but thresholds were commonly used to report values for inclusion of search terms into search filters (i.e. terms of min. 50% sensitivity were included), rather than as guides to interpreting the operating characteristics of the index or reference test. No study was identified that established a threshold prospectively and tested against this.

6.3.5 Discussion

Six metrics and one method to calculate literature search effectiveness were identified in this study. In the absence of definitive guidance, the decision on which of the metrics or methods identified in this study should be used to calculate effectiveness will continue to be determined by what researchers aim to achieve, demonstrate or explore. It is unclear how researchers selected their methods to calculate effectiveness (329).

Formative methods

Capture Re-capture was the only formative method identified and it can be used to estimate the potential number of studies to be identified from the outset of a review. This has plausible utility for allocating resources and searching time, as well as planning time to screen the number of studies identified. The Capture Re-capture method has, however, been criticised by Sampson et. al given that issues of sample independence have not been adequately explored (190).

Summative methods

The summative methods all have specific purposes when used alone: sensitivity aims to demonstrate the comprehensiveness of a literature search and NNR demonstrates the screening-rate required to identity relevant studies, for instance. When these summative methods are used in combination, researchers are able to report on effectiveness (e.g. sensitivity (371)) and efficiency (e.g. precision and NNR ((371-373)).

Handsearching: the 'gold standard' search method for effectiveness evaluation?

In the review, 92% of included studies used handsearching to develop their reference standard, a finding similar to a review by Jenkins (328). Handsearching aims to ensure the complete identification of studies or publication types that are not routinely indexed in, or identified by, searches of bibliographic databases, including recently published studies (3, 6). Whilst studies show that handsearching will identify studies missed by database searching (35, 208, 231, 242, 291, 296, 300), they also show that studies can be missed by handsearching (35, 208, 231, 242, 291, 296, 300), that handsearching offers low precision (35, 242) and that it is costly in terms of time (277, 374). This raises some potentially troubling questions on the suitability of handsearching as a reference standard (6, 354, 374).

Sampson et al. propose an alternative to handsearching, namely the use of relative recall (277). Sampson et al. define relative recall as 'the proportion that any specific system retrieves of the total or pooled relevant documents retrieved by all systems considered to be working as a composite (277).' Sampson et al's approach is a composite approach, which uses a combined set of studies as a surrogate for a reference standard and, as such, this study did not meet the inclusion criteria for this study. The disadvantages of Sampson et al's method are similar to those of

handsearching: that the reference set becomes only as good as the searches that underpin it (277). Sampson et al's method would, however, mediate the concerns that calculating effectiveness using handsearching bears little relation to "real life" and it might make testing effectiveness easier, increasing the number of potential data sets available against which to test. Furthermore, since relative recall relies on underlying reviews, it might increase the transparency of methods, which would be of considerable benefit.

F Score

In peer review, a reviewer queried the absence of the F score (sometimes F-measure or F1 score) as a measure of literature search effectiveness in our review's findings. One study using F Score was identified in the main searches (375) but it did not meet inclusion at title/abstract since it did not report a calculation of literature search effectiveness between a supplementary search method and bibliographic database searching. Additional literature searches were undertaken In MEDLINE (OVID), Embase (OVID) and LISTA (EBSCOHost) to identify studies meeting our inclusion criteria and in reply to the reviewer's query. The search strategy and a PRISMA flow diagram are included in supplementary material. Thirty-nine studies were identified and double-screened. No studies met the inclusion criteria of the review.

The F Score aims to summarise precision and recall into one single number presenting a balanced mean between the two measures (376-378). As we demonstrate in this review, its application would appear to be limited in health and allied topics, and as a measure to examine literature search effectiveness. Whilst studies indicate that its use is common in information retrieval (376, 378), we found no evidence to support this.

Determining effectiveness:

Determining how effectiveness was defined in the studies was not straight-forward. We explore the issues we found, which are chiefly methodological, but this issue raises some challenging questions on the purpose of calculating effectiveness and what researchers learn by undertaking an analysis of literature search effectiveness.

Terminology:

The language used to calculate literature search effectiveness is unclear. The language used is typically borrowed from the evaluation of diagnostic tests (346) but

the terms have been adopted to calculate literature search effectiveness and are used inter-changeably, often inconsistently, and sometimes confusingly between studies (336). This impairs understanding not only of what is being measured and calculated, but also what is reported and what the purpose of the calculation(s) is. Adoption of a specific and consistent language to report the calculation of literature search effectiveness would improve the transparency of effectiveness evaluation. Where possible, we have attempted to codify the language used in attempt to define the key terms relevant to the purpose of evaluating literature search effectiveness (Figure 10 and Figure 12).

Reporting and validation within studies

Whilst study quality was not formally examined, the reporting of methods to develop reference standards or index tests, and the corresponding searches undertaken, was considered poor. Only 52% of studies in the reference standard group, and 6% in the index test group, reported validating the methods and/or searches used to develop their reference standard or index test. By validation, we mean that the methods of the underlying literature search (either for the reference standard or index test) were checked or validated by another researcher. Our findings here compare with, and are arguably even worse than, those observed in a study by Patrick et al., which concluded that peer review must be developed by authors to report evidence of effectiveness of their retrieval strategies (379).

Sampson et. al have proposed a method ('Inquisitio validus Index Medicus') for search validation (190), and the Peer Review of Electronic Search Strategies (PRESS Checklist) exists for the review of electronic search methods (183). A study by Hausner et al. recorded the time taken to quality appraise searches used in effectiveness evaluation as between 0.5 to 6.75 hours (380). Reporting the validation of methods used to develop reference standard or index tests, and their corresponding searches, should be a particular focus of studies seeking to calculate or estimate effectiveness of literature searching. Errors generated in producing a 'test set' will necessarily impact on the accuracy of their effectiveness estimate.

Use of thresholds

Whereas the design of studies comparing the index and reference test is self-evident, none of the studies reported a threshold beyond which they determined 'effectiveness' to have been achieved. Thirty-four percent of studies reported

effectiveness thresholds (Table 13, see supplementary material), but these studies typically indicated the threshold at which search terms were included in the search strategy, rather than a prospective indication of what constituted effectiveness for the overall retrieval strategy. Gehanno et al. usefully defined thresholds in their study (minimum sensitivity 65% and minimum precision 20%: NNR <5) and this approach is of benefit (335).

Diagnostic tests determine and report thresholds to indicate the point at which results are classified as either negative or positive (3). The prospective and clear reporting of thresholds in evaluation studies of search strategies would aid interpretation of the studies and would inform corresponding estimates of effectiveness generally, if the reporting of thresholds was clearer. Glanville et al. prospectively determined 'ideal performance' levels for search filters through discussion with the project team. Whilst these levels were not realised within the study, their evaluation of literature search effectiveness was consequently easier to understand and analyse relative to their objectives (381).

Confidence intervals

Confidence intervals were reported in 52% of studies. Confidence intervals offer the reader an estimate of certainty (and conversely of uncertainty) in connection with the estimate of effect. Confidence intervals should, in our opinion, be calculated and reported in all studies that seek to calculate search effectiveness.

Sample size

Harbour et al. reported that sample size calculations were not reported in their evaluation of search filter performance and our study shares similar conclusions (329). The number of studies included in the reference standard impacts upon the reliability of the effectiveness estimate. The reporting of sample size calculations, or alternatively why it was not considered possible to generate a reliable sample, is recommended.

Value

Effectiveness, reported in purely quantitative terms, tells researchers little about the value of the studies identified or missed, or what the effect of missing studies means (6). It is unclear what proportion of relevant studies identified represents an adequate literature search, so researchers are presently required to make their own

judgements of sensitivity (382-384). Sensitivity values do not help researchers understand this problem. It is acknowledged that no search can record 100% sensitivity (113, 385), so what does a 90% value demonstrate, other than that 10% of studies might be missing? Determining steps to identify the missing 10% (where comprehensive study identification is important to the review), or why a search was stopped, would be of benefit when reporting literature searches (386). The more pressing issue appears to be whether to revisit assumptions of the usefulness of evaluating literature searches by measuring comprehensiveness, since comprehensiveness may not be an appropriate indicator of search quality (113).

This also raises the question of what metrics or methods are most useful to record and report. Different researchers put effectiveness estimates to different purposes (331, 341, 351, 387), and it is not clear why study authors select the metrics or methods they do (329). As researchers and information specialists are being required to identify studies in new and more efficient ways, particularly in the context of abbreviated and accelerated reviews, thinking further about how effectiveness is evaluated and why, and also about what would be useful to report for other researchers, may be more important (388). Booth has called for an evaluation agenda (113). Such an agenda should be extended to include evaluating the usefulness of variables to be recorded (for instance, the time to search (6, 363) or sift is seldom recorded in studies) but it could also include different methods to capture effectiveness data (6).

Researchers may also consider how current metrics or methods may be used specifically for literature searching or making decisions on literature searching (389). A study by White et al. (published after the literature searches and screening had been completed and whilst this study was in final draft) evaluates the number needed to retrieve to justify inclusion of a database in systematic review search. This study offers 'proof of concept' testing of a metric, demonstrating that researchers can useful adapt metrics to demonstrate effectiveness, making transparent and evidence-based decisions on literature searching using data (387).

Limitations

Literature searching for this study was conducted in two bibliographic health-focused databases (MEDLINE and Embase). This limits the scope of this study to studies that evaluate literature search effectiveness in health or allied topics. Whilst it is a limitation in terms of scope, this limit was necessary to manage the work of the

review and, methodologically, the metrics or methods identified are not limited in application to health topics. The results and discussion above apply equally to other topic areas.

This study compared effectiveness calculations between supplementary search methods and bibliographic database searching since it offered a pragmatic way to limit the scope to the resources available. The studies identified in this study are, therefore, a representative, rather than comprehensive, sample of relevant studies.

Conclusions of the systematic review

The review identified 50 studies that sought to calculate the effectiveness of literature searching. Whilst all 50 studies calculated the effectiveness of literature searching, what constitutes an effective result was unclear. This leaves the question of what constitutes effectiveness in literature searching unresolved.

Studies evaluating effectiveness need to identify clearly the threshold at which they will define effectiveness and how the evaluation they report correlates to this threshold. We found that this is not yet common practice.

Studies that attempt to investigate literature search effectiveness should be informed by the reporting of confidence intervals, which aids interpretation of uncertainty within the result, and the search methods used to derive effectiveness estimates should be clearly reported and clearly validated in studies.

6.4 Chapter findings

Having reproduced the journal article above, I now situate the findings in the context of this thesis.

The objective of the systematic review was to review the published literature on the evaluation of literature search effectiveness and, in so doing, develop an understanding of how the research community measure, evaluate and understand the effectiveness of literature searching. As it relates to the thesis, the key conclusions of the systematic review were:

The combination of sensitivity, specificity and precision were the prevailing metrics used by researchers to measure and report the evaluation of literature search effectiveness in the studies identified and included in the systematic review.

Recording and reporting these metrics allows a researcher to evaluate and understand the 'effectiveness' of their literature search in the context of the number

of studies identified as relevant relative to the number of studies that may exist (sensitivity), the number of irrelevant studies excluded or not identified (specificity) and the proportion of retrieved studies identified as relevant (precision). From this, a researcher can evaluate and then generate an understanding of how 'effective' their literature search was by comparing it to a 'gold' or reference standard.

I identified little evidence that the research community have a consistent understanding of what effectiveness means in literature searching. It is also not clear what constitutes an effective literature search. I set out my learning on this point below by way of reflective critique when considering the contributions of this chapter to the thesis and how I will compare the conventional approach to the tailored approach in the case studies that follow.

6.5   Contributions of this chapter to the thesis

Authors of the studies included in the systematic review reported the language and methods of evaluation interchangeably and often inconsistently and occasionally incorrectly (7). There is further support for this finding in a review by Jenkins, cited in my systematic review, and a recent Health Technology Assessment report by Lefebvre et al., published after my systematic review (193, 328). This suggests that, if the terminology and methods vary between studies, the understanding of how to evaluate effectiveness is perhaps unclear in the minds of researchers. A contribution to knowledge arising from my systematic review is the meta-definitions (Figure 10) and the table I produced reporting definitions and calculations (Figure 12) which aims to clarify this.

It remains unclear what effectiveness in literature searching means and how it should be reported. Jenkins, in an evaluation of methodological search filters, suggests that reporting effectiveness estimates allows researchers to make an informed judgement on the optimal search filter to use (328). This is an assumption that has not been tested and, with reference to the variation of understanding in terminology and calculations identified in the systematic review and set out above, it is not clear that the research community can actually make informed decisions based on the limitations in practice and a lack of clarity between studies. This challenges Jenkins' assumption that researchers make an informed judgement.

On the basis of my systematic review, I suggest that what constitutes effectiveness or an effective result in literature search evaluation is also unclear. No studies meeting

inclusion criteria in the systematic review were identified that pre-specified an effectiveness threshold and compared their study findings to it. In fact, there was a distinct lack of clarity as to what effectiveness in literature searching meant, either generally or within the studies identified. How effectiveness was defined in included studies was specifically a research question of the systematic review (see 6.3.3) but I was unable to address it since no studies were identified that addressed this research question. Moreover, as I explore in the discussion section of the systematic review, it is not clear what constitutes a 'good' search much less an effective one. As it is acknowledged that no search can record 100% sensitivity (113, 385), what a 90% value demonstrates is unclear, other than that 10% of studies might be missing.

Taking this learning forward

I aimed to develop my understanding of this field whilst also formulating how I will evaluate effectiveness when I compare the conventional approach to the tailored approach in the two case studies that follow in Chapter 7 and Chapter 8.

This critical reflection not only allows me to situate the evaluation of my case studies in the context of current practice but also it allows me to consider refinements or developments to the problem of how to evaluate and report effectiveness evaluations of literature searching. As I set out in my introduction to this chapter, there are potential limitations to framing the effectiveness evaluation of literature searching in terms of diagnostic test accuracy and whether this is suitable or useful to researchers and stakeholders.

The systematic review provided a deeper understanding of this field than before but with many more questions. Some of these questions have been identified by other authors, namely:

Jenkins, who calls to improve the quality of search reporting (328). The discussion section of the systematic review (reported above: see 6.3.5) makes a contribution to knowledge here and the inclusion of Figure 10 with the calculations, and the meta-definitions in the results 6.3.4, respond to this gap in knowledge.

Other questions I seek to explore below, and where I set out how I measure effectiveness in the case studies that follow, namely:

Beale et al., who called for presenting the reporting of effectiveness of literature search filters in different ways and testing searchers' understandings of filter performance (192). This is explored below, situating it in my understanding that

grounding the evaluation of literature search evaluation in the terminology and methods of diagnostic test accuracy does not aid uptake of knowledge or clarity of understanding.

Booth, has called for an evaluation agenda which includes the exploration of the specific impact of missing relevant studies (113). I seek to take this forward both below and in the case studies presented in Chapter 7 and Chapter 8. I aim to explore the value of individuals studies in information retrieval and on the synthesis of studies.

Brettle et al. propose that a 'successful' strategy is one that retrieves a manageable number of references (219). This identifies a gap in the literature since the question of effectiveness is seldom viewed also in the context of producing manageable search returns – although this is crucial in practice for the conduct of systematic reviews. It is possible for a search filter (for example) to be effective but not useful if it does not reduce the screening burden for the researcher sufficiently. The idea potentially shares a similarity with the Cochrane handbook and the idea that study identification should be conceived in 'resource limits' (25). Jenkins notes that, whilst Brettle et al. call for an evaluation of the manageability of searches, this has yet to be investigated (328). No further evidence of manageability as a metric or any evidence of a method to do this was identified in the systematic review (5). Brettle et al. also question if different users have differing needs of information retrieval outputs (219). This is something I conclude myself and I wonder if the presentation of effectiveness evaluation might not be a way to develop Brettle et al's concept of 'manageability' whilst also aiming to measure literature search effectiveness in such a way as different users groups can derive different understandings from easy to understand data. I will attempt to elucidate this below and I will pick it up in Chapter 7 and 8 where I evaluate the case studies.

Evaluating effectiveness in the case studies

These calls, and the findings of my systematic review, identify an opportunity to re-consider not only how effectiveness is measured and determined but also to articulate an idea of what effectiveness means. In Table 14, I set-out the three domains I will capture and report data on to determine the effectiveness of the tailored approach when compared to the conventional approach. For each domain identified in Table 14, I set out my critical reflection below.

Table 14 Measures of effectiveness used in the case studies

Effectiveness (i)	The total number of studies identified by each search
Effectiveness (ii)	The number of studies uniquely identified by each search
Efficiency	The time taken to screen each output
Value	The contribution of any unique studies on the synthesis

Search effectiveness

The clearest initial marker when evaluating the effectiveness of literature search approaches is yield. Yield refers to the number of studies identified by a literature search method (5). Only 8% (n=4) of studies included in the systematic review reported yield and used it to inform evaluation of their literature search (104, 231, 294, 332).

Yield is a key value to report in a systematic review. PRISMA reporting guidance, for instance, requires that the number of studies identified from bibliographic database searching or supplementary search methods are reported in the PRISMA flow-chart (196). Understanding the yield of studies allows a researcher to plan and make time available to screen the number of studies identified, as I identified in the Key Stage Three of the conventional approach.

Comparing the yield between two search approaches offers an initial and intuitive indication of the effect of one search approach versus another. It is a clear way to demonstrate the 'effect' of a search, perhaps responding to Beale's call to present results in different ways and it would allow a researcher to determine the 'manageability' as identified by Brettle et al, since yield can be contextualised in determining how long it would take to sift the studies identified and therefore if the search produces a 'manageable' return. Although a straight-forward initial indication of effectiveness, a lower number of studies identified may not represent an optimal search over a higher number of studies. Just as screening a higher number of studies does not necessarily guarantee the quality of the literature search.

Effectiveness is not only the total number of studies identified by either search approach but also the total number of studies identified uniquely by either approach after the process of screening studies for inclusion is completed. The smaller number of studies identified uniquely will give a clearer sense of what each search approach offers uniquely, if there is any difference. This develops a better understanding of the total number of studies identified as a measure of effectiveness.

Search efficiency

As 'new styles' of rapid and scoping reviews emerge, and where decisions on where to search, or approaches to literature searching more generally are developed, there is an emerging need to understand how long search methods and approaches to literature searching take (6, 187).

As shown in the review of supplementary search methods (Chapter 5) (6), the time taken to undertake literature search methods, or the process of literature searching more generally, is seldom reported in studies, although this is an emerging area of study (187). If it were, it could be used to articulate time-saving in literature search approaches, or even the cost effectiveness of search approaches. Shemilt et al. for instance undertook a cost effectiveness analysis of approaches to screening studies in a systematic review (390). Such an idea has not been undertaken as it relates to literature searching, perhaps because the unit data on how long it takes to undertake individual literature search method is so poorly understood.

Efficiency represents an interesting evaluative point for evaluating literature search strategies since it can be used to highlight the differences in resource allocation between search approaches whilst identifying the suitability of using literature search methods depending on the time available for literature searching and review. The latter, in particular, will be of interest to researchers as they explore the challenges of rapid reviews, where choosing how to identify studies is an acknowledged challenge, and so understanding how long searching takes can inform decision making. Understanding the efficiency of literature search methods will also inform the idea of when to use literature search methods and when not to use literature search methods, and it will aid the planning of when to use literature search methods in a systematic review. Researchers are already considering aspects of efficiency, such as the number of databases searched or the comprehensiveness of the literature searches, as these debates lead into the idea of efficiency (92, 98, 111, 113, 164, 180, 203, 220, 223, 391-393).

This identifies a gap in current knowledge. As I found in my review of supplementary search methods (see Chapter 5), a single study by Gibson et al. reported average timing for the length of time between contacting a study author and receiving a reply (293). Knowing that the likely time involved in sending and receiving a request is 14 ± 22 days allows researchers to plan time to account for the delay between making a request and receiving data (293). Contextualising this with effectiveness data (as I

also do in Chapter 5), demonstrates not only the variation in response rate, but also that repeated attempts to contact study authors after this period are unlikely to result in a response (307).

Whilst efficiency has been a motivating factor in the design of study design literature search filters, actually measuring and evaluating efficiency of search methods has received little attention. Again, as with understanding effectiveness, it does not follow that the quickest search approach is optimal or that investing months developing a literature search strategy for a bibliographic database is a worthwhile investment of time. Incorporating effectiveness, as defined above, begins to contextualise efficiency, showing how many studies are identified against how efficient or inefficient a process of literature searching is. This would also allow a researcher to understand manageability too, contextualising an unscreened search result against a screened search result to determine the "number needed to read". This is helpful but it still does not articulate which would be an optimal search approach since a researcher can understanding nothing of the value of the studies identified.

The value of identified studies

Jenkins (328), Brettle et al. (219), Ingui and Rogers (341) note that users may have differing information needs of literature searches. Whilst the identification of all relevant studies may be important in some reviews, there is survey evidence from Wagner et al. that decision makers may be willing to accept a lower certainty in the answer to a research question for a swifter review (181). Moreover, study authors have questioned the utility of comprehensive literature searching in systematic reviews (c.f. (112)). This is a limitation in the use of diagnostic test evaluation methods for evaluating literature search effectiveness, since it assumes that comprehensive literature searching is the aim of the literature search which is not always the case. The notion of demonstrating comprehensiveness might usefully be re-developed into an understanding of value, which is articulated through the contribution of any study identified, or not identified, to the estimate of effectiveness (quantitative study designs) or synthesis of studies (qualitative study designs).

The recognition and measurement of study value reflects my earlier work (see 4.4.1) and the belief that information specialists need to move beyond simply finding more studies to demonstrating why the studies they do identify matter or are valuable studies for the evidence synthesis. I think that the wider idea of demonstrating the value of the studies identified is situated in Beale et al's call to propose new ways of

presenting search evaluation results and Booth's call for a new evaluation agenda to move the debate on comprehensive study identification forward in a pragmatic way (113, 192).

Value is not, however, an easy metric to demonstrate. There is also a concern that value could be confused or conflated with study quality. What I would aim to capture by measuring value in study identification is some way to determine which studies are important from all of those identified that meet the inclusion criteria of a systematic review. Or, thought of another way, if we repeat the review, which studies could have been missed with no impact to the findings or estimate of effect. The idea of effectiveness and efficiency could be reconsidered in this context. Capturing and reporting these data would begin to re-shape the idea of comprehensive literature searching (identify everything) to optimal literature searching (identify what matters). It would also start to re-orientate the way in which researchers identify and report data. We might see a move away from diagnostic accuracy methods to reporting methods specific to information retrieval.

6.6 Implications for practice

As I set out above, a contribution to knowledge arising from the systematic review, is that I identified a number of methodological weaknesses in the primary studies I reviewed. Whilst I have not undertaken an appraisal of study quality, the impact of these identified weaknesses is that it makes the individual studies harder to understand and the reliability of the effectiveness estimates less certain.

Researchers need to make clearer statements about what constitutes an effective result in the context of their studies to aid the understanding of their work. Setting *a priori* thresholds for effectiveness, and testing against these, would be an example of this, and an area for further study.

Researchers also need to raise the methodological standards of their evaluations, for instance: reporting peer review of the search strategies or second-checking of search approaches, so the robustness of their evaluations are clear and the effectiveness estimates they derive are clearly situated in the methods used. I set these particular quality improvements in the discussion section of the review (see 6.3.5).

I raise a broader question as to the continued use of diagnostic terminology and methods of evaluation to determine the effectiveness of literature searching. The

implication of this is illustrated in the systematic review. It demonstrates that even study authors are inconsistent in their use of terminology and evaluation.

Section One: A summary

Section One comprises Chapters 1-6, and contains the key definitions and the preparatory work needed to take forward the evaluation of the conventional approach compared to the tailored approach in the two case studies that follow in Section Two.

In **Chapter 3**, I identified and defined the conventional approach to literature searching for systematic reviews through a critical literature review of nine guidance documents and 119 supporting studies. I determined the suitability of the conventional approach which developed areas of refinement to the search process which I then examined in Chapter 4. This is the first time that a shared approach to literature searching process has been identified and acknowledged between guidance handbooks.

In **Chapter 4**, I set out the tailored approach to literature searching. Developed from the critical review of the conventional approach, and my experience as an information specialist, I explored the origin of the tailored approach, I set out what tailored literature searching is, and how it might work.

Tailored literature searching potentially draws more heavily on supplementary search methods than the conventional approach and I undertook the first review of supplementary search methods in order to address this knowledge gap, and to ensure that tailored literature searches were based on the best available evidence. This review, reported in **Chapter 5**, sets out the advantages, disadvantages and resource requirements of various supplementary search methods. This chapter was used to inform the development of tailored literature search approaches, used in the case studies reported in Section Two of the thesis, in addition to making a novel contribution to research in through publication in an open access journal.

To compare the conventional approach to the tailored approach it was necessary to understand the evidence available on determining the effectiveness of literature searching. The systematic review reported in **Chapter 6** examines the methods and metrics currently used to determine effectiveness in literature searching and I examine and I respond to calls to identify new ways to evaluate and report effectiveness

Section Two: The case studies

In this section of the thesis, I turn my attention to two case studies. In these case studies I compare a conventional approach to a tailored approach for literature searching. The two case studies relate to two systematic review projects in which I was the information specialist in the review team.

Key definitions

For ease of reference, I repeat the key definitions:

Conventional approach to literature searching: Identified and defined in the literature review (Chapter 3), the conventional approach is an eight-stage process of literature searching (Figure 1), common to nine leading guidance documents on the process of systematic reviews (Table 3).

The conventional approach begins with comprehensive searches of bibliographic databases, followed by supplementary search methods. The same configuration of search methods (databases followed by supplementary search methods) is used irrespective of the research question posed. The aim of the conventional approach is for a comprehensive and transparent identification of relevant studies, aiming to minimise the introduction of bias in study identification (7).

Tailored literature searching: Tailored literature searching aims to 'tailor' the selection of literature search methods to the research question (or information need(s)), the likely location of the evidence and the publication status of study data. It proposes that the order and extent to which literature search methods are used should be decided by the research team and the information need(s) of the review (see Chapter 4).

Tailored literature searching represents an alternative approach to the conventional approach to literature searching and it is what this thesis seeks to explore.

Methods for comparing the search approaches

To compare the approaches, I will measure the effectiveness, efficiency and value. The precise calculations (and the work supporting this) are set out in Chapter 6. I repeat here for ease of reference.

Table 15 Measures of effectiveness used in the case studies

Effectiveness (i)	The total number of studies identified by each search
Effectiveness (ii)	The number of studies uniquely identified by each search
Efficiency	The time taken to screen each output
Value	The contribution of any unique studies on the synthesis

Section Two outline

In Chapter 7, I compare a conventional approach to a tailored approach for a systematic review to inform the development of NICE public health guidance. This review assessed the performance of a tailored literature search represented as novel literature search filter.

In Chapter 8, I compare a conventional approach to a tailored approach for a Cochrane systematic review of intervention effectiveness, and qualitative evidence synthesis about people's perceptions and experiences of undertaking environmental enhancement volunteering. This review focused on assessing the differences between an approach to literature searching informed by following the Cochrane Handbook and developing a novel tailored approach based on supplementary search methods.

Developing ideas from Section One, these case studies explore the suitability of the conventional approach compared to the tailored approach for literature searching in systematic reviews. The principal lens to determine suitability will be effectiveness (as defined in Chapter 6) but suitability is also considered in the context of delivering literature searching that meets the needs of decision makers at NICE (case study one, Chapter 7) and the methodological expectations of Cochrane (case study two, Chapter 8).

7 Case study one: Developing and testing a search filter using a tailored approach

This first case study originated in the development of a public health guideline for the National Institute for Health and Care Excellence (NICE) and as part of a set of systematic reviews I was working on in 2010 on tuberculous (TB) screening in hard-to-reach populations in the United Kingdom.

The specific context of this study is reported in the paper which follows (see 7.4) and it relates to how I designed a literature search strategy to capture a hard-to-reach population group that was difficult to define. Below, I set out the broader challenge that this work represented, and I explain why it is of relevance to the research question of this thesis.

7.1 Background to the case study and relevance to this thesis

The systematic reviews and research questions relating to this study were conceived by NICE to address specific questions that would aid decision makers when writing public health guidance on TB screening in hard-to-reach population groups in the United Kingdom. It was important that the systematic reviews we produced followed NICE guidance, so that the reviews met the methodological standards required by NICE, and that decision-makers had the best decision-making tools available in the form of robust reviews of the evidence available.

NICE have produced a number of guidance documents in the form of methodological handbooks. These handbooks set out how a researcher should produce research and systematic reviews that meet the requirements of their decision makers. The handbook in use at the time of this study was the 2009 'Methods for the development of NICE public health guidance' (second edition) (184). Just like the 2013 edition, reviewed in the literature review (see Chapter 3) (394), the 2009 handbook advocates a process of literature searching which is represented as the conventional approach. This is, it focuses searching effort and time on bibliographic databases as the primary search method and it aims for a comprehensive identification of evidence.

Following a conventional approach for this review, where a broad and comprehensive literature search focused on bibliographic database, I estimated that we would have had to have screen approximately 27,500 studies. The challenge of this work was that, in order to deliver a review that was manageable within the time

frame demanded for tis review, we had to deviate from NICE guidance and therefore the conventional approach. By investing time to think through the problems we faced, especially how to define the hard-to-reach population, spending time scoping and discussing preliminary searches as a team, and more generally exploring this topic area with our topic experts, I was able to develop a search filter which reduced the number of studies to screen by approximately 17,000 studies. In reducing the number of studies to screen, it allowed more time to undertake supplementary search methods, which ultimately led to the identification of studies that we might otherwise have missed. I explore this further in the case study that follows, which also includes definitions of what is a literature search filter.

This case study represents my first test of a tailored search approach since it contains the basic structure of the tailored search approach and it seeks to challenge the existing guidance to produce a pragmatic literature search that meets the needs of NICE and its decision-makers. It also represents my belief that exploring the approach to literature searching in detail and as a team led by the information professional can have benefits to the overall process of study identification in systematic reviews. This is another important part of tailored literature searching.

By setting out a comparison between the conventional approach and a tailored approach, this case study directly addresses the research question of this thesis. It also builds on the other chapters of my thesis, namely: the identification of the conventional approach (Chapter 3); a test of the tailored approach (Chapter 4) the literature review of supplementary search methods (Chapter 5); and I continue to develop the approaches to measuring effectiveness that I set out in Chapter 6.

7.2   Chapter objective and research question

The chapter objective is:

> to report a case study in which I compare a tailored approach to literature searching to a conventional approach and to examine the findings of this case study in context of the research question of this thesis.

To address this objective, I identified the following research question for this chapter:

How does the tailored literature search method compare to the conventional approach in practice?

To address this, I set out the final accepted, pre-publication version of this case study below.

7.3 A population search filter for hard-to-reach populations increased search efficiency for a systematic review

7.3.1 Abstract

Objectives: This article discusses how hard-to-reach population groups were conceptualized into a search filter. The objectives of this article were to (1) discuss how the authors designed a multi-stranded population search filter and (2) retrospectively test the effectiveness of the search filter in capturing all relevant populations (e.g. homeless people, immigrants, substance misusers) in a public health systematic review.

Study Design and Setting: Systematic and retrospective analysis via a case study. Retrospective analysis of the search filter was conducted by comparing the MEDLINE search results retrieved without using the search filter against those retrieved with the search filter. A total of 5,465 additional results from the unfiltered search were screened to the same criteria as the filtered search.

Results: No additional populations were identified in the unfiltered sample. The search filter reduced the volume of MEDLINE hits to screen by 64%, with no impact on inclusion of populations.

Conclusions: The results demonstrate the effectiveness of the filter in capturing all relevant UK populations for the review. This suggests that well-planned search filters can be written for reviews that analyse imprecisely defined population groups. This filter could be used in topic areas of associated comorbidities, for rapid clinical searches, or for investigating hard-to-reach populations.

Keywords: Tuberculosis; Equity; Disadvantaged; Vulnerable; Hard to reach; Search filter; Systematic review methodology; Literature searching; Information retrieval; Information science

7.4 Introduction

This case study originated in a systematic review of qualitative evidence on the barriers to tuberculosis screening (395). The review's purpose was to inform guidance on identifying and managing tuberculosis among hard-to-reach groups (396), which presented some challenges to the review team. Chiefly, how should a population group, which, by its very nature, is hard to define, be turned into a concept

that could be sensitively reviewed and, at the same time, not become so broad, as to become diffuse?

This article will explore how the review team responded to this challenge, explaining why and how a multi-stranded population search filter was constructed. The article will also demonstrate how this filter has been tested and validated, as well as analysing its strengths and weaknesses.

The population search filter (reported in volume two of this thesis) discussed in this article can be used either

1. in the form presented here (e.g. as a filter for rapid clinical enquires on hard-to-reach populations) or in analogous topic areas (such as other infectious diseases associated with similar populations); or

2. as a basis for further research, in which the remit of the search requires a detailed analysis of the population groups discussed here.

The review discussed in this article was commissioned by the UK's National Institute for Health and Care Excellence (NICE) to inform the development of public health guidance in England (396). The searches, data extraction, and other methods used to compile the review were conducted according to the second edition of the NICE public health methods manual (184).

7.5   Background

7.5.1   Definition: what does hard to reach mean?

The definition of ''hard to reach'' was modified during the review process as the evidence was searched, quality appraised, and synthesized. The completed review (395) used a more detailed definition than the original scope setting out the parameters of the review (397), reflecting the iterative nature of the process and the work that this search filter engendered. The published NICE guidance defined the hard-to-reach population in the United Kingdom as ''adults, young people and children from any ethnic background, regardless of migration status. They are 'hard-to-reach' if their social circumstances, language, culture or lifestyle (or those of their parents or carers) make it difficult to: recognise the clinical onset of TB; access diagnostic and treatment services; self-administer treatment; or attend regular appointments for clinical follow-up''. ''The main groups considered in this guidance

were people who are homeless, substance misusers, prisoners and vulnerable migrants'' (396).

Given that the review did not start from a precise definition of the population (397), the search strategy had to reflect the diversity of meanings attached to hard to reach and recognize that the population under review could be described as underserved, or hidden, difficult to locate, and difficult to engage with and treat. These distinctions are important because they each have implications for the ways in which services are organised and delivered.

7.5.2   Definition: what are search filters?

Search filters form part of a search strategy, and they are designed to retrieve specific types of results, often those reporting on a certain study type or outcome (347). Search filters are commonly used for locating reports of randomized controlled trials (RCTs), where the concept of an RCT is well understood and a shared definition is held by those involved in the process (347). There are certain characteristics that always make up an RCT, and a search filter can be designed to capture these essential and static components such as the fact that the trial has to be randomized and controlled. This method of information retrieval requires the author to identify their article as an RCT, the database producer to index it as an RCT, and the searcher to know how to identify an RCT. The key is to match these three stages in the process, and when the concepts are well understood, this can be done with high levels of accuracy (347, 398).

The issue in reviewing population-level interventions such as tuberculosis screening is that the naming of, and characteristics associated with, the search terms are frequently changing. It is difficult for the literature searcher to translate fluid and difficult-to-define concepts into the strict controlled vocabulary of a database and to conceptualise definitions, which are ever changing. For example, the phrase ''community-based interventions,'' although a popular phrase in the epidemiologic literature, does not have a universally agreed or single point of understanding, and so it does not match any Medical Subject Headings (MeSH) in MEDLINE or the controlled indexing vocabulary of other databases. The literature searcher has to use a variety of free-text terms and indexing vocabulary to capture all potential interpretations of this idea. Furthermore, evidence relating to ''community-based interventions'' could be drawn from a range of disciplines, including psychology, education, and sociology (399), with each one using its own terminology and sources

of information (400). It is because of these points that search filters are not commonly used in public health reviews, particularly when a priori definitions of concepts do not exist (391).

7.5.3 The structure of this article

This article will now break into two parts. Section 7.6 details why this population filter was required and how a multi-stranded population search filter was written. Section 7.7 records an evaluation of the filter's effectiveness and covers the methods used to test the filter and the results found. In Section 7.8, the article analyses the strengths and limitations of the filter.

7.6 Part 1

7.6.1 The hard-to-reach population search filter

Why use a filter?

Initially, we had proposed running the search without recourse to any population search syntax. This would have allowed the reviewers to define the population during the title and abstract screening, as it emerged from the evidence. This was attractive as it meant a clean and open search syntax could be used: simply, a search strategy combining terms for tuberculosis and qualitative methods. No population of potential interest would have been excluded at the search stage as the population was unspecified. The qualitative evidence was isolated using another search filter, this time one for retrieving studies reporting the relevant outcomes. The filter was based on current research on qualitative search filters (401, 402), and its construction does not form part of the present study as the initial testing suggested that it worked adequately.

Testing this approach produced too many results to screen in the time available for the project, and sampling of the 8,549 results returned by MEDLINE produced poor returns. It became clear that something was required to manage the volume of false-positive literature from these test searches. As no current hard-to-reach population search filter was found, we wrote one.

Construction of the search filter

Structure.

It became clear quickly that the population filter had to take a multi-stranded approach to the topic, an idea used by Glanville in the study by Whiting et al. (216). The population filter would have to capture all the population groups of potential interest, and it would simultaneously have to cover them with adequate sensitivity. This dual purpose meant the filter was unusual in that it had to be sensitive both within and between the concepts. Once a relevant population group was identified (e.g. prisoners), it was essential to list the synonyms (e.g. convicts), the settings where they might be found (e.g. prisons), and also to take into account any interlinked concepts (e.g. substance abuse). We also needed to assess how the concept was described in the literature from other relevant countries to the United Kingdom and the different terminology used in those other contexts (e.g. penitentiary). The search filter had to achieve adequate sensitivity (both ensuring no populations were missed and adequately covering the groups it had found), while still retrieving a workable volume. There would be a point of diminishing returns in which the search filter became so sensitive, and the results so numerous, that it would have been quicker to revert to the original plan of searching without any population filter at all.

The population filter was structured using four strands:

1. Drawing together the high-level concepts surrounding the population (e.g. hard to reach, difficult to locate);

2. Systematically listing the groups that were known to be likely to face lifestyle or social issues associated with tuberculosis in the United Kingdom (e.g. homeless people);

3. Defining relevant synonyms pertaining to the known groups and their circumstances; and

4. Analysing the social factors that could hinder access to tuberculosis services. The purpose of this being to draw in literature that was not specifically about named groups but that might include settings causally linked to the topic (e.g. poverty and poor housing).

Constructing the filter: iterative searching and scoping.

Intensive iterative scoping searches were conducted with a view to locating likely population groups. Importantly, we operationalized these searches not on a topic level (i.e. by only looking at the tuberculosis literature) but across a broad subject base, taking a view on all associated comorbidities in the biomedical field, in which the hard-to-reach population might be referenced. The population content was more important than the subject context at this stage of the process.

This approach led to a variety of literature and produced a broad list of terms, as well as suggesting various controlled indexing terms. Once identified, a new population group was tagged for follow-up and then searched as a term in its own right, until saturation had been achieved. Saturation in this case being a subjective measure as to whether any further synonyms or variants of the new search term could be identified.

The wide range of sources outside the primary topic area was important to develop a comprehensive list of free-text terms, which included not only American and British variations in spelling but also differences in the way the topics were described (e.g. penitentiaries and prisons). Drawing from the controlled indexing thesauri of various databases beyond MEDLINE (e.g. National Criminal Justice Reference Service) was of particular use in capturing terminology that is not usually used in medical databases. Concepts taken from other sources were then tested in MEDLINE to ensure saturation of each concept.

Constructing the filter: snowballing.

The next step was to locate any articles that had actually used the term ''hard to reach'' and to examine how these had been handled by the indexers (e.g. (403)). This helped to compile a list of synonyms that was incorporated into the search filter in numerous combinations, for example, ''hard'' and ''difficult'' were combined with ''reach,'' ''locate,'' ''find,'' ''treat,'' and ''engage.'' Once a relevant article had been identified, it was thoroughly examined to extract maximum benefit from it. The purpose of this exercise was not necessarily to identify articles relevant to the review but to find includable population groups, synonyms, or associated themes relating to the population. For example, the report by Flanagan and Hancock (403) contains the target phrase ''hard to reach'' in its title, and it was a productive source of information, although it was not directly concerned with tuberculosis or infectious

diseases. The MEDLINE record shows how the term "hard to reach" has been translated by the indexers into MeSH terms, and we then searched for other articles with these subject headings. Similarly, Dixon-Woods et al. (404) mentioned neither tuberculosis nor "hard to reach," but the article deals with the central concept of our review (barriers to accessing healthcare services), and it was useful for identifying the MeSH term "vulnerable populations." These articles were also subjected to backward citation chasing (where we looked at the references they cited) and forward chasing (where we used Web of Knowledge to find later articles citing the ones we already knew were useful), which helped locate further terminology.

Constructing the filter: tuberculosis epidemiology.

It was important to draw on the evidence relating to tuberculosis itself. We undertook several scoping searches, which identified social factors that could increase the risk of tuberculosis, and then, we were able to investigate free text and MeSH descriptors for them. We started with the broad patterns in the UK epidemiology (405) and followed this up by looking in more detail at specific groups (406). From these studies, we were able to identify a core set of risk factors, including homelessness, substance misuse, and imprisonment (407).

Constructing the filter: contact with experts.

A multidisciplinary team drawing on information specialists, systematic reviewers, and experts in tuberculosis designed, tested, and quality assured the filter. The combination of searching expertise and subject-specific knowledge has been shown to be an effective approach to strategy development (146). The topic experts were invaluable at identifying issues within the tuberculosis field such as outmoded terms, which had recently fallen from usage but were still required in the search, given the date parameters of the review (1990-current). The experts were also essential in identifying new terms that were emerging at conferences but that had not made it firmly into the literature. For example, the phrase "non-UK born" was a different way of describing migrant populations that was gaining currency before the search (408).

7.7 Part 2

7.7.1 Testing the search filter

The obvious concern with any literature search is missing includable articles, and this search had the added concern of missing any includable populations. The filter was

extensively piloted in MEDLINE before being used in the original review to check if it was retrieving articles potentially eligible for inclusion. The article now moves to the retrospective analysis that was developed to verify that all includable populations had been captured. Did the filtered search miss any population groups that would have been identified by screening the unfiltered search results?

7.7.2   Methods

The original unfiltered search yielded 8,549 results in MEDLINE (OVID interface). The population search filter reduced the number of results screened for the review (395) down to 3084, a difference of 5,465. The purpose of this retrospective analysis was to investigate whether any relevant populations were missed by not screening these 5,465 MEDLINE references.

The original MEDLINE strategy from the review was rerun both with and without the filter. The 5,465 references were identified and entered into a new database, using Reference Manager software (version 12; Thomson Reuters, New York, NY). This identified that 10 of the articles were duplicates, meaning that the analysis was conducted on a set of 5,455 items. The articles were then screened according to the criteria that had been used in the review (395), with an additional question added at the end in this case study, which asked, ''Does the study include data from any hard-to-reach group that has not already been included in the search filter?'' (See appendix (volume 2 of this thesis) for full searches and criteria)).

The first 100 articles were extracted and independently screened by two reviewers. There were 100% agreement on the final decision whether to include or exclude the article (k 5 1) and 79% agreement on the codes used to exclude articles. The discrepancies were discussed and a consensus reached on the coding decisions. This led to several minor amendments to the screening criteria such as changing the order of the questions. A second batch of 100 articles was then chosen for double screening against the revised criteria, and this time, there were 100% agreement on the final decision (k=1) and 98% agreement on the coding used. The references were then split in half (one database of 2,727 and one of 2,728), and two reviewers single screened one of these each.

7.7.3 Results

The screening did not identify any additional population groups that should have been included in the search filter (Table 16).

The population filter achieved a 64% reduction in volume of studies to be screened for the review without missing any population groups, which would have been includable in the review. In practice, this means that the review located the same relevant studies from screening 3,084 as opposed to screening 8,549 MEDLINE hits. Extrapolating the filterless MEDLINE figure to the other 16 databases used in the review suggests that the reviewers would have sifted approximately 27,500 instead of the 9,800 articles they actually screened. The reviewers who worked on the review were able to screen at a rate of about 500 articles a day. Screening the additional 17,700 articles at this rate suggests that the filter saved approximately 35 days or 7 working weeks of time. The time and resources available for the project could not have accommodated this additional workload.

Table 16 Exclusion and inclusion of the 5,455 articles

Exclusion criteria		Percentage of 5,455
1	Not in English	22.8
2	Published before 1990	0
3	Non-OECD country	18
4	Nonqualitative study	57.86
5	Not a tuberculosis service	0.44
6	Not a hard-to-reach group	0.59
7	Hard-to-reach group already in the filter	0.31
8	Includable populations that should have been in the filter	0

Abbreviation: OECD, Organisation for Economic Co-operation and Development.

The sifting did identify a number of additional population groups, including funeral directors (409), farmworkers (410), Russian factory workers (411), air force dentists (412), and a church gospel choir (413). These rather eclectic groups did not meet the definition of ''hard to reach'' adopted in the final review, and these articles would have been excluded by the reviewers. The search filter did not retrieve articles on every population mentioned in the tuberculosis literature, but it did not need to be modified to incorporate these groups as they were not relevant to the definition of hard-to-reach populations adopted for the UK context of the guidance development.

Calculating a precise sensitivity and specificity score to validate the search filter, as is often done with methodological filters, has not been done in this case as it may be misleading. The aim of this study was to test whether any population groups had been missed by the filtered search, and so the sensitivity, in these terms, was 100%, and precision was unaltered as no new populations emerged.

7.8 Discussion

7.8.1 Strengths

The strength of the filter was that it named the populations that were already known and simultaneously searched for those that had not yet been identified but that might have been of value to the review. The reviewers were unable to give a fixed definition of ''hard to reach'' until they had reviewed the evidence and the search filter coped with this inherent fluidity. The definition of ''hard to reach'' became much more robust in the final review because it emerged from the evidence, and the reviewers were not forced to list explicitly the populations that were, and were not, includable at the beginning of the project.

The filter facilitated this fluidity because it adopted a multi-stranded approach that incorporated themes and situations associated with lifestyle and culture (e.g. poverty), rather than conceptualizing population merely as the names for specific groups of people (e.g. homeless people). Excluding themes such as poor-quality housing (not being a population but being a risk factor for tuberculosis) would have weakened the semantic nature of the search and missed those populations of interest that had not yet been identified.

The reduction in the screening workload (which we have estimated to be up to 35 days) allowed time for further gray literature searching and screening. Booth (113) suggested that searchers should conduct a preliminary mapping of a field, target their resources on the productive areas, and make full use of non-database techniques such as citation searching and snowballing. The search filter meant that the reviewers did not have to spend time sifting through 5,455 MEDLINE search results, which, we have subsequently proven, did not contain any additional at-risk groups.

The time saved on screening MEDLINE results was available for other activities such as contacting experts, which had a noticeable impact on the final review. The review (395) included 25 studies: 19 of these were journal articles, with 18 identified by MEDLINE and 1 by CINAHL. The remaining six studies were reports, a thesis, and other pieces of gray literature that were not indexed on the bibliographic databases most often used in reviews. The time saved on MEDLINE searching was invested in other techniques, which in turn retrieved around a quarter of the studies included in the review. The need for optimal searching approaches is particularly important in

topics in which the range of databases and potential sources to be considered is much wider than those traditionally used in systematic reviews (400).

7.8.2 Limitations

The retrospective case study presented here has only considered the MEDLINE results, and a fuller analysis would require unfiltered searches on all 16 of the other databases used in the review. No funding was available to screen the 17,700 articles this could have required. It is unlikely that a wider study would reach different conclusions given that MEDLINE retrieved 18 of the 19 journal articles included in the review. MEDLINE was believed to be a sufficiently robust benchmark to judge the impact of the filter.

Search filters should not be unquestioningly copied from one review to another as they need to be suitable to the specific research questions being asked. The search filter discussed here was designed for the OVID interface to MEDLINE, and it would require some amendments before being used in other interfaces as search results have been shown to differ according to the database provider (109). The filter was deliberately constructed for English language databases, and there is, naturally, the potential to explore terms in other languages. The review focused on evidence relevant to producing guidance for England, as required by the NICE public health methods manual (184), and so articles on Hispanic Americans and aboriginal Canadians were excluded (414, 415), which suggests that the filter would have to be reviewed to check its appropriateness to the culture or country under investigation. The search filter should also be crosschecked against the epidemiology and other risk factors when applied in other disease areas, for example, reviews on barriers to human immunodeficiency virus or hepatitis testing would benefit from the filter, but they might emphasize different aspects of lifestyle and culture.

This search filter will need to be reviewed over time, as new terminology is developed in the literature; MeSH headings are constantly under review and search interfaces change.

7.9 Conclusions

The retrospective analysis has shown that the decision to design a search filter was justified on the basis that it saw a 64% reduction in the number of MEDLINE results for screening, without any impact on the conclusions of the review. The time saved

on screening had a demonstrable impact on the quality of the review as it freed time to explore other sources of evidence and resulted in other literature being captured.

This study has demonstrated that it is possible to search for undefined and unknown populations. A search was constructed for "hard-to-reach" populations without knowing which groups were actually captured by that definition. The search approach enabled the definition of relevancy to emerge from the evidence, rather than demanding an a priori definition. The multi-stranded search approach combining named populations, settings, and themes is recommended for uncovering any hidden or undefined populations.

The specific filter is offered and tested for populations relating to tuberculosis, and its value is transferable to associated disease areas (such as blood borne infections) in which hard-to-reach populations are of interest.

7.10 Chapter findings

Having reproduced the journal article above, I now situate the findings in the context of this thesis.

This chapter reports the first example of a case study which compares a tailored approach to literature searching against a conventional approach and it sets out a specific information retrieval challenge in a public health review. Faced with a potentially unmanageable number of studies to screen, due to an uncertain understanding and imprecise definition of the population under review, I was required to develop an approach to literature searching that met the needs of decision makers and NICE and that produced a manageable number of studies in a comprehensive and transparent way.

The findings indicate that the literature search filter, developed following the early principles of a tailored approach, was successful, since it reduced the volume of studies to screen without missing any potentially relevant studies. The early principles of the tailored approach, in this study, were: where I took the lead as the decision maker, working with the team and topic experts to develop a multi-stranded search filter which met the requirements of stakeholders, and investing time at the start of the review to develop the approach to literature searching and match a search approach (the search filter) to the needs and timeline of the project. The time invested in scoping, and the team-work, contributed to the overall success of the filter.

Having summarised the findings of this work as they relate to the thesis and its research aims, I now focus on a critiquing the findings in the context of the relevant literature and the process of literature searching in systematic reviews. I will focus on the application of the tailored approach and the evaluation of effectiveness. This will help me address the research aims of this thesis and chapter objective two before setting out the implications to practice arising from the work reported.

7.11 Contributions of this chapter to the thesis

The aims of this thesis are to: 1) examine approaches to systematic literature searching for systematic reviews and, 2) propose and test a method of systematic literature searching for reviews that do not focus on the effectiveness of clinical interventions.

Through the first case study, this chapter demonstrates that the idea of tailored literature searching, as set out in Chapter 4, can develop alternatives to the conventional approach and with positive results. Investing time to explore the approach to literature searching led to the development of a literature search filter which identified all the relevant population groups with a lower overall number of studies to screen when compared to the conventional approach. This created more time to undertake supplementary search methods which proved valuable since six unpublished pieces of evidence were identified.

This case study demonstrates that alternatives to the conventional approach are possible; it articulates a clear role for the information professional in the team, developing the idea of the information professional as 'decision maker' in the instrumental and conceptual approaches to literature searching; that researchers should challenge guidance and work with decision makers to develop alternative approaches to literature searching, and it articulates some of the benefits to be found.

The case study reported in this chapter also articulates a test of an alternative way to demonstrate the effectiveness of the tailored approach when compared to the conventional approach. Whilst I was unable to calculate efficiency, and I was unable to demonstrate value, I responded to the call of other researchers to report effectiveness findings in clearer and more simple ways.

I now situate my learning in the relevant literature by way of critique.

7.11.1 Application of the tailored approach

The tailored approach to literature searching in this case study involved investing time at the start of a systematic reviews to scope the evidence, explore the research questions, and generate a clear understanding of the information need (4, 327). The interaction with the review team is important, since a team has a greater collective history, skills and experience to draw upon and they can work together to understand different aspects of the information need(s) (4, 131, 132, 134). This turned out to be of particular importance to the case study given the uncertainty around defining the population group under review.

I played the role of the 'decision maker' in this context (as suggested in Chapter 4), as I was responsible for developing this period of review immersion and structuring the approach to literature searching. I also developed the definition of the information need, translating this into search query and approach (327).

A challenge in the application of the tailored approach was determining how long to spend developing the tailored approach since any investment in developing an alternative approach must be balanced against the time needed to screen the higher number of studies generated by a conventional approach. It is not clear how long the period of immersion should last. The 2009 'Methods for the development of NICE public health guidance' (second edition) handbook included guidance on the length of time that the research team should budget for when planning the literature search process. The suggested timing was 3-4 weeks (184).

It is not clear how this time-frame was derived and there were no sources of reference to situate it in empirical research. The guidance on timing was not included in the 2014 revised handbook; the reasons for this are not clear.

I followed this guidance on timing in the NICE handbook to design the literature search filter I report in this case study. The period of review immersion included frequent scoping searches, discussion on relevant studies identified, and also exploration of any 'near miss' studies (studies that almost met inclusion but did not evaluate the intervention, for example). Meeting with our topic expert for a briefing on current interventions and trials, and receiving feedback from her on studies we had identified, was helpful. As was sourcing grey literature from the NICE call for evidence.

Investing this time, and exploring the problem as a team, I believe paid dividends. This can be seen in the results of the study (specifically, the lower number of studies to screen) and in the development of team's understanding of the population prior to screening, as is set out in the study (50). By developing the population filter to 'filter in' population groups and 'filter out' populations or studies that were not of interest (328), we were able to budget more time for supplementary search methods which identified six studies uniquely (50).

The tailored approach achieved a 64% reduction in the volume of studies to be screened in the review without missing any population groups which should have been included in the original review. I estimate that this saved approximately 35 days based on screening 500 studies a day.

Arguments are also made in the study that the investment of time to develop the tailored approach were of benefit to understanding the population and evidence we were tasked to review. The team discussion and immersion into the evidence for the review question developed a more cohesive understanding of the information needs of the search, helping to articulate a clearer search query (327) and improve the understanding of the screening and rigour in the review (50). This has further benefits for the review too, such as a clearer data extraction table.

The results of the evaluation demonstrate a number of contributions to knowledge, which I set out below. In practical terms, the search filter made the systematic review manageable. As I set out in Chapter 6, Brettle et al. propose that a successful search strategy is one that retrieves a manageable number of references (219). Using this case study to examine the issues raised by Brettle et al, the search filter reduced the number of studies to screen to within resource limits and it was successful since the case study demonstrates that we did not also miss any relevant populations groups based on a search of MEDLINE.

There are limitations to the application of the tailored approach and this case study which should also be considered. This approach to literature searching and the potential impact on the review meant challenging NICE methods guidance (in the form of the conventional approach) and with no guarantee that the approach would be successful. Reasonable questions were raised on how I could be certain that the search would not exclude relevant population groups. This was a particularly sensitive issue as the population under review was already considered marginalised and at risk. There was no mechanism or test to prospectively examine this challenge.

The claims for time-saving and the reduction of studies to screen reported in the study are speculative. It should also be noted that, whilst I claim that 35 days screening time were saved by applications of the tailored approach, 30 days were taken to develop the search approach, suggesting that only 5 says were saved overall. This highlights a further limitation of this case study. Precise timings of the stages of development, of searching, and of screening, were not routinely kept. I call for researchers to record the timings of individual search methods and the timing of the overall search process itself in the literature review of supplementary search methods. These timings would help inform the development of future search approaches and perhaps begin to provide an evidence-base for the 3-4 weeks that NICE formally suggested were appropriate for planning a process to literature searching.

The scope of the comparison reported in the case study is also open to criticism. As noted in the reported limitations of the case study 7.8.2, by comparing the tailored approach to a MEDLINE only search, I present an incomplete summary of the other 15 databases searched in the review. It may have been more appropriate to include a social science database (such as ASSIA) alongside MEDLINE with reference the topic and population under review however, 18 (of 19) studies identified from database searching were identified in MEDLINE (the 19th study was identified in CINAHL).

7.11.2 Measuring the comparison between search approaches

In Chapter 6, I set out how I proposed to measure the effectiveness of the conventional approach compared to the tailored approach. I repeat this again here, for ease of reference.

Table 17 Measures of effectiveness used in the case studies

Effectiveness (i)	The total number of studies identified by each search
Effectiveness (ii)	The number of studies uniquely identified by each search
Efficiency	The time taken to screen each output
Value	The contribution of any unique studies on the synthesis

I was not able to compute efficiency since the timings of the various stages of the process of literature searching were not recorded at the time of conducting the review. I acknowledge that this is a limitation of this work. I was also not able to examine the idea of value, since I did not identify any new studies missed by the original work. For this case study then, I was only able to report on search effectiveness i and ii.

In peer review, a reviewer challenged me on how I reported the differences between approaches. They requested that sensitivity, specificity, precision and accuracy be calculated and reported since, in their view, the comparison presented in the study was 'essentially a diagnostic test accuracy study.' As I set out in Chapter 6, I am not comfortable in nesting effectiveness evaluation in the terminology and methods of diagnostic test accuracy since I do not feel that this represents a clear reporting of the evaluation undertaken. Whilst I appreciate the comments of the peer reviewer, the case study, and my response, set out why it is not entirely appropriate or satisfactory to demonstrate effectiveness evaluation of search approaches in this way.

I did not calculate the sensitivity or precision since I did not feel that this aided interpretation of the findings. Moreover, whilst I call this a 'search filter' in the study, this is possibly misleading as it is unlikely that the filter would be repeatedly used without being changed, so I would reasonably expect the operating characterises of the filter to change on a case-by-case basis, rending a report on sensitivity, specificity and precision within limited context.

The reviewer's comment reflected what I set out in Chapter 6, and other searchers have called for, which is alternative ways to present effectiveness evaluations (192, 193). In my case study, reporting sensitivity and precision, for example, would not have informed understanding of the findings of the case study, since sensitivity is 100% as no new studies were identified, and precision was unaltered for the same reasons. Presenting yield (effectiveness i: see Table 17) and the number needed to read (effectiveness ii: see Table 17) made for a more accessible comparison of differences between the search approaches. As I set out above though, had timing been captured, presenting efficiency data too may have presented a closer comparison between the two approaches.

7.12 Implications for practice

Reflecting on this case study in the context of this thesis, the principle contribution to practice is that, whilst guidance to literature does exist, it demonstrates that researchers can work with research commissioners and decision makers to develop alternative approaches to literature searching that do not necessarily follow this guidance. The case study sets out the benefits of this approach.

Researchers call for investing time at the start of a systematic review to develop the understanding of the evidence base and this case study contributes to that call (208). I link spending time to prepare to search, working with the review team specifically, to the development of a search approach that reduced the number of studies to screen without missing any studies of relevance to the review and which created more time to undertake supplementary search methods.

The question of how much time to invest is uncertain since I did not (and no other authors have either) recorded the time invested overall. Researchers should record the time taken relative to tasks undertaken and make clearer links between investment of time and end benefit should be explored. On the basis of this case study, it would appear that this is not time saving but rather a re-allocation of how time is spent. This is not necessarily a negative finding.

The publication and validation in context of the search filter identifies a gap in knowledge. The contribution to knowledge is that the search filter can be used or adapted when researchers attempt to locate studies relating to this broad population group. The search filter is reported on the Information Specialists Sub-Group (ISSG) search filters resource. This is a world-leading resource for validated literature search filters.

8 Case study 2: Developing and testing a tailored approach in a Cochrane Public Health systematic review

This second case study originated in a project funded by the National Institute of Health Research School for Public Health Research. The project sought to determine the health and wellbeing impacts of participating in environmental enhancement activities (416). The project was registered with the Cochrane Public Health group and it included a Cochrane systematic review of intervention effectiveness, together with an associated qualitative evidence synthesis about people's perceptions and experiences of volunteering for environmental enhancement activities.

8.1 Background to the case study and why it is relevance to this thesis

In the case study I set out below, I compare the conventional approach, which in this case study is the approach to study identification as set out in the Cochrane Handbook, to a tailored approach, which I developed with the research team for use in this study. This case study directly addresses the research aims of the thesis.

At the start of this systematic review, it seemed likely that the relevant evidence would be largely unpublished and that it was unlikely to indexed in bibliographic databases. Furthermore, given the breadth of quantitative and qualitative research to be identified, and the lack of a consistent definition for the interventions under investigation, that searching bibliographic databases would led to us over-screen a number of studies to little advantage in terms of identifying relevant studies and that it would cause us to run out of time to undertake supplementary search methods. The tailored approach was developed, which focused on the supplementary search methods I thought would identify relevant studies and data.

By setting out a comparison between the conventional approach and a tailored approach, this case study directly addresses the research question of this thesis. It also builds on the other chapters of my thesis, namely: the identification of the conventional approach (Chapter 3); a test of the tailored approach (Chapter 4) the literature review of supplementary search methods (Chapter 5); and I continue to develop the approaches to measuring effectiveness that I set out in Chapter 6. This case study was undertaken and published three years after the case study reported in Chapter 7, so I was also able to build on and develop from this experience and findings of this work too.

8.2 Chapter objective and research question

The chapter objective is:

> to report a case study in which I compare a tailored approach to literature searching to a conventional approach and to examine the findings of this case study in context of the research question of this thesis.

To address this objective, I identified the following research question for this chapter:

How does the tailored literature search method compare to the conventional approach in practice?

To address this, I set out the final accepted, pre-publication version of this case study below.

8.3 Supplementary search methods were more effective and offered better value than bibliographic database searching: a case study from public health and environmental enhancement.

8.3.1 Abstract

Background: We undertook a systematic review to evaluate the health benefits of environmental enhancement and conservation activities. We were concerned that a conventional process of study identification, focusing on exhaustive searches of bibliographic databases as the primary search method would be ineffective, offering limited value.

The focus of this study is comparing study identification methods. We compare: (i) an approach led by searches of bibliographic databases to (ii) an approach led by supplementary search methods. We retrospectively assessed the effectiveness and value of both approaches.

Methods: 'Effectiveness' was determined by comparing: 1) the *total number of studies* identified and screened and, 2) the number of includable studies *uniquely identified* by each approach.

'Value' was determined by comparing included study quality and by using qualitative sensitivity analysis to explore the contribution of studies to the synthesis.

Results: The bibliographic databases approach identified 21,409 studies to screen and two included qualitative studies were uniquely identified. Study quality was moderate and contribution to the synthesis was minimal.

The supplementary search approach identified 453 studies to screen and nine included studies were uniquely identified. Four quantitative studies were poor quality but made a substantive contribution to the synthesis; Five studies were qualitative: three studies were good quality, one was moderate quality, and one study was excluded from the synthesis due to poor quality. All four included qualitative studies made significant contributions to the synthesis.

Conclusions: This case study found value in aligning primary methods of study identification to maximise location of relevant evidence.

Keywords: information science; literature searching; sensitivity analysis; Cochrane systematic reviews; Public health.

8.3.2 Background

With the increased interest in evidence-informed environmental policy (417), researchers have explored the suitability of applying the explicit methods of systematic review to the field of conservation research (418-423). Whilst collectively researchers agree that a systematic process to identify and review studies is of benefit, they helpfully highlight several issues. A primary concern is the appropriateness and application of a process and methodology which was originally developed to systematically review studies reporting randomised controlled trials indexed within bibliographic databases, to the systematic review of the myriad of study designs used to evaluate conservation, and other complex interventions, the results of which are widely dispersed throughout bibliographic databases and 'grey literature' (420, 422, 423).

In 2012, we began a mixed-methods systematic review to evaluate the health and wellbeing impacts for different groups of people undertaking environmental enhancement and conservation activities. We encountered issues highlighted by Pullin and Knight, Fazey et al., and Stewart et al. (420, 422, 423) as we began scoping our review, namely: a relative absence of studies using controlled or otherwise 'higher order' study designs (420, 421, 423); a difficulty in accessing primary studies to review, due to: delays in publication, limited publication, or simply no attempt to formally publish completed research (418, 419, 421, 424); and a recognition that a variety of sources would need to be searched to identify studies

(420, 424). Our project reference group (PRG[4]) validated these concerns, while anticipating that many of the studies that might address our research question would likely be found in the grey literature.

We were concerned that a conventional approach to study identification, described in the leading handbooks for the process of systematic review (3, 25) that focuses on sensitive searches of bibliographic databases as the primary method of study identification, could yield an overwhelming number of studies to screen, with low numbers of includable studies identified, and potentially diverting time away from identification of grey literature. Facing similarly challenging searches, other researchers have explored the successful adaptation of conventional search methods to the identification of studies within disparate bodies of grey literature (88, 93, 295). Accordingly, we developed a tailored study identification protocol. The tailored study identification protocol was designed *a priori* to ensure the systematic identification of studies and minimise the introduction of bias in study selection, whilst also seeking to allocate time to supplementary study identification methods that were anticipated to offer a more productive yield of studies for inclusion than searches of bibliographic databases.

During the process of protocol development, we registered our systematic review with Cochrane's Public Health Group (425). Cochrane provides specific methodological guidance for the systematic review of intervention effectiveness. Typically, in Cochrane Reviews of interventions, studies reporting randomised controlled trials are sought (25) but, in public health reviews and/or reviews of conservation interventions such as this one, a range of study designs may be included (426). The process of study identification for Cochrane Reviews is set out in detail in chapter six of The Cochrane Handbook, 'searching for studies,' and summarised for reviews in public health topics in chapter 21, 'reviews in public health and health promotion' (25, 426). The aim of study identification within the Cochrane model is the comprehensive identification of published and unpublished studies; this is a sequential process of study identification, led by comprehensive searches of bibliographic databases and followed by searches of non-bibliographic databases sources (e.g. handsearching, searches of conferences).

---

[4] practitioners, experts in the field and academics brought together to oversee the development of the review

As Cochrane authors, we were committed to following this Cochrane process of study identification but, given the need to interpret this process within conservation science and public health, and our awareness of the need for more time and effort to identify grey literature than is typical for a Cochrane Review, we decided to employ a hybrid approach. This augmented the Cochrane method for study identification (with bibliographic database searches as its primary method of study identification) with a tailored study identification protocol (with supplementary searches as its primary method of study identification and a focus on extensive grey literature searches). This adaptation provided us with the opportunity to compare the effectiveness of the two study identification protocols.

8.3.3 Study aims

To assess the effectiveness and value of a search approach led by supplementary search methods (the tailored study identification protocol) compared to a search approach led by bibliographic databases (The Cochrane study identification protocol).

In this study, we determined 'effectiveness' by comparing (i) the *total number of studies* identified and screened and (ii) by comparing the number of included studies *uniquely identified* by each study identification protocol. We determined 'value' by comparing the study quality across included studies retrieved for each study identification protocol and by analysing the contribution of studies to the synthesis.

8.3.4 Developing the Cochrane study identification protocol and tailored study identification protocol

This section describes how we developed the Cochrane study identification protocol and the tailored study identification protocol and the methods used to measure the effectiveness of study identification and the evaluation of study quality and contribution to the synthesis of each approach.

The Cochrane study identification protocol

The Cochrane study identification protocol was developed and peer-reviewed as a required component of our overall systematic review protocol by The Cochrane Public Health Group (425).

The primary method of study identification in the Cochrane study identification protocol involved searches of 22 bibliographic databases (see Figure 13). The multi-

disciplinary nature of conservation/public health topics means that studies can be identified from diverse databases, not necessarily limited to health topics, so it is common practice to search a greater number of bibliographic databases than for clinical topics (89-91, 427). These 22 databases included: MEDLINE (Ovid), Embase (Ovid) and The Cochrane Library (Wiley interface) as well as Social Policy and Practice (Ovid), IBSS (Pro Quest) and ASSIA (Pro Quest), CAB Abstracts and Greenfile. The full list of bibliographic databases searched, and our MEDLINE search strategy, is included in the published Cochrane Review (416). The Trial Search Co-Ordinator of The Cochrane Public Health Group checked and approved our searches.

Figure 13 Databases searched

- Assia (ProQuest);
- BIOSIS (ISI);
- British Education Index (ProQuest);
- British Nursing Index (ProQuest);
- CAB Abstracts (CAB Direct);
- Campbell Collaboration;
- Cochrane Public Health Specialized Register;
- DOPHER (EPPI);
- EMBASE (Ovid);
- ERIC (ProQuest);
- Global Health (Ovid);
- GreenFILE (EBSCO);
- HMIC (Ovid);
- MEDLINE in Process (Ovid);
- MEDLINE (Ovid);
- OpenGrey;
- PsycINFO (Ovid);
- Social Policy and Practice (Ovid);
- SPORTDiscus (EBSCO);
- TRoPHI (EPPI);
- Social Services Abstracts (ProQuest);
- Sociological Abstracts (ProQuest);
- The Cochrane Library (all via Wiley Interface);
- TRIP Database; and
- Web of Science (including conference citations index) (ISI).

The tailored study identification protocol

The tailored study identification protocol included the same methods of study identification as set out in The Cochrane Handbook (and used in the Cochrane

protocol) but with a revised focus for study identification methods. We changed the primary focus of study identification from bibliographic database searching to contacting organisations and searching web-sites (see appendix (volume 2 of this thesis) for supplementary material) thereby affecting the weighting of the methods in the process of study identification as it relates to searching time. Studies evaluating the use of supplementary search methods were useful in informing this discussion (95).

The study identification protocols are outlined in Figure 14.

Figure 14 Schematic of Cochrane protocol and the Tailored protocol, showing the primary and supplementary methods of study identification, and the chronological order and investment in study identification methods.
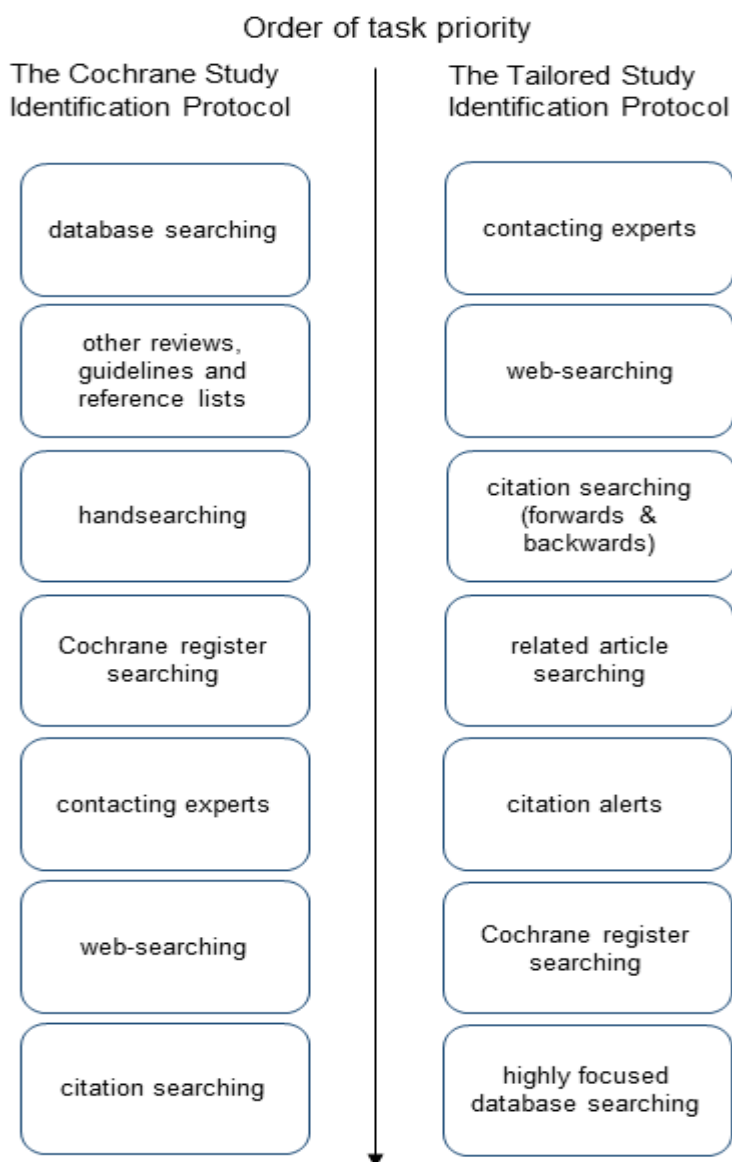
The design of the tailored study identification protocol

We sought to sensitise the team to the disparate evidence for this review before designing the tailored study identification protocol. We aimed to understand what types of studies (by design, publication type and publication status) may exist and where (and how) they could be identified. We sought to achieve this in two ways:

1. scoping searches were undertaken by the review team. Scoping searches took the following structure: ((search terms for possible interventions) and (search terms for review-relevant outcomes)). The aim was to identify candidate studies in bibliographic databases (published) and through web-searching (grey literature). The purpose of these searches was early identification of studies and organisations as well as to explore how and where potentially includable studies were being identified; and

    2. a project reference group (**PRG**) was formed, made up of a wide range of key organisations, such as: the Conservation Volunteers, Mind, Local Authorities and Groundwork. We met with the PRG at a preliminary stage in our review to hear from topic experts about the types of interventions and participants we were aiming to find/identify. This helped generate search terms and it developed our understanding of the evidence base for the review, in particular the nature of the grey literature.

Whilst the process described above was iterative and informal, it identified two key factors that ultimately informed the order of study identification methods in the tailored study identification protocol. First, the PRG advised that the types of studies that would meet our inclusion criteria were likely to be identified in the grey literature and, secondly, our scoping searches of bibliographic databases suggested that a sensitive search strategy for this review would yield approximately 20,000 studies to screen. Piloting our inclusion/exclusion criteria on these 20,000 studies suggested low specificity and precision suggesting the need to prioritise grey literature searches as a way to further refine the bibliographic database search strategy.

The tailored study identification protocol was designed therefore to concentrate searching time on grey literature searches as the primary method of study identification, specifically contacting organisations and experts in the field to identify studies, supplemented with web searching. In contrast to the Cochrane study identification protocol, we planned that bibliographic database searching would be a supplementary search method to identify published studies and reviews.

8.3.5 Methods

This is a retrospective comparison of the effectiveness and value of the two study identification protocols.

Effectiveness

Effectiveness is a term used in literature searching to describe the impact of study identification when two (or more) search approaches are compared. Whilst methods exist to calculate search effectiveness (e.g. sensitivity, specificity and precision), there is no agreed understanding as to what actually constitutes effectiveness in study identification. In this study 'effectiveness' will be determined by: 1) comparing the *total number of studies* identified and screened by each of the two study identification protocols and 2) comparing the number of included studies *uniquely identified* by each of the two study identification protocols. We are able to make this comparison since the same inclusion and exclusion criteria were used to screen studies returned by each study identification protocol.

Value and contribution

Determining effectiveness in purely quantitative terms as the number of studies identified and included in the review (as above) makes no acknowledgement of the *value* of the studies identified uniquely by each study identification protocol, nor how studies may substantively *contribute* to the synthesis or alter the conclusions of the review. In this study, we seek to link the idea of effectiveness (defined above) to the concept of study value (defined below), so that we can determine not only the effect of each study identification protocol but also the value. Value will be determined by comparing a measure of study 'quality' and by assessing the unique contribution from each study identified to the synthesis and the confidence in the findings.

Study quality

The assessment of study 'quality', using standardised and validated tools, is a key component in a systematic review (428). Quality assessment of studies included in a review examines the risk of bias in studies using quantitative study designs, and subjective interpretation in qualitative studies, and the impact on results (315), guiding the interpretation of findings (429). In this way, study quality is integral to interpreting the value of studies identified.

Study quality was assessed using the Effective Public Health Practice Project (EPHPP) tool for studies using quantitative study designs (430). Study quality was rated over six categories from being very strong (scoring the minimum of 6) up to very weak (scoring the maximum of 18). Scoring for these six categories where, 1 = strong, 2 = moderate and 3 = weak.

Cochrane's risk of bias tool was not used in the absence of any includable RCTs (426). The Wallace criteria were used to appraise qualitative studies (431).

Contribution to the synthesis (qualitative studies only)

We are not aware of any formal or standardised approach to identifying the 'contribution' of any individual study to the findings in a qualitative synthesis, although researchers describe the use of 'sensitivity analysis' (432). We developed an alternative approach and we test this idea here for the first time in an attempt to link methods for study identification to study value.

Contribution to the synthesis was evaluated by re-examining the qualitative synthesis (e.g. the documentation of the results of each of the individual stages of the qualitative synthesis) to understand which papers substantively contributed data, concepts and understanding to identification and development of the overarching themes and sub-themes. The synthesis of qualitative studies as reported in our Cochrane Review was used (416). Once each paper's contribution to the overarching and sub-themes was identified in the synthesis, we determined which studies were: 1) fundamental and necessary to the specific overarching and/or sub-theme (we term these 'key studies'), and 2) which papers merely added confirmatory validity or data richness (we term these 'additional studies'). This contributed an understanding of the relative contribution of each paper to the overall synthesis. The Confidence in the Evidence from Reviews of Qualitative Research (CERQual) approach was then used to appraise the confidence in review findings with and without the studies that were missed by each study identification protocol (433). The CERQual tool helps assess how much confidence to place in the findings from a qualitative evidence synthesis (433). In this study, we make the link between confidence and attempt to interpret this as value.

8.3.6 Results

Effectiveness

The number of studies identified and screened by each study identification protocol

The Cochrane study identification protocol resulted in the identification of 21,409 studies to screen at the title/abstract stage, compared with 453 studies identified via the tailored study identification protocol searches. At full text, 166 studies were screened from the Cochrane study identification protocol and 211 were screened from the tailored study identification protocol

The number of studies uniquely identified by each study identification protocol

Twenty-one studies met our review inclusion criteria and were included in the review (Figure 15). By study identification protocol these were:

Figure 15 Schematic of source of study identification. Key: TSIP = Tailored study identification protocol and CSIP = Cochrane study identification protocol.
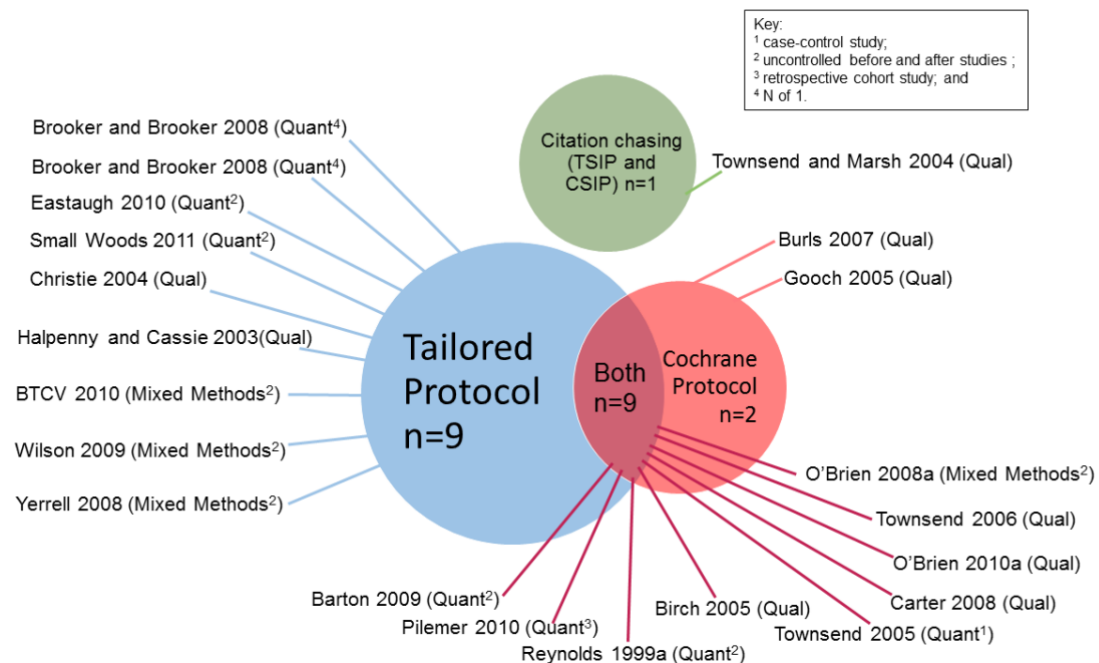
Studies identified by the Cochrane study identification protocol only: two

Two included studies were uniquely identified by the Cochrane study identification protocol through bibliographic database searching (434, 435) (Figure 15). Burls (434) was identified twice: once in Social Policy and Practice (OVID) and again in British

Nursing Index (Pro Quest). Gooch (435) was identified once, in the International Bibliography of the Social Sciences (IBSS, Pro Quest).

Studies identified by the tailored study identification protocol only: nine

Nine included studies were uniquely identified by the tailored study identification protocol (Figure 15)(436-444). These studies were uniquely identified by the tailored study identification protocol and were not indexed in any of the bibliographic databases. These studies could only have been identified by author contact or web-searching.

Study identified by citation chasing (Cochrane study identification protocol and tailored study identification protocols): one

One included study was identified uniquely by citation chasing, a method of study identification shared by both search protocols (Figure 15). Townsend et al. (445) was identified through backwards citation chasing Moore et al. which was identified by both search protocols (446).

Studies identified by both study identification protocols: nine

Nine included studies were identified by both the tailored protocol and the Cochrane protocol (Figure 15) (445, 447-454). These studies were identified by bibliographic searching in the Cochrane study identification protocol and, separately, through organisation contact and web-searching in the tailored study identification protocol.

Effectiveness summary

The tailored study identification protocol identified all studies included in our Cochrane Review with the exclusion of two studies: a study by Burls and a study by Gooch, both qualitative studies (434, 435). The tailored study identification protocol uniquely identified nine studies missed by the Cochrane study identification protocol (436-444).

8.3.7 Value

8.3.8 Study quality

Quantitative studies: The EPHPP Tool

The EPHPP tool scores study quality using a global rating summarised in three domains: Strong, Moderate and Weak (430). The tailored study identification protocol

uniquely identified seven studies using quantitative study designs and the quality was
scored weak for all (between 12-18. Table 18). Two of these seven studies were
included in our review but were excluded from the actual synthesis due to poor study
quality (primarily due to small study samples) (436, 437). No studies using
quantitative study designs were identified uniquely by the Cochrane study
identification protocol (Table 18).

Table 18: Study Quality

Study	Study Type	Identification Method	EPHPP	Wallace
Brooker and Brooker 2008* (436)	Quantitative	TSIP	Weak	
Brooker and Brooker 2008* (437)	Quantitative	TSIP	Weak	
Eastaugh 2010 (440)	Quantitative	TSIP	Weak	
Small Woods 2011a (442)	Quantitative	TSIP	Weak	
Barton 2009 (447)	Quantitative	CSIP + TSIP	Weak	
Pillemer 2010 (451)	Quantitative	CSIP + TSIP	Weak	
Reynolds 1999a (452)	Quantitative	CSIP + TSIP	Weak	
Townsend 2005 (454)	Quantitative	CSIP + TSIP	Weak	
Christie 2004 (439)	Qualitative	TSIP		Good
Halpenny and Cassie 2003 (441)	Qualitative	TSIP		Good
Burls 2007 (434)	Qualitative	CSIP		Moderate
Gooch 2005 (435)	Qualitative	CSIP		Moderate
Birch 2005 (448)	Qualitative	CSIP + TSIP		Moderate
Carter 2008 (449)	Qualitative	CSIP + TSIP		Moderate
O'Brien 2010a (450)	Qualitative	CSIP + TSIP		Good
Townsend 2006 (453)	Qualitative	CSIP + TSIP		Moderate
Townsend and Marsh 2004 (445)	Qualitative	Citation chase		Moderate
BTCV 2010 (438)	Mixed Methods	TSIP	Weak	Moderate
Wilson 2009 (443)	Mixed Methods	TSIP	Weak	Good
Yerrell 2008 (444)	Mixed Methods	TSIP	Weak	
O'Brien 2008a (450)	Mixed Methods	CSIP + TSIP	Weak	Good

* studies were included in the review but excluded from the synthesis due to poor study quality.
Key: TSIP = tailored study identification protocol and CSIP = Cochrane study identification protocol.

Qualitative studies: The Wallace Criteria

Where seven or more of the Wallace criteria were answered positively, studies were
scored as 'good', if studies met between four and six criteria positively, a 'moderate'
score was awarded.

In total, nine qualitative studies were identified (Table 18). The two studies uniquely
identified by the tailored study identification protocol were scored as 'good' (439, 441)

whereas the two studies uniquely identified by the Cochrane study identification protocol were scored as 'moderate' (434, 435). This data, and the quality appraisal of the studies identified by both the tailored study identification protocol and the Cochrane study identification protocol, is set out in Table 18.

8.3.9 Contribution to synthesis

The contributions of the quantitative and qualitative studies have been appraised separately. For the mixed method studies, these studies (Wilson 2009, Yerrell 2008 and O'Brien 2008) have been appraised separately for their contributions of quantitative and qualitative data.

Quantitative

No studies reporting quantitative data were uniquely identified by the Cochrane study identification protocol so the results reported here focus on the seven studies uniquely identified by the tailored study identification protocol and the five studies identified by both protocols. The heterogeneity of outcomes assessed by the study authors, the general lack of studies using controlled study designs, and the poor study quality overall, prohibited meta-analysis. The results are therefore summarised narratively and tabulated in Table 19 below.

Table 19: Quantitative results

Study	Identification Method	Mental and Emotional Wellbeing			HRQoL			Physical Activity Measures		
		Reported	Tool	Outcome	Reported	Tool	Outcome	Reported	Tool	Outcome
Barton 2009 (447)	CSIP + TSIP	✓	RSES + PMSS	No change	x			x		
O'Brien 2008a (450)	CSIP + TSIP	✓	ESS	Significant improvement	x			x		
Pillemer 2010 (451)	CSIP + TSIP	✓	NR	Reduction	✓	Retrospective comparison	Improvement with volunteers	✓	Unique to study	PA sig. associated with volunteers
Reynolds 1999a (452)	CSIP + TSIP	x			✓	SF-36	Improvements*	x		
Townsend 2005 (454)	CSIP + TSIP	✓	NR	Some differences	✓	Likert scale	Some improvements	x		
BTCV 2010 (438)	TSIP	x			✓	SF-12	Little/no change	x		
Eastaugh 2010 (440)	TSIP	x			✓	SF-36	Little/no change	x		
Small Woods 2011a (442)	TSIP	x			✓	SF-36	Improvements*	x		
Wilson 2009 (443)	TSIP	✓	WEMWBS	Increased or no change	✓	SF-12	Little/no change	✓	SPAQ	Increased PA
Yerrell 2008 (444)	TSIP	x			✓	PCS/MCS-12	Improvements	x		

Key: Emotional State Scale (ESS); Rosenberg self-esteem scale (RSES); Profile of Mood States scale (PMSS); physical activity (PA); Warwick-Edinburgh Mental Well-being Scale (WEMWBS); Scottish Physical Activity Questionnaire (SPAQ). CSIP = Cochrane study identification protocol and TSIP = tailored study identification protocol.

Notes: *very small sample sizes so robustness of results is questionable

Five outcome domains were of interest in this review:

1. physiological outcomes,

2. physical health measures,

3. mental and emotional wellbeing,

4. quality of life, and

5. physical activity measures

The tailored study identification protocol identified studies that contributed data to three of these outcomes: mental and emotional wellbeing (443); quality of life (438, 440, 442-444) and physical activity measures (443).

In the first domain (mental and emotional wellbeing), the identification and inclusion of Wilson et al. did not alter the overall conclusion of improvements of mental and emotional wellbeing (425, 443).

In the second domain (quality of life), one study reported HRQoL improvements (444). Two studies also reported improvements in HRQoL, one from the tailored study identification protocol (442) and another identified by the tailored study identification protocol and the Cochrane study identification protocol (452), but both studies had small sample sizes (Small Woods n=7 & Reynolds n=15 compared with Yerrell n=194) which limits the robustness of the findings (416). The findings of Yerrell would therefore appear valuable in this domain, in relation to their findings and relative to their sample size, although the uncontrolled before-and-after study design is considered of limited value in assessing causation (416, 444).

One study was unique to the tailored study identification protocol in the final domain (physical activity measures) (443). Wilson et al. reported increased physical activity, measured using a validated tool,12 weeks after participating in environmental enhancement activities (443). Only one other study evaluated physical activity measures (451). The study by Pilemer, identified by both the tailored and the Cochrane study identification protocols, also found improvements in physical activity scores but this was appraised retrospectively and through a scale created especially for their study (451). The findings of Wilson et al. would therefore appear valuable in this domain (416, 443).

Quantitative summary

Whilst the quality of each study (and therefore of the overall pool of studies) was weak regardless of study identification protocol, the value of each of the studies to the synthesis is clear. To generate a reliable understanding of intervention effectiveness, it was important that all studies reporting effectiveness outcomes are identified and the Cochrane study identification protocol would have missed studies and, thus, study data.

Qualitative

The findings of the qualitative studies were used to understand the links, as perceived by participants, between participation in environmental enhancement activities and health and wellbeing outcomes (416, 455).

Nine overarching themes were identified in the qualitative synthesis:

1. Physical activity

2. Personal achievement

3. Personal/ social identity

4. Developing knowledge

5. Benefits of place

6. Social Contact

7. Spirituality

8. Psychological benefits

9. Risks/negatives

Evidence available per theme

Table 20 records the study data available per theme. Eight of the nine themes were present in one or more of the studies rated as 'good' quality (Table 18) (455).

Table 20: Presence of qualitative themes in each study

Author	Identification Method	Personal Achievement	Personal/ Social Identify	Developing Knowledge	Benefits of place	Social Contact	Physical Activity	Spirituality	Psychological benefits	Risks/ negatives
Townsend & Marsh 2004 (445)*	Citation chase	✓	X	✓	✓	✓	✓	X	✓	X
		✓	X	✓	X	✓	✓	X	✓	X
Burls 2007 (434)	CSIP	✓	✓	✓	✓	✓	✓	✓	✓	X
Gooch 2005 (435)	CSIP	✓	✓	✓	✓	✓	X	X	✓	✓
Birch 2005 (448)	CSIP + TSIP	✓	X	X	✓	✓	✓	X	✓	X
Carter 2008 (449)	CSIP + TSIP	✓	✓	✓	✓	✓	✓	X	✓	X
O'Brien 2008a (450)	CSIP + TSIP	✓	✓	✓	✓	✓	✓	✓	✓	X
O'Brien 2010a (456)	CSIP + TSIP	✓	X	✓	✓	✓	X	✓	✓	X
Townsend 2006 (453)	CSIP + TSIP	✓	X	X	✓	✓	✓	X	✓	X
BTCV 2010 (438)*	TSIP	✓	X	✓	✓	✓	X	✓	✓	✓
		✓	X	✓	✓	✓	X	X	✓	✓
Christie 2004 (439)	TSIP	✓	✓	✓	✓	✓	X	✓	✓	✓
Halpenny & Cassie 2003 (441)	TSIP	✓	X	X	✓	✓	X	X	✓	X
Wilson 2009 (443)	TSIP	✓	✓	✓	✓	✓	X	X	X	✓

*there were two sub-groups for each of these citations
Key: TSIP = tailored study identification protocol and CSIP = Cochrane study identification protocol.

Contribution of studies per theme

The results of the analysis to determine the contribution of individual studies to the synthesis are recorded below. The first theme, Physical Activity, is summarised narratively and through Figure 16. The remaining eight themes are summarised narratively but with the corresponding figures being included in the supplementary file.

Studies are categorised as 'key studies' where they provide sufficient validity and richness to identify key concepts and develop primary and sub-themes. If a study provides either data richness, through a participant quotation to support a sub-theme, or a study confirms validity through identifying the themes and being cited in the final review, we categorise this as an 'additional study' since it provides additional but not unique contributions. If a study is identified as a 'key study' but it is also an additional study for another sub-theme, it is only counted once as a key study in the narrative since the synthesis is dependent on it.

Figure 16 Contribution of data to physical activity theme (qualitative studies)
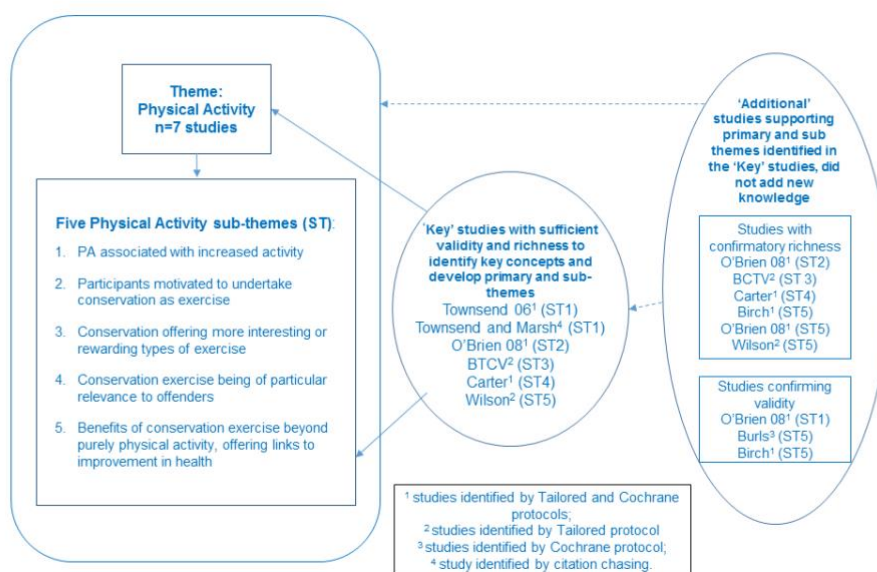
Physical activity

Figure 16 summarises the contribution of studies to this theme. Overall seven studies contributed data to this theme. Analysis of the sub-themes shows that five of the seven studies were 'key studies' with sufficient validity and richness to identify key concepts and develop primary and sub-themes (438, 443, 445, 449, 450, 453). Two

studies provided data that reinforced the primary theme or sub-themes identified from the key studies but did not contribute new knowledge to the synthesis (434, 448).

Personal achievement ((see the appendix (volume 2 of this thesis) for summary figure))

Overall, twelve studies contributed data to this theme. Analysis of the sub-themes shows that two studies were 'key studies' with sufficient validity and richness to identify all key concepts and develop primary and sub-themes (439, 443). Five studies provided data that reinforced the primary theme or sub-themes identified from the key studies but did not contribute new knowledge to the synthesis (434, 435, 438, 445, 453).

Personal/ social identity

Overall, six studies contributed data to this theme. Analysis of the sub-themes shows that three of the five studies were 'key studies' with sufficient validity and richness to identify key concepts and develop primary and sub-themes (439, 449, 450). Three studies provided data that supported the primary theme or sub-themes identified from the key studies but did not contribute new knowledge to the synthesis (434, 435, 443).

Developing knowledge

Overall, nine studies contributed data to this theme. Analysis of the sub-themes shows that three of the nine studies were 'key studies' with sufficient validity and richness to identify key concepts and develop primary and sub-themes (438, 450, 456). Six studies provided data that supported the primary theme or sub-themes identified from the key studies but did not contribute new knowledge to the synthesis (434, 435, 439, 441, 443, 449, 453).

Benefits of place

All 12 studies contributed data to this theme. Analysis of the sub-themes shows that five studies were 'key studies' with sufficient validity and richness to identify all key concepts and develop primary and sub-themes (439, 441, 443, 445, 450). Two studies provided data that supported the primary theme or sub-themes identified from the key studies but did not contribute new knowledge to the synthesis (434, 435).

Social contact

All 12 studies contributed data to this theme. Analysis of the sub-themes shows that five studies were 'key studies' provided sufficient validity and richness to identify all key concepts and develop primary and sub-themes (438, 441, 449, 450, 456). One study provided data that supported the primary theme or sub-themes identified from the key studies but did not contribute new knowledge to the synthesis (435).

Spirituality

Overall, five studies contributed data to this theme. Analysis of the sub-themes shows that two studies were key studies with sufficient validity and richness to identify all key concepts and develop the primary theme and sub-themes (439, 456). Three studies provided data that supported primary or sub-themes identified from the key studies but did not contribute new knowledge to the synthesis (434, 438, 450).

Psychological benefits

Overall, eleven studies contributed data to this theme. Analysis of the sub-themes shows that two studies were key studies with sufficient validity and richness to identify key concepts and develop the primary theme and sub-themes (439, 443). Four studies provided data that supported primary or sub-themes identified from the key studies but did not contribute new knowledge to the synthesis (434, 435, 441, 448).

Risk and negative impacts

Overall, four studies contributed data to this them. Analysis of the sub-themes shows that one of the five studies provided sufficient validity and richness to identify key concepts and develop primary and sub-themes (439). Two studies provided data that supported the primary theme or sub-themes identified from the key studies but did not contribute new knowledge to the synthesis (434, 435).

Qualitative summary

Within the nine overarching themes, 37 sub-themes were identified from nine studies (438, 439, 441, 443, 445, 449, 450, 453, 456). These nine studies were fundamentally key to the synthesis since they provided sufficiently rich data to identify key concepts and develop all the overarching themes and sub-themes. If any of these studies had been missed, the findings of the review would have been different since potentially unique data from sufficiently rigorous studies would have

been omitted from the synthesis. The identification and contribution of these nine studies was therefore key to the qualitative review. These nine studies were all identified by the tailored study identification protocol.

Studies supporting either overarching or sub-themes were included in the synthesis. Whilst the identification and inclusion of these studies increase the validity of the overall synthesis, two studies were only used in the synthesis to increase validity and they did not identify primary or sub-themes uniquely (434, 435). The omission of these studies from the synthesis would not alter the synthesis or change the findings of the review. These studies were uniquely identified by the Cochrane study identification protocol (434, 435).

The CERQual tool was used to appraise how much confidence could be placed in the findings listed above and its application in this study extends the work undertaken in our Cochrane Review. In this study, we first applied CERQual to all findings and included all studies in the analysis (Table 22). Secondly, we applied CERQual to all findings but excluded the study by Burls and the study by Gooch, since we sought to measure the contribution of bibliographic database searching in the Cochrane study identification protocol and the potential impact of missing these studies on the synthesis of studies (Table 22). Thirdly, we applied CERQual to all findings but excluded the study by Christie and the study by Halpenny and Cassie, since we sought to measure the contribution of author contact in the tailored protocol and the potential impact of missing these studies on the synthesis of studies (Table 23).

Table 21: CERQual all studies included

Review finding	Studies contributing to the review finding	Assessment of methodological limitations	Assessment of relevance	Assessment of coherence	Assessment of adequacy	Overall CERQual assessment of confidence	Explanation of judgement
Physical activity	Seven studies. (Townsend & Marsh 2004[1*]; Burls 2007[2]; Birch 2005[3]; Carter 2008[3]; O'Brien 2008a[3]; Townsend 2006[3]; Wilson 2009[4])	Minor methodological limitations Two studies were rated as good (O'Brien 2008a[3]; Wilson 2009[4]) Five studies were rated as moderate (Townsend & Marsh 2004[1*]; Burls 2007[2]; Birch 2005[3]; Carter 2008[3]; Townsend 2006[3])	No concerns	No concerns	Minor concerns	Moderate confidence	This theme was graded as moderate confidence since there were minor concerns on study quality and adequacy of data.
Personal achievement	Twelve studies (Townsend & Marsh 2004[1*]; Burls 2007[2]; Gooch 2005[2]; Birch 2005[3]; Carter 2008[3]; O'Brien 2008a[3]; O'Brien 2010a[3]; Townsend 2006[3]; BTCV 2010[4*]; Christie 2004[4]; Halpenny & Cassie 2003[4]; Wilson 2009[4])	No concerns Five studies rated as Good (Christie 2004[4]; Halpenny & Cassie 2003[4]; O'Brien 2008a[3]; O'Brien 2010a[3]; Wilson 2009[4]) Seven studies rated moderate (Townsend & Marsh 2004[1*]; Burls 2007[2]; Gooch 2005[2]; Birch 2005[3]; Carter 2008[3]; Townsend 2006[3]; BTCV 2010[4*])	No concerns	No concerns	No concerns	High confidence	This theme was graded as high confidence since there were no concerns in the four CERQual domains.
Personal/ Social Identity	Six studies (Carter 2008[3]; Christie 2004[4]; O'Brien 2008a[3]; Gooch 2005[2]; Wilson 2009[4]; Burls 2007[2])	No concerns Three studies were rated as good (Christie 2004[4]; O'Brien 2008a[3]; Wilson 2009[4])	No concerns	No concerns	No concerns	High confidence	This theme was graded as high confidence since there were no concerns in the four CERQual domains.

		Three studies were rated as moderate (Carter 2008[3]; Gooch 2005[2]; Burls 2007[2])					
Developing knowledge	Nine studies	No concerns	No concerns	No concerns	No concerns	High confidence	This theme was graded as high confidence since there were no concerns in the four CERQual domains.
	(Townsend & Marsh 2004[1*]; Burls 2007[2]; Gooch 2005[2]; Carter 2008[3]; O'Brien 2008a[3]; O'Brien 2010a[3]; BTCV 2010[4*]; Christie 2004[4]; Wilson 2009[4])	Four studies rated as good (Christie 2004[4]; O'Brien 2008a[3]; O'Brien 2010a[3]; Wilson 2009[4]) Five studies rated as moderate (Townsend & Marsh 2004[1*]; Burls 2007[2]; Gooch 2005[2]; Carter 2008[3]; BTCV 2010[4*])					
Benefits of place	Twelve studies	No concerns	No concerns	No concerns	No concerns	High confidence	This theme was graded as high confidence since there were no concerns in the four CERQual domains.
	(Townsend & Marsh 2004[1*]; Burls 2007[2]; Gooch 2005[2]; Birch 2005[3]; Carter 2008[3]; O'Brien 2008a[3]; O'Brien 2010a[3]; Townsend 2006[3]; BTCV 2010[4*]; Christie 2004[4]; Halpenny & Cassie 2003[4]; Wilson 2009[4])	Five studies rated as Good (Christie 2004[4]; Halpenny & Cassie 2003[4]; O'Brien 2008a[3]; O'Brien 2010a[3]; Wilson 2009[4]) Seven studies rated moderate (Townsend & Marsh 2004[1*]; Burls 2007[2]; Gooch 2005[2]; Birch 2005[3]; Carter 2008[3]; Townsend 2006[3]; BTCV 2010[4*])					
Social contact	Twelve studies	No concerns	No concerns	No concerns	No concerns	High confidence	This theme was graded as high confidence since there were no concerns in the four CERQual domains.
	(Townsend & Marsh 2004[1*]; Burls 2007[2]; Gooch 2005[2]; Birch 2005[3]; Carter 2008[3]; O'Brien 2008a[3]; O'Brien 2010a[3]; Townsend 2006[3]; BTCV	Five studies rated as Good (Christie 2004[4]; Halpenny & Cassie 2003[4]; O'Brien 2008a[3]; O'Brien 2010a[3]; Wilson 2009[4]) Seven studies rated moderate (Townsend & Marsh 2004[1*]; Burls					

Theme	Studies	Methodological limitations	Coherence	Relevance	Adequacy	CERQual confidence	Explanation
	2010[4*]; Christie 2004[4]; Halpenny & Cassie 2003[4]; Wilson 2009[4])	2007[2]; Gooch 2005[2]; Birch 2005[3]; Carter 2008[3]; Townsend 2006[3]; BTCV 2010[4*])					
Spirituality	Five studies (Burls 2007[2]; O'Brien 2008a[3]; O'Brien 2010a[3]; BTCV 2010[4*]; Christie 2004[4])	No concerns three studies were rated as good (O'Brien 2008a[3]; O'Brien 2010a[3]; Christie 2004[4]) two studies were rated as moderate (Burls 2007[2]; BTCV 2010[4*])	No concerns	No concerns	No concerns	High confidence	This theme was graded as high confidence since there were no concerns in the four CERQual domains.
Psychological benefits	Twelve studies (Townsend & Marsh 2004[1*]; Burls 2007[2]; Gooch 2005[2]; Birch 2005[3]; Carter 2008[3]; O'Brien 2008a[3]; O'Brien 2010a[3]; Townsend 2006[3]; BTCV 2010[4*]; Christie 2004[4]; Halpenny & Cassie 2003[4]; Wilson 2009[4])	No concerns Five studies rated as Good (Christie 2004[4]; Halpenny & Cassie 2003[4]; O'Brien 2008a[3]; O'Brien 2010a[3]; Wilson 2009[4]) Seven studies rated moderate (Townsend & Marsh 2004[1*]; Burls 2007[2]; Gooch 2005[2]; Birch 2005[3]; Carter 2008[3]; Townsend 2006[3]; BTCV 2010[4*])	No concerns	No concerns	No concerns	High confidence	This theme was graded as high confidence since there were no concerns in the four CERQual domains.
Risks and negative impacts	Four studies (Gooch 2005[2]; BTCV 2010[4*]; Christie 2004[4]; Wilson 2009[4])	No concerns Two studies were rated as good (Christie 2004[4]; Wilson 2009[4]) two studies were rated as moderate (Gooch 2005[2]; BTCV 2010[4*])	No concerns	No concerns	Minor concerns	Moderate confidence	This theme was graded as moderate confidence since there were minor concerns on the adequacy of data.

[1] Citation Chasing; [2] Cochrane study identification protocol; [3] Cochrane study identification protocol & Tailored study identification protocol, and; [4] Tailored study identification protocol. [*] there were two sub-groups for each of these citations.

Table 22: CERQual Burls and Gooch removed

Review finding	Studies contributing to the review finding	Assessment of methodological limitations	Assessment of relevance	Assessment of coherence	Assessment of adequacy	Overall CERQual assessment of confidence	Explanation of judgement
Physical activity	Six studies. (Townsend & Marsh 2004[1*]; Birch 2005[3]; Carter 2008[3]; O'Brien 2008a[3]; Townsend 2006[3]; Wilson 2009[4])	Minor methodological limitations Two studies were rated as good (O'Brien 2008a[3]; Wilson 2009[4]) Four studies were rated as moderate (Townsend & Marsh 2004[1*]; Birch 2005[3]; Carter 2008[3]; Townsend 2006[3])	No concerns	No concerns	No concerns	Moderate confidence	This theme was graded as moderate confidence since there were minor concerns on study quality. In this theme, Burls provides confirmatory validity alongside Birch for the same sub-theme. The loss of Burls would therefore be insignificant.
Personal achievement	Ten studies (Townsend & Marsh 2004[1*]; Birch 2005[3]; Carter 2008[3]; O'Brien 2008a[3]; O'Brien 2010a[3]; Townsend 2006[3]; BTCV 2010[4*]; Christie 2004[4]; Halpenny & Cassie 2003[4]; Wilson 2009[4])	No concerns Five studies rated as Good (Christie 2004[4]; Halpenny & Cassie 2003[4]; O'Brien 2008a[3]; O'Brien 2010a[3]; Wilson 2009[4]) Five studies rated moderate (Townsend & Marsh 2004[1*]; Birch 2005[3]; Carter 2008[3]; Townsend 2006[3]; BTCV 2010[4*])	No concerns	No concerns	Minor concerns The loss of Burls removes some confirmatory richness as a participant quote would be lost. The study that defines the sub-theme of 'payback' (Christie 04) remains, so the underlying data is not lost. This theme is well supported by studies.	High confidence	This theme was graded as high confidence since the loss of confirmatory richness in the form of Burls, was considered a minor point in the identification of the theme and contribution to the synthesis. Similarly, Gooch provides confirmatory validity to a sub-theme already supported by other studies one of which (Christie 04) is of better methodological quality.
Personal/ Social Identity	Four studies	No concerns	No concerns	No concerns	No concerns	High confidence	This theme was graded as high confidence since there were no concerns in

	(Carter 2008[3]; Christie 2004[4]; O'Brien 2008a[3]; Wilson 2009[4])	Three studies were rated as good (Christie 2004[4]; O'Brien 2008a[3]; Wilson 2009[4]) One study was rated as moderate (Carter 2008[3])			Neither the study by Burls or the study by Gooch provided either confirmatory richness or validity in this sub-theme. Moreover, neither study uniquely identified any subthemes.		the four CERQual domains. The omission of both Burls and Gooch would not alter this theme.
Developing knowledge	Seven studies (Townsend & Marsh 2004[1*]; Carter 2008[3]; O'Brien 2008a[3]; O'Brien 2010a[3]; BTCV 2010[4*]; Christie 2004[4]; Wilson 2009[4])	No concerns Four studies rated as good (Christie 2004[4]; O'Brien 2008a[3]; O'Brien 2010a[3]; Wilson 2009[4]) Three studies rated as moderate (Townsend & Marsh 2004[1*]; Carter 2008[3]; BTCV 2010[4*])	No concerns	No concerns	No concerns The loss of Burls removes some validating richness. The loss of Gooch removes some confirmatory richness as a participant quote would be lost.	High confidence	This theme was graded as high confidence since the change in assessment of adequacy was felt to be minor resulting in no change to the synthesis.
Benefits of place	Ten studies (Townsend & Marsh 2004[1*]; Birch 2005[3]; Carter 2008[3]; O'Brien 2008a[3]; O'Brien 2010a[3]; Townsend 2006[3]; BTCV 2010[4*]; Christie 2004[4]; Halpenny &	No concerns Five studies rated as Good (Christie 2004[4]; Halpenny & Cassie 2003[4]; O'Brien 2008a[3]; O'Brien 2010a[3]; Wilson 2009[4]) Five studies rated moderate (Townsend & Marsh 2004[1*]; Birch 2005[3]; Carter 2008[3]; Townsend 2006[3]; BTCV 2010[4*])	No concerns	No concerns	No concerns The loss of Burls removes some confirmatory richness as the study is quoted three times. On each occasion, it is only to confirm or validate	High confidence	This theme was graded as high confidence since there were no concerns in the four CERQual domains. The loss of Burls was considered more important than the loss of Gooch but neither studies were sufficiently valuable to alter the synthesis

	Cassie 2003[4]; Wilson 2009[4]				studies providing richer data.		since neither study directly supported the identification of any sub-themes.
Social contact	Ten studies (Townsend & Marsh 2004[1*]; Birch 2005[3]; Carter 2008[3]; O'Brien 2008a[3]; O'Brien 2010a[3]; Townsend 2006[3]; BTCV 2010[4*]; Christie 2004[4]; Halpenny & Cassie 2003[4]; Wilson 2009[4])	No concerns Five studies rated as Good (Christie 2004[4]; Halpenny & Cassie 2003[4]; O'Brien 2008a[3]; O'Brien 2010a[3]; Wilson 2009[4]) Five studies rated moderate (Townsend & Marsh 2004[1*]; Birch 2005[3]; Carter 2008[3]; Townsend 2006[3]; BTCV 2010[4*])	No concerns	No concerns	Minor concerns Burls is not referenced in the synthesis. Gooch provides validating richness to one sub-theme.	High confidence	This theme was graded as high confidence. The minor concerns on adequacy are very minor concerns since neither study identified a sub-theme or provided confirmatory richness in the form of participant quotes.
Spirituality	Four studies (O'Brien 2008a[3]; O'Brien 2010a[3]; BTCV 2010[4*]; Christie 2004[4])	No concerns three studies were rated as good (O'Brien 2008a[3]; O'Brien 2010a[3]; Christie 2004[4]) one study was rated as moderate (BTCV 2010[4*])	No concerns	No concerns	No concerns The loss of Burls removes some validating richness but it is one of four studies cited in the identification of a sub-theme so the contribution of Burls is questionable.	High confidence	This theme was graded as high confidence since there were no concerns in the four CERQual domains.

Psychological benefits	Ten studies (Townsend & Marsh 2004[1*]; Burls 2007[2]; Gooch 2005[2]; Birch 2005[3]; Carter 2008[3]; O'Brien 2008a[3]; O'Brien 2010a[3]; Townsend 2006[3]; BTCV 2010[4*]; Christie 2004[4]; Halpenny & Cassie 2003[4]; Wilson 2009[4])	No concerns Five studies rated as Good (Christie 2004[4]; Halpenny & Cassie 2003[4]; O'Brien 2008a[3]; O'Brien 2010a[3]; Wilson 2009[4]) Five studies rated moderate (Townsend & Marsh 2004[1*]; Birch 2005[3]; Carter 2008[3]; Townsend 2006[3]; BTCV 2010[4*])	No concerns	No concerns	No concerns	High confidence	This theme was graded as high confidence since there were no concerns in the four CERQual domains.
Risks and negative impacts	Three studies (BTCV 2010[4*]; Christie 2004[4]; Wilson 2009[4])	No concerns Two studies were rated as good (Christie 2004[4]; Wilson 2009[4]) one study was rated as moderate (BTCV 2010[4*])	No concerns	No concerns	Minor concerns	moderate confidence	This theme was graded as moderate confidence since there were minor concerns on the adequacy of data.

[1] Citation Chasing; [2] Cochrane study identification protocol; [3] Cochrane study identification protocol & Tailored study identification protocol, and; [4] Tailored study identification protocol. [*] there were two sub-groups for each of these citations.

Table 23: Christie and Halpenny & Cassie removed

Review finding	Studies contributing to the review finding	Assessment of methodological limitations	Assessment of relevance	Assessment of coherence	Assessment of adequacy	Overall CERQual assessment of confidence	Explanation of judgement
Physical activity	Six studies. (Townsend & Marsh 2004[1*]; Birch 2005[3]; Carter 2008[3]; O'Brien 2008a[3]; Townsend 2006[3]; Wilson 2009[4])	Minor methodological limitations Two studies were rated as good (O'Brien 2008a[3]; Wilson 2009[4]) Four studies were rated as moderate (Townsend & Marsh 2004[1*]; Birch 2005[3]; Carter 2008[3]; Townsend 2006[3])	No concerns	No concerns	No concerns	Moderate confidence	This theme was graded as moderate confidence since there were minor concerns on study quality. Christie and Halpenny and Cassie did not contribute to this theme so there are no changes to the CERQual judgement.
Personal achievement	Eight studies (Townsend & Marsh 2004[1*]; Birch 2005[3]; Carter 2008[3]; O'Brien 2008a[3]; O'Brien 2010a[3]; Townsend 2006[3]; BTCV 2010[4*]; 2003[4]; Wilson 2009[4])	Moderate concerns Three studies rated as Good (O'Brien 2008a[3]; O'Brien 2010a[3]; Wilson 2009[4]) Five studies rated moderate (Townsend & Marsh 2004[1*]; Birch 2005[3]; Carter 2008[3]; Townsend 2006[3]; BTCV 2010[4*])	No concerns	Minor concerns The loss of Christie represents the loss of relevant data to support and identify sub-themes. The loss of Christie therefore raises questions about the coherence of the sub-themes since Christie identifies sub-themes that are supported by other weaker studies.	Major concerns The loss of Christie represents the loss of relevant data and a key study. Sub-themes would have been missed.	Low confidence	This theme was graded as low confidence. The loss of Christie & Halpenny and Cassie represent the loss of two 'good' quality studies from this theme. The loss of Christie, specifically, represents the loss of what we consider a key study to this theme which, in terms of adequacy would mean two sub-themes would have been missed.

Personal/ Social Identity	Three studies (Carter 2008[3]; O'Brien 2008a[3]; Wilson 2009[4])	Moderate concerns Two studies were rated as good (O'Brien 2008a[3]; Wilson 2009[4]) One study was rated as moderate (Carter 2008[3])	No concerns	Moderate concerns The data on the sub-theme of identity being linked to the impact in the environment was incoherent. Christie was the only 'good quality' study in the identification of this sub-theme and it provided data that contrasted with other studies.	Minor concerns In comparison to other themes, this theme was weakly supported by study data. The loss of Christie as a key study raises concerns.	Moderate confidence	This theme was graded as moderate confidence. The omission of Christie would alter the understanding of this theme in the synthesis of studies.
Developing knowledge	Six studies (Townsend & Marsh 2004[1*]; Carter 2008[3]; O'Brien 2008a[3]; O'Brien 2010a[3]; BTCV 2010[4*]; Wilson 2009[4])	No concerns Three studies rated as good (O'Brien 2008a[3]; O'Brien 2010a[3]; Wilson 2009[4]) Three studies rated as moderate (Townsend & Marsh 2004[1*]; Carter 2008[3]; BTCV 2010[4*])	No concerns	No concerns	No concerns	High confidence	This theme was graded as high confidence.
Benefits of place	Eight studies (Townsend & Marsh 2004[1*]; Birch 2005[3]; Carter 2008[3]; O'Brien 2008a[3]; O'Brien 2010a[3]; Townsend 2006[3]; BTCV 2010[4*]; Wilson 2009[4]	Minor concerns Three studies rated as Good (O'Brien 2008a[3]; O'Brien 2010a[3]; Wilson 2009[4]) Five studies rated moderate (Townsend & Marsh 2004[1*]; Birch	No concerns	No concerns	Minor concerns Removing Christie removes some validating richness through the loss of participant	Moderate confidence	This theme was graded as moderate confidence since there were minor concerns in the two CERQual domains.

		2005[3]; Carter 2008[3]; Townsend 2006[3]; BTCV 2010[4*])			quotes to support sub-themes. Other, weaker, studies do provide data, however.		
Social contact	Eight studies (Townsend & Marsh 2004[1*]; Birch 2005[3]; Carter 2008[3]; O'Brien 2008a[3]; O'Brien 2010a[3]; Townsend 2006[3]; BTCV 2010[4*]; Wilson 2009[4])	Minor concerns

Three studies rated as Good (O'Brien 2008a[3]; O'Brien 2010a[3]; Wilson 2009[4])

Five studies rated moderate (Townsend & Marsh 2004[1*]; Birch 2005[3]; Carter 2008[3]; Townsend 2006[3]; BTCV 2010[4*]) | No concerns | No concerns | Minor concerns | Moderate confidence | This theme was graded as Moderate confidence |
| *Spirituality* | Three studies (O'Brien 2008a[3]; O'Brien 2010a[3]; BTCV 2010[4*]) | No concerns

two studies were rated as good (O'Brien 2008a[3]; O'Brien 2010a[3];)

one study was rated as moderate (BTCV 2010[4*]) | No concerns | No concerns | Major concerns

The loss of Christie would prohibit the identification of one (out of two) sub themes. | Low confidence | This theme was graded as low confidence since there was major concerns on data adequacy. |

Psychological benefits	Eight studies (Townsend & Marsh 2004[1*]; Burls 2007[2]; Gooch 2005[2]; Birch 2005[3]; Carter 2008[3]; O'Brien 2008a[3]; O'Brien 2010a[3]; Townsend 2006[3]; BTCV 2010[4*]; Wilson 2009[4])	No concerns Three studies rated as Good (O'Brien 2008a[3]; O'Brien 2010a[3]; Wilson 2009[4]) Five studies rated moderate (Townsend & Marsh 2004[1*]; Birch 2005[3]; Carter 2008[3]; Townsend 2006[3]; BTCV 2010[4*])	No concerns	No concerns	No concerns	High confidence	This theme was graded as high confidence since there were no concerns in the four CERQual domains.
Risks and negative impacts	Two studies (BTCV 2010[4*]; Wilson 2009[4])	No concerns One study was rated as good (Wilson 2009[4]) one study was rated as moderate (BTCV 2010[4*])	No concerns	No concerns	Minor concerns	moderate confidence	This theme was graded as moderate confidence since there were minor concerns on the adequacy of data.

[1]Citation Chasing; [2] Cochrane study identification protocol; [3] Cochrane study identification protocol & Tailored study identification protocol, and; [4] Tailored study identification protocol. [*] there were two sub-groups for each of these citations.

The use of CERQual allows us to measure the impact of potentially missing studies from either search protocol and to explore any possible changes to the synthesis of studies. It also helps demonstrate the utility of both search approaches, helping us to interpret the value of studies and, therefore, the search protocols or search methods.

CERQual: excluding the study by Burls and the study by Gooch (Table 22)

We found no difference in the overall confidence of findings in any of the nine domains if the study by Burls and the study by Gooch were removed. We observed small changes in the assessment of adequacy in three cases but these changes did not alter the overall confidence using CERQual. These changes were:

- physical activity: minor methodological limitations were consistent between both analyses. This did not change the overall CERQual assessment of moderate confidence;

- personal achievement: the removal of Burls raised minor concerns in the assessment of adequacy but the overall CERQual assessment of high confidence remained unchanged;

- social contact: the use of Gooch to provide validating richness was a minor concern in the assessment of adequacy but the overall CERQual assessment of high confidence remained unchanged; and

- risks and negative impacts: minor methodological limitations were noted in the assessment of adequacy, since the removal of Gooch would potentially remove a sub-theme. This would not, however, change the overall CERQual assessment of moderate confidence in this domain. Overall, this domain was of limited importance to the synthesis.

This analysis would appear to confirm our finding that the study by Burls and the study by Gooch did not materially affect the synthesis of qualitative studies. This would suggest that in missing these particular studies the synthesis, as presented in our Cochrane Review, would remain unchanged.

CERQual: excluding the study by Christie and the study by Halpenny & Cassie (Table 23)

We observed a difference in the overall confidence of findings in five of the nine domains if the study by Christie and the study by Halpenny & Cassie were removed. These changes significantly altered the confidence in findings and, therefore, would

appear to impact negatively on the synthesis of studies had these two studies been missed by our searches. The changes were in the following domains:

- personal achievement: the CERQual assessment was altered by the removal of these two studies, being downgraded from high confidence to moderate confidence. The loss of Christie (specifically) raised major concerns in the assessment of adequacy and minor concerns in the assessment of coherence. Furthermore, minor concerns were raised in methodological limitations, since both the removed studies were 'good quality' studies;

- personal/social identity: the CERQual assessment was altered by the removal of these two studies, being downgraded from high confidence to moderate confidence. The loss of Christie raised concerns on adequacy and coherence specifically;

- developing knowledge: there was no change in the CERQual assessment. This theme was graded as high confidence even in spite of the omission of Christie;

- benefits of place: the CERQual assessment was altered by the removal of Christie, being downgraded from high confidence to moderate confidence. The loss of Christie raised concerns on adequacy specifically;

- social contact: the CERQual assessment was altered by the removal of these two studies, being downgraded from high confidence to moderate confidence;

- spirituality: the CERQual assessment was altered by the removal of Christie, being downgraded from high confidence to low confidence. The loss of Christie raised concerns on adequacy; and

- risks and negative impacts: minor methodological limitations were noted in the assessment of adequacy. This would not, however, change the overall CERQual assessment of moderate confidence in this domain. Overall, this domain was of limited importance to the synthesis.

This additional analysis would appear to confirm our finding that the study by Burls and the study by Gooch did not materially affect the synthesis of qualitative studies, whereas the studies by Christie and Halpenny and Cassie did.

8.3.10 Discussion

This section seeks to highlight the differences between the tailored study identification protocol and the Cochrane study identification protocol as they relate to (i) the effectiveness of study identification, measured here by the number of studies identified and the number of studies identified uniquely, and (ii) the differences in the value of the studies, measured here by differences in study quality and the contribution to the synthesis of the studies identified. We focus on the primary study identification methods of the Cochrane study identification protocol (database searching) and the tailored study identification protocol (contacting organisations/web-searching), since these are ultimately the approaches by which the studies were uniquely identified in each case.

Effectiveness

Number of studies identified

The Cochrane study identification protocol identified 21,409 studies to screen compared to 453 studies identified by the tailored study identification protocol. Interpreting the difference between the tailored study identification protocol and the Cochrane study identification protocol in strictly numerical terms should be treated with caution since it risks overstating the efficiency of the tailored study identification protocol.

Prior to registering the review with The Cochrane Public Health Group, we had queried the utility of undertaking exhaustive and sensitive bibliographic database searches at the start of the review process. Researchers have found that even sensitive search strategies will not identify all studies in topics where a standardised or controlled terminology does not yet exist (179, 457), and key topic search terms for this review, nature or natural (for example), have multifarious application both as descriptors of place (i.e. adjectives) and also as definers of activity (i.e. adverbs). Defining a sufficiently sensitive literature search strategy, that produced a manageable number of search results to screen, represented a challenge, which was further compounded as standard techniques to improve efficiency in bibliographic database searches, such as the use of study design literature search filters, are not recommend in public health topics or reviews of conservation interventions (89, 90).

Contacting study authors and organisations as a primary method of study identification ameliorated some of these issues in the tailored study identification

protocol. Previous studies have evaluated the effectiveness of contacting study authors to identify studies or study data (240, 293, 297, 307) but they have focused on the effectiveness of contact to identify data (as supported by our case study). We identified a further advantage: contacting study authors or organisations allowed us to explain our research question and inclusion criteria through conversation, circumventing the ambiguity of the search terms used in bibliographic database searching. Database hosts do not presently permit semantic searching, meaning that most search terms (indexing terms aside) do not differentiate retrieval based on meaning. Contacting relevant authors and organisations involved in the types of interventions under review allowed us to explain our research questions and this explains the lower number of studies identified. A positive side effect was to develop awareness and interest in our review from practitioners and policy makers.

In terms of effectively identifying studies and study data, our findings accord with other study authors who also report that contacting authors and experts will identify studies missed by bibliographic database searching (241, 421). Improved effectiveness should not, however, be confused with improved efficiency. We are comparing the searches retrospectively, and did not record the time taken to identify included studies using the Cochrane study identification protocol or the tailored study identification protocol at the time of the original review, but we conservatively estimate that the process of searching and screening in the Cochrane study identification protocol, and contacting organisations and web searching in the tailored study identification protocol, were approximately equal. The process of contacting organisations and web-searching is time intensive (88, 307) with accompanying problems of data management and replicability (307). Bibliographic databases, almost without exception in this review, have export facilities to bibliographic management tools, whereas managing and de-duplicating studies identified through organisation contact and web-searching required manually entering study data into a bibliographic tool for screening (308).

Number of studies identified uniquely

After screening, the Cochrane study identification protocol identified two studies uniquely (434, 435) and the tailored study identification protocol identified nine studies uniquely: four using quantitative study designs (436, 437, 440, 442), two qualitative studies (439, 441) and three mixed-methods studies (438, 443, 444).

All studies using quantitative designs were identified by the tailored study identification protocol, whereas two qualitative studies were missed by the tailored study identification protocol. Understanding why the two qualitative studies were missed by the tailored study identification protocol would be almost impossible to unpick, since it would require re-contacting 288 organisations to ask them why they did not recommend those two studies. We explore the value of these two missed studies to the synthesis, and therefore develop our understanding of the significance of missing these studies in the tailored study identification protocol below, under study value.

Methodologically, the process of screening the 21,409 studies (31 days work at 7hrs a day/ screening at a rate of 100 studies per hour) identified in the Cochrane study identification protocol in order to identify two unique studies validates our initial concern that this topic was not necessarily suitable – or perhaps the topic area was not yet mature enough – for relying upon the application of sensitive, systematic bibliographic database searching. Researchers have previously questioned the utility of extensive online searches when compared with contacting organisations likely to collect review-relevant data (89, 421), and our findings in this study would support the usefulness of contacting organisations. Indeed, it could be worth questioning the practicable need for exhaustive bibliographic database searches in topics which are multidisciplinary and have a diverse evidence base, such those at the intersection of environmental management and health, since the comprehensive identification of studies is often not an attainable goal. More research needs to be done to understand the value of alternative approaches in different topic areas, including public and environmental health.

It should be noted that the tailored study identification protocol did not directly compete against use of bibliographic database searches. As shown in Figure 14, we proposed to undertake bibliographic database searches as a supplementary (i.e. adjunct), rather than as a primary method of study identification. We intended to use focused bibliographic database searches (217), informed by our earlier grey literature searches. These searches were not ultimately required, since we used the bibliographic database searches of the Cochrane study identification protocol as a surrogate.

Changing the chronological order of study identification methods from the Cochrane study identification protocol to the tailored study identification protocol may initially

appear to be superficial but what we really seek to alter is the allocation of searching effort. This study confirms the value of aligning the primary method of study identification to where studies are most likely to be identified. In this case, the belief of our expert panel, that grey literature studies would be important to this review, meant we prioritised identification and searching effort for such studies over formally published studies indexed in bibliographic databases. The idea that the chronological order of study identification methods, led by a primary method of study identification, reflects the likely location of studies and affects the distribution of searching effort is not without precedent, since it forms the basis of the Cochrane study identification protocol. In the Cochrane study identification protocol, the information need (typically for studies reporting RCTs) is matched to a corresponding process of study identification. Generically, the process of study identification, as conducted by an expert searcher, can be perceived as starting from the methods most likely to identify relevant studies (and most likely to identify the most studies) to methods least likely to identify studies. Searching end-to-end of this methodological process seeks to address the risk of publication bias, since even those studies that are more difficult to identify are still sought, although in reality the time spent searching, using each individual search method, is often different and decreases after the primary method is undertaken. Hartling et al. explore the possibility of prioritising which databases to search in systematic reviews (172) but we believe that this study is the first to prioritise and allocate search methods, in particular, supplementary search methods, in a review.

Studies have demonstrated (Helmer et al. (458)) or explored (Greenhalgh and Peacock (111)) the use of supplementary search methods but our findings would suggest that categorising study identification methods as primary or supplementary is unhelpful, since no guidance exists on which search methods should be used for different review needs (241). Our findings suggest that matching methods of study identification to the evidence base proved valuable in this case study and this approach may hold value not only for similar topics but also for other topic areas with a disparate evidence base.

Study value

Studies that evaluate search effectiveness commonly interpret effectiveness as the identification of studies missed when measured against a comparator or alternative search approach (113). Additional studies identified by alternative search methods

can provide valuable information to researchers but the perceived value of those newly identified studies is seldom established and is difficult to measure accurately (179).

Study quality

Quantitative

As Table 18 illustrates, all identified quantitative studies, both formally published (identified by the Cochrane study identification protocol and tailored study identification protocol) and grey literature studies (tailored study identification protocol only) were appraised as being of weak study quality in our Cochrane Review. There is no perceivable improvement in study quality between the grey and published studies identified by the tailored study identification protocol, a finding that is consistent with other studies (28).

Qualitative

Conversely, there was a difference in study quality between the tailored study identification protocol and the Cochrane study identification protocol (Table 18). Three grey literature studies identified only by the tailored study identification protocol (439, 441, 443) scored one category higher on the Wallace criterion than the two published studies identified only in the Cochrane study identification protocol (434, 435). It is possible that the unpublished nature of the grey literature, with no limitation on the use of tables or words count, meant that greater detail was provided on the methods and results than would be possible in a journal article study. We interpret this idea cautiously, since the number of studies concerned is limited, and there is no wider empirical evidence to aid interpretation of this finding. Moreover, it does not follow that because greater detail is provided on the methods and results, that the study is generally of better quality.

Contribution to the synthesis

Quantitative

Comprehensive study identification is an important part of evaluating intervention effectiveness as it is linked to producing a reliable estimate of intervention effectiveness (28). The fact that the Cochrane study identification protocol would have missed nine studies (four quantitative and three mixed-methods) evaluating the effectiveness of environmental enhancement and conservation activity interventions

is an important finding when considering the contribution of the tailored study identification protocol to the synthesis of effectiveness studies in this field. It highlights the importance of so-called 'supplementary search methods', perhaps suggesting that they are in fact complementary (possibly primary) methods of study identification.

Qualitative

With the qualitative studies, we found that two studies made no significant contribution to the synthesis and we therefore question the value of these studies in the synthesis and the impact of identifying them. We conclude that, had these studies been missed in study identification, the impact on the synthesis would have been negligible.

The study by Burls and the study by Gooch were uniquely identified by the Cochrane study identification protocol and after screening a significant number of non-relevant studies. We initially questioned the need for, and utility of, comprehensive bibliographic database searches in this review. Whilst this perception is only now clear through retrospective analysis, the research waste in searching, screening and ordering full-text in the Cochrane study identification protocol is potentially troubling, especially since we questioned the utility of comprehensive searching at the outset. We lacked the metric to test or demonstrate our concerns beyond suspicion. A metric to formatively test the effectiveness of study identification would be a valuable contribution to the process of systematic review.

Our findings in this case study raises further questions as to whether it is possible to conduct truly "comprehensive" searches for reviews (or topics) in which the evidence is widely dispersed across both bibliographic databases and the 'grey literature,' and it highlights the need for so-called supplementary study identification methods (458). Given the specific findings from the qualitative studies, this argument could be extended to reviews of qualitative studies: specifically that comprehensive study identification is unlikely to prove an attainable goal in most cases (112).

In retrospectively analysing both study identification protocols, we feel that the time invested in scoping, working with the PRG, and the make-up of our research team and team discussion, was of great benefit in developing the tailored study identification protocol. Linking the methods and process of study identification to study quality, or contribution of studies to synthesis, could help researchers better

understand the value of investing in the process of study identification or selecting more appropriate study identification methods. Matching methods of study identification to studies, and potentially working out when (or how) not to search, could yield benefits in the efficiency of study identification in systematic reviews.

8.3.11 Study limitations

The use of a case study research design to report this study means that the findings should be interpreted with caution since they relate to a single case study.

A limitation of this study is that time taken to undertake each individual search method was not recorded. This limits any interpretation as to the efficiency of the tailored study identification protocol and Cochrane study identification protocol. Recording time taken to search more generally would develop the evidence on the effectiveness and efficiency of searching in systematic reviews.

The quality of the studies identified and included in our Cochrane Review was variable, which prohibits not only the interpretation of results and the conclusions that can be drawn from The Cochrane Review but also, it inhibits our ability to interpret the contribution of the study identification and to make links to study value. Better quality studies would aid interpretation and discussion.

Our use of CERQual to explore the contribution of the qualitative studies might be considered a limitation since its discriminant validity is yet to be established. Nevertheless, the use of CERQual in a supportive capacity reduces the dependence of the results on this specific tool.

8.4 Conclusions

In this study, we sought to link the idea of search effectiveness to study value. We retrospectively found that, in the case of a mixed methods review of a topic that crossed environmental and public health boundaries, extensive bibliographic database searching was of limited value in terms of contribution to synthesis but that grey literature searching was valuable and identified studies that made unique contributions to both the quantitative and qualitative synthesis.

What we demonstrate in this case study is that the sequential order of study identification methods can be altered from a conventional study identification protocol. This, in effect, gives study identification methods different weighting depending upon how much effort and time is invested in them relative to the

anticipated value. In the tailored study identification protocol, our primary methods of study identification were grey literature searching and contacting experts, which we demonstrate contributed valuable studies and study data. We valued bibliographic database searching as lower priority, so aimed to treat it as a supplementary study identification method, which, by comparing with the Cochrane study identification protocol, was valid.

8.5   Chapter findings

Having reproduced the journal article above, I now situate the findings in the context of this thesis.

This case study represents the second test of a tailored approach compared to the conventional approach and the findings directly address the research aims of this thesis.

The principal finding is that the tailored approach was more effective when compared to the conventional approach. This finding is based upon: a lower total volume of studies to screen and a greater number of studies uniquely identified by the tailored approach. Furthermore, the tailored approach identified studies missed by the conventional approach, which meant that the value of the tailored approach could be illustrated. The use of CERQual to identify, assess and report the value of the conventional and the tailored approach is a unique contribution to knowledge arising from this work since I move the dialogue of effectiveness evaluation in literature searching beyond purely quantitative measurements of effect.

8.6   Contributions of this chapter to the thesis

This chapter addresses the research aims of the thesis and it builds upon the work presented in the previous chapter and Section One of the thesis. This case study demonstrates that the idea of tailored literature searching can be used as alternative process to literature searching, even in 'gold standard' Cochrane reviews.

I further extend the concepts of a tailored literature search and reflect on them, namely: that the team work together to shape the information needs, that the information specialist plays the role of 'decision maker' to select the search methods used and the order in which they will be undertaken.

This chapter also reports a second examination of an alternative way to demonstrate the effectiveness of the tailored approach when compared to the conventional

approach. I was able to extend my evaluation here to capture the idea of value by measuring the contribution of studies to the synthesis of studies.

8.6.1 Application of the tailored approach

The development and application of the tailored approach in this case study built upon my learning and experience from the first case study (see Chapter 7). Identifying the need for the tailored approach, based upon a potentially unmanageable number of studies to screen with a low number of anticipated studies meeting inclusion, happened early in the scoping stage of the review. Discussion with the research team led to further discussion with the Project Reference Group at our first advisory meeting. The approach to literature searching, the types of evidence available (by design and publication status), where the evidence was available, and how to access it, were key points of discussion in this meeting. The advisory meeting also provided a forum for discussion on search terms. This meeting informed the development of the tailored approach, in particular the decision to prioritise author contact and grey literature searching, and it builds upon my experience in the first case study and the critique of Key Stage Three of the conventional approach that I set out in Chapter 3. Specifically, the work of Armstrong et al. and Arksey et al. who encourage investing time in the early scoping stage of a new project (207, 208).

Developing the conventional approach was, conversely, a matter of following the Cochrane handbook with input from the Cochrane Public Health Group and peer review of the Cochrane Review protocol. The peer review of the Cochrane protocol included sign-off from the Cochrane group on our bibliographic database search strategy but there was little engagement with, or comment on, the use of or need for supplementary search methods.

Reflecting on the process of developing the tailored and conventional approaches for this review, it is possible to draw out a distinction between the methodological interaction of the Cochrane group in the conventional approach, which ensured that our Cochrane review met the methodological expectations of the Cochrane public health group but entirely lacked any guidance on topic, and the topical interaction with our expert advisory group, which informed development of the tailored approach but lacked any guidance on methods. It is potentially beneficial to have both the topic AND methodological input but it is not necessarily a requirement of Cochrane guidance when undertaking a Cochrane review to do so. As I demonstrate in the case study, the topic input led to identifying studies we might otherwise have missed,

and there are notable benefits in terms of perception of our work, and potentially of review up-take by publishing with Cochrane.

In peer review of the published case study, I was questioned about the weight placed on the database first (i.e. conventional) approach which underpins the comparison of search approaches in this chapter. The peer reviewer said:

'in the guidance the authors may discuss database searching first before all the other search methods, I am not sure that this reflects a rule, more a suggested approach that for most topics is likely to yield the highest volume of records relatively quickly. It is guidance and not a commandment. The authors should distinguish between the guidance and how people choose to interpret it.'

The criticism is worthy of reflection. The peer reviewer alludes to a tension which I have not addressed so far in the case studies.

I establish in the critical literature review (see Chapter 3) what guidance recommends as the process of literature searching in systematic reviews (the conventional approach), and I attempt to explore how this is understood by comparing guidance to published studies. The critical literature review setting out the conventional approach had not been published at this time, and so I could not refer the peer reviewer to this work in my thesis.

Further work is indicated to explore how guidance is followed in systematic reviews and if the conventional approach that I identify in Chapter 3 is actually followed in 'real life.' This reflects a challenge arising in this case study, which was the tension between the guidance on topic from the expert advisory group and the guidance on methods from Cochrane. These were quite simply disparate unions. We did not identify a mechanism to combine the topic and methods guidance, rather the tailored approach was conducted quietly and alongside the Cochrane process. Whilst Cochrane are more recently open to the inclusion of observational and qualitative studies, there is no indication that Cochrane groups would accept different approaches to literature searching.

I reference the peer reviewer comment above because it is a fair criticism of the case study. In reality, it more broadly supports the work of the case study than it challenges it, since I am arguing through tailored literature searching that guidance can and should be challenged and I demonstrate this. Whilst we (as a research team) were unable to alter the process as it relates to use of methods and delivery of

our Cochrane review, publishing this case study begins to add weight to the idea more generally. To change practice, it is necessary to explore, to demonstrate and evaluate approaches, which this case study contributes too, but it is likely necessary to 'win over' Cochrane groups, adopting their key principles which I identify in the literature review, namely: that systematic reviews aim for a comprehensive identification of studies to minimise bias and that the search methods used are transparently reported and replicable.

Whilst there is valuable critique to be found in exploring the above, the criticism made by the peer reviewer's comment is open to criticism by reply. There are three immediate problems with the reviewer's comment:

1) The Cochrane handbook is not guidance, it is a form of guidance-led instruction. Authors who wish to publish Cochrane reviews have very little room to interpret the Cochrane Handbook (as the peer reviewer suggests) and Cochrane can, and they do, refuse to publish reviews which do not follow guidance or meet their standards of review. All attempts to vary the guidance of the handbook in our review were met with a form of resistance, so it is not correct to say that we could have interpreted the handbook.

2) As I found in my review of supplementary search methods (Chapter 5), and the critical literature review in Chapter 3, there is more guidance available on searching bibliographic databases than on using supplementary search methods. Whilst I accept that a bibliographic database is more likely to yield the highest volume of studies relatively quickly, as the peer reviewer observes, it does not follow that this is a reason to use this search method first, which I demonstrate in the literature review (Chapter 3) is presently recommend best practice.

What I demonstrate in the results of this case study is that there is benefit to prioritising other search methods, or more specifically aligning search methods to the task in hand, since it allows a researcher to control the process of literature searching, deploying those search methods that may take longer to perform to generate relevant studies. How long search methods take is a gap in knowledge I identified in Chapter 3 and I attempted to address in the literature review of supplementary search methods (Chapter 5) by including 'resource requirements' of the supplementary search methods. My proposal was to identify approximate timings of supplementary search methods so that,

as was the case in this case study, if a researcher was planning to undertake author contact, they wold know that the task takes between 14-22 days between request and response (6, 293). It may therefore be practical to start making contact with authors and, whilst waiting for response, run the bibliographic database searches.

3) The peer reviewer also misses a principal finding of the case study. Whilst I accept that the highest volume of studies was returned with the greatest efficiency by searching bibliographic databases, this should be situated in the context of the findings of the case study. Namely, that no relevant studies were identified by searching bibliographic databases. The implication of the Cochrane process (and the conventional approach) is bibliographic database searching is implicitly an important and unmovable part of literature searching in systematic reviews. As I demonstrate in this case study, and it is important to note that this based only on this case study, this black-letter law is questionable.

A key finding for reflection is also a key contribution to knowledge. In Chapter 6 (and in Chapter 7), I set out my idea that the value of missed or identified studies is an important concept to capture and to demonstrate in literature searching. It is, I suggest, situated in the call to identify new ways to report the effectiveness of literature searching which I summarise in Chapter 6.

The use of CERQual to compare the search approaches is a novel contribution in terms of methodology. It could be seen as comparable to Egger et al's use of statistical meta-analysis to determine the effect of missing studies for systematic reviews of intervention effectiveness. This was an important contribution to the case study since it helped to capture the differences between a conventional approach and a tailored approach. It also help demonstrate, for the first time, my feeling that information professionals need to find new ways to demonstrate the importance of their work (see 4.4.1 and also the critique of key stage three)

8.6.2 Measuring the comparison between search approaches

I aimed to measure the differences between the tailored approach and the conventional approach as I set out in Chapter 6. I repeat the metrics measured and the methods used here for ease of reference.

Table 24 how I will measure effectiveness in the case studies

Effectiveness (i)	The total number of studies identified by each search
Effectiveness (ii)	The number of studies uniquely identified by each search
Efficiency	The time taken to screen each output
Value	The contribution of any unique studies on the synthesis

Given the relative difference between the numbers of studies identified by the conventional approach (21, 409) and the tailored approach (453), there is a concern that the case for effectiveness in the tailored approach could be over-stated in purely quantitative terms. The time taken to search would mediate this concern but the data was not accurately retained by the research team. This prohibits further critique of the effectiveness in the context of relative efficiencies by either approach. This is a limitation of this case study.

Identifying value, however, extends the work presented in the previous case study. I was not able to explore the value of a tailored approach in the previous study since no unique studies were identified by the tailored approach or (therefore) missed by the conventional approach. Determining the value of studies in this case study, and the use of CERQual to do so, makes a unique contribution to knowledge. I situate the importance of demonstrating value in my idea beyond the tailored approach (see 4.1) and the belief that information professionals need to move beyond simply identifying more studies to explaining why the studies that they identify are important and worthy of the effort invested. This work is situated in the work of Savoie et al. who suggested that demonstrating value was key to determining effectiveness of search methods (459). This application of method, and findings, I suggest is situated in the discussion on the need for comprehensive literature searching in qualitative evidence synthesis. Whilst this finding is the result of only one case study, and there are limitations in the work (as I set out above), I demonstrate not only the importance of literature searching but also that the contribution of studies in qualitative evidence synthesis not equal. The publication of this study advances the work reported by myself and colleagues in 2012 where we raised similar themes but without the benefit of empirical testing (112).

8.7   Implications for practice

This case study builds on the work of the previous case study in testing and evaluating a tailored literature search when compared to a conventional approach. It represents a second successful test of the tailored approach when compared to the

conventional approach. The publication of this work makes a valuable contribution not only to systematic reviews in the environmental enhancement field but also in reviews of complex interventions and qualitative evidence synthesis since it demonstrates that alternative approaches to literature can be developed and undertaken.

The principal contribution to knowledge arising, and therefore implication to practice, is the worked example of CERQual to examine the reliability of findings and demonstrate the 'effect' of missing studies. This represents a possible new method of assessing study value in qualitative evidence synthesis and it moves the debate on in literature searching as to the effect of 'missing' studies. With researchers questioning the need for, and possibility of, comprehensive literature searches for qualitative evidence synthesis, my work represents a potential change in practice.

Section Two: A summary

In Section Two, I compared the conventional approach to the tailored approach in two case studies based on two separate systematic reviews. This directly addressed the research aims of this thesis, since it compares an existing approach to literature searching (the conventional approach) to a new and proposed method of literature searching (the tailored approach).

The work reported in this section builds on the work reported in Section One, in which I defined the conventional approach and the tailored approach. The review of supplementary search methods (Chapter 5) informed the selection and prioritisation of literature search methods in the second case study (Chapter 8), and the systematic review of methods and metrics to determine literature search effectiveness (Chapter 6) informed the metrics I used to evaluate the comparison between the two search approaches in both the case studies that I report in Chapter 7 and Chapter 8.

The findings of the two case studies reported in this section demonstrated that the tailored approach identified a lower total volume of studies to screen when compared to the conventional approach. It also demonstrated that the tailored approach did not miss studies in the first case study (Chapter 7) and the studies potentially missed in the second case study (Chapter 8) would not have made a difference to the synthesis of studies in a qualitative evidence synthesis.

I established the principles of undertaking a tailored literature search in Chapter 4 and I tested this in Chapter 7, the first case study. I suggest that these guiding principles are linked to the success of the tailored approach in both case studies. Namely, that investing time to prepare to literature search, working with the research team and with the information specialist leading on decision making for literature searching, and working closely with topic experts to develop literature search approaches, is an effective approach to developing literature search approaches beyond systematic reviews of intervention effectiveness. This points, and findings, were further developed and re-enforced in the second case study reported in Chapter 8.

I also sought to simplify the reporting of effectiveness evaluation between the two approaches to literature searching. Focusing on yield provides a clear statement of the effect of each approach to literature searching as it relates to screening. In the

second case study (Chapter 8), I was able to identify and report the idea of study value by determining which studies were valuable and which were less valuable in the synthesis of qualitative studies. This extends the analysis of the two searches approaches and it is a novel contribution to research knowledge.

Section Three: Discussion and Conclusions

In this section of the thesis, I aim to draw together the work I have presented in Section One and Section Two. Having defined and examined the conventional and tailored approaches to literature searching, and reviewed supplementary literature search methods and methods and metrics used to determine the effectiveness of literature searching in Section One, I directly compared the conventional approach to the tailored approach in Section Two. I will summarise my findings alongside my objectives and in the context of the relevant literature.

In Chapter 9, I will re-visit the aims and objectives of this thesis that I initially set out in Table 2 (Chapter 2). I will set out how I have addressed the nine objectives. I will then set out the discussion section of this thesis.

In Chapter 10, I set out the contributions to knowledge arising from my work and the conclusions of the thesis; and

In Chapter 11, I set out a future research agenda based on the work arising from this thesis.

9   Discussion

In this chapter, I begin by revisiting the aims of the thesis and its 9 objectives which I initially set out in Table 2. Table 2 has been amended in this chapter (Chapter 9) to incorporate findings and it sets out how these findings contribute to addressing the aims of this thesis. I set this out in Table 25 below.

Table 25 How the aims and objectives of this thesis have been addressed

Section, Chapter & Methods	Research question(s)	Objective(s)	How the findings addressed the objective(s)	Contribution to the thesis aims
Section 1 *Chapter 3* Critical literature review	Can a shared model of the literature searching process be detected across guidance documents and, if so, how is this process reported and supported? If a shared model can be identified, is it suitable to meet the needs of researchers undertaking systematic reviews?	1) To conduct a critical review of influential systematic review guidance documents to assess whether they share an understanding of the approach to searching for systemic reviews. 2) To use the findings of this review to articulate a model of current approach(es) to searching in systematic reviews. 3) To conduct a structured critical review of the empirical studies that have investigated each stage of this model in order to understand if there is support for the proposed model and whether it is suitable to meet the needs of researchers.	Through a critical literature review of nine guidance documents and 119 supporting studies, I demonstrated the existence of a shared model of literature searching for systematic reviews. I illustrated this as a conceptual process model, made up of eight key stages, which I set out and I explored in Chapter 3. I named this process the 'conventional approach' since the process identified was common convention in all handbooks. Through critique of the guidance and studies set out above, I explored the suitability of the conventional approach to meet the needs of researchers undertaking systematic reviews. Of the eight key stages which make up the conventional approach, I determined that three key stages appear suitable to meet the needs of researchers; three key stages partially appear suitable to meet the needs of researchers; and in two key stages, it is unclear if they are suitable and meet need the needs of researchers. For the three key stages that partially agreed, and the two keys stages in which it was unclear, I indicated refinements to the key stage or further work within the thesis.	Chapter 3 establishes the existence of the conventional approach in literature searching for systematic reviews. I demonstrate that the conventional approach is based on recommend best practice in the form of nine leading guidance documents and that it constitutes the conventional approach between guidance for literature searching in systematic reviews. In this way, I establish the conventional approach as the reference standard for the purposes of comparison in this thesis. By critiquing the suitability of the conventional approach, I began to examine how researchers literature search and I consider if it is appropriate to use the same process of literature searching to identify RCTs for systematic reviews of intervention effectiveness as for qualitative evidence synthesis.
Section 1 *Chapter 4*	What is the origin of thinking behind	4) To discuss the original thinking behind the need for tailored literature searching.	In this chapter I set out the origin for my idea of tailored literature searching. The idea originated in public health systematic reviews for NICE,	This Chapter establishes the tailored literature search approach in the thesis. The tailored

Section, Chapter & Methods	Research question(s)	Objective(s)	How the findings addressed the objective(s)	Contribution to the thesis aims
Literature review	tailored literature searching? What is tailored literature searching? How does tailored literature searching work in theory?	5) To theorise how tailored literature searching could work conceptually.	where I was becoming increasingly convinced that researchers were over-searching bibliographic databases and missing studies by focusing on guidance and not understanding how supplementary search methods worked. I illustrate tailored literature search models and I define and explore how tailored literature searching might work. Tailored literature search believes that the process of searching and the search methods used should match the information needs of the review. This casts the information professional it the role of the decision maker where, working closely with the research team, an approach to literature searching is tailored. By developing the critique of suitability for the conventional approach in Chapter 3, I explore potential refinements to the conventional approach that might meet the needs of reviews that looks before effectiveness, I also undertake a scoping search for other models of information retrieval theory to help explore how other researchers have developed and presented alternative models of information retrieval. This work allowed me to build off the illustrated process model and explain how tailored literature searching can work in the in theory.	approach is, in effect, the index test for the purposes of this thesis.
Section 1 *Chapter 5*	How do empirical studies of supplementary search techniques compare to the	6) To review the review methodology handbooks and empirical studies on supplementary search methods, in order to	In Chapter 5, also published as a literature review, I reviewed the guidance of four review methodology handbooks and 35 empirical studies for five supplementary search methods.	The conventional approach and the tailored approach to literature searching both utilise supplementary search methods. The tailored approach perhaps more so than the conventional approach since it is focused on

Section, Chapter & Methods	Research question(s)	Objective(s)	How the findings addressed the objective(s)	Contribution to the thesis aims
Literature review – publication one	recommendations in the review methodology handbooks?	determine the advantages, disadvantages and resource requirements of supplementary search methods.	I categorised the findings into advantages and disadvantages and I identified resource requirements, where data permitted.	matching information needs to search methods as opposed to the conventional approach which matches comprehensive literature searching to a research question.

For my thesis, this literature review resulted in development of an evidence-based tool-box of literature search methods. By way of example, this review informed the selection of the supplementary search methods in Chapter 8. It thereby develops the comparison of the conventional approach to the tailored approach to literature searching in systematic reviews.

The chapter has since been published and it represents a unique contribution to knowledge since it is the first literature review of supplementary search methods. That the chapter is published open-access means other researchers have a literature searching 'tool box' too. |
| *Section 1*

Chapter 6

Systematic review – publication two | What metrics or methods are used to calculate literature search effectiveness?

Which metrics or methods are used formatively or summatively?

How is effectiveness defined by study authors? | 7) To conduct a systematic review to identify all metrics or methods currently used by researchers to demonstrate literature search effectiveness.

8) To use the findings from this systematic review to identify which methods are formative and summative and how study authors define what effectiveness means in literature searching. | I identified six metrics and one method used to calculate literature search effectiveness by undertaking a systematic review of studies.

Six metrics were summative (Sensitivity, Specificity, Precision, Accuracy, Number Needed to Read and Yield) and one method was formative (Capture recapture);

None of the studies identified explicitly defined effectiveness or clearly reported what the threshold (or cut-off) was for an "effective" result in the context of their evaluation. On this basis, I | In order to compare which of the conventional approach or the tailored approach is most effective for the particular review in question, it is necessary to explore what underlying values are used to calculate effectiveness in literature searching and how other researchers have determined effectiveness in literature searching to date.

I explore these aspects and use my findings to develop the comparison between the conventional approach and the tailored approach in Chapter 7 and 8. My findings from this review directly informed how I compared |

Section, Chapter & Methods	Research question(s)	Objective(s)	How the findings addressed the objective(s)	Contribution to the thesis aims
			was unable to conclude how study authors defined effectiveness in their studies.	the conventional approach to the tailored approach in this thesis.

This systematic review also developed my understanding of this topic so it contributes to my advancement as a researcher. I am left with a clearer impression than before that the use of diagnostic test accuracy terminology and methods is not helping researchers understand clearly what effectiveness means in literature searching. |
| *Section 2*

Chapter 7 and 8

Case studies – publications three and four | How does the tailored literature search method compare to the conventional approach in practice? | 9) to report a case study in which I compare a tailored approach to literature searching to a conventional approach and to examine the findings of this case study in context of the research question of this thesis. | In this comparative case study, I propounded working definitions of effectiveness and value for this systematic review (as above). Overall, I found that a tailored approach to literature searching was more effective and offered better value than the conventional approach to literature searching. See Chapter 8. | These publications directly address the aim of my thesis, since they compare the tailored approach to the conventional approach on an empirical basis.

These case study allows me to test and demonstrate the criteria I set out in Chapter 6 to evaluate and demonstrate effectiveness.

These case studies also validate the refinements set out in Chapter 3, namely: the role of the information professional as decision maker, the value of the team in helping design the approach to literature searching, and the value of tailoring the approach to literature searching on research waste. |

9.1 Issues and challenges arising from this thesis

I will now set out the issues and challenges arising from the use of tailored literature searching when compared with the conventional approach. My objective in this chapter is to situate my thesis findings, summarised in Table 25, in the relevant literature, to consider the implications of my findings as it relates to literature searching and, more broadly, systematic reviews, and to the set out the next steps for work which is beyond the scope of this thesis.

The issues and challenges identified in this chapter are informed by the preceding chapters of this thesis and the findings of the 9 objectives identified in Chapter 2 and repeated in Table 25 in Chapter 9. I also draw upon my experience as an information professional and researcher. Considering the findings alongside the objectives will allow me to set out how I have addressed the research aims of this thesis.

As a reminder, the aims of this thesis are to: 1) examine approaches to systematic literature searching for systematic reviews and, 2) propose and test a method of systematic literature searching for reviews that do not focus on the effectiveness of clinical interventions.

9.2 Who should literature search in systematic reviews and an 'evolving' role for the information professional?

9.2.1 The issue: the role of the information professional

In Chapter 3, Key Stage One, I questioned who should undertake literature searching in a systematic review (7). The guidance documents I reviewed (Table 3) make clear recommendations that information professionals are appropriate researchers to undertake literature searching in a systematic review. The studies I cited to support this idea agree but they do not provide empirical evidence for these claims.

Even if only a hypothetical straw-man, because I believe that the number of completed systematic reviews with literature searching undertaken by information professionals establishes a point similar to that of guidance, I suggest that the information professional is potentially in transition from the role of a librarian to an information professional and researcher in the context of systematic reviews. There is good support for this idea in the studies that I cited in Key Stage One and the work of other authors (140, 149, 158, 460).

In considering the refinements to the conventional approach (see 4.4.8), I situate the information professional as 'the decision maker' in the process of literature searching. I demonstrate the effect of this 'decision making' role specifically in my case studies, presented in this thesis in Chapter 7 and Chapter 8, since I can assess my involvement, and the 'effect' of my decision-making, on review-level outcomes, namely, a reduction in the volume of studies to screen without missing any relevant studies.

I see the information professional's role as comparable to a statistician or health economist within a systematic review - in that they offer specialist, methodological input. The involvement of an information professional is manifestly linked to the success of the resulting review (147) as I identify in my case studies (see Chapter 7 and Chapter 8), since identification of relevant studies is a critical success factor in conducting the review. The information professional brings specific practical and methodological skills, which the guidance and supporting studies suggest are necessary to success in undertaking a systematic review. These arguments are explored in Chapter 3, which discusses the literature on both expert searchers (460-466) and the role of information professionals in systematic reviews (140-152, 157).

In the conventional approach, which I argue in Chapter 3 informs the current process of literature searching in systematic reviews, it is unclear what 'active role' the information professional can (or should) 'actively play' in defining a process of literature searching, since there is (or should be) little variation in the process of literature searching. In Chapter 3, I make a distinction between the instrumental and conceptual literature searching tasks, suggesting that practical tasks are more generally understood as information professional tasks to the exclusion of the theory-based tasks. For instance, decision making in the conventional approach tends to centre on the selection of databases (80, 92), design of the search strategy, and engagement with supplementary search methods. These are all valuable inputs into the design of the process, and they do influence the outcomes of literature searching (95, 111, 182, 241), but the information professional is working within a tightly defined process and in a passive way.

In a tailored approach, the information professional is more clearly the decision maker, since I demonstrate in Chapter 7 and Chapter 8, that the information professional can, and should, select the model of information retrieval (131, 132) and participate in an active way in the overall design of the process of literature

searching. The rationale for this is similar to the guidance documents I reviewed: the information professional is best placed to lead on information retrieval decisions. I would suggest that this finding builds on the idea of that information professionals are in transition. As I indicate in Chapter 3, support will be needed, but the idea of the information professional as the 'decision maker' would seem to consolidate the advantages of working with an information professional whilst situating the importance of the role in systematic reviews.

9.2.2 The challenge: The information professional as the decision maker

The role of the information professional as decision maker represents a challenge to researchers generally and information professionals specifically. The active role of the information professional in the tailored approach calls for experienced information professionals (146, 467, 468). This potentially requires the development of the information professional role (158), with more advanced research training than in my experience was available at library school (UCL, 2006-2008 (p/t)), since decisions are being taken which may influence the identification of relevant studies and, therefore, the quality of the resulting review (469). Training on methods and study designs would seem the most important next steps, since the information professional needs to be able to delineate between, for example, methods of data collection and study design when making information retrieval decisions.

There are considerable benefits to this development in the process of systematic review (147, 155), as I demonstrate in the case studies reported in Chapter 7 and Chapter 8. The practical next steps would seem to indicate knowledge brokering. Researchers need to work with information professionals, understanding the nature of their work and trusting their professional judgement. Information professionals need to work with researchers to understand more about the nature of their work and the impact of literature searching on systematic reviews. This should orientate within the debate on the meaning and purpose of comprehensiveness and also centralise on the impact of screening large data sets. This leads to my next issue and challenge.

9.3 Working with information professionals: the role of the team

9.3.1 The issue: how to work with a team

As suggested above, and demonstrated in the case studies (Chapters 7 and 8), the tailored approach makes the information professional the decision maker in selecting

an approach to literature searching. I suggest that this works best in the context of a team-based approach (141, 462, 470).

I used a team-based approach in the projects which resulted in the work presented in Chapter 7 (50) and Chapter 8 (77). In the work presented in Chapter 8, the case study based on The Cochrane Review, a team-based approach informed the selection of search methods in the tailored approach, which I demonstrate was ultimately more effective than an approach led by searching bibliographic databases. My findings in the work presented in Chapter 7 adds to the findings of Savoie et al. who also found 'value' beyond database searching (459), and it develops the work of Haddaway and Bayliss who found working with experts was an effective way to identify unpublished studies and data (421).

The issue here is in planning a tailored search approach (see 9.4) and managing the process of engagement with the review team. It requires working with the review team to reduce the anomaly in knowledge (131) and producing the clearest possible understanding of information needs (4, 130). It also requires that a research team is willing to accept and explore new approaches and accept the risks and challenges of presenting these new approaches to searching (c.f. (177, 471, 472)).

9.3.2 The challenge: working as a team

The challenge is assembling a team with sufficient and requisite experience to inform discussion and allow enough time to undertake the preliminary investigation of the topic area. Whilst teams commonly produce systematic reviews, there is little evaluation of working models and best practice in managing the effectiveness and efficiency of the literature search process within this context. This may leave researchers and systematic reviewers unclear about the tasks and the time it takes to deliver these tasks. The next steps need to include a clearer outline of the variety of tasks, including the challenges, and efforts to start developing data on the amount of time systematic reviews tasks can take.

9.4 The importance of planning

9.4.1 The issue: how to plan for literature searching

The structure of the conventional approach removes some of the need for planning since the process of literature searching is prescriptively described. The guidance reviewed in Chapter 3 (see Table 3) considered planning to include assessing the

need for the review, or checking that it had not been done before, rather than actually planning 'how to literature search'.

From experience, planning in the conventional approach includes selecting which bibliographic databases to search and when to undertake searching. In the tailored approach, planning is of greater importance, since the approach to literature searching is determined by the information need(s) of the review (131). Planning in the tailored approach is likely to include: determining an approach to literature searching based on scoping (see Chapter 8), team discussion (see Chapter 7 and Chapter 8) and developing a timeline for literature searching (which may be an iterative process) during the life-course of a systematic review.

9.4.2 The challenge: further research is needed on how to plan for literature searching

The challenge here is determining how to plan for literature searching and this is linked to the experience of the information professional (Section 9.2.1) and the role of the team (Section 9.3.1) in informing development. The work presented in Chapter 8 sets out a worked example but there is undoubtedly more work to be done, not least since this is finding is based on the strength of one case study. As I found in Chapter 3, Key Stage 3, there is a limited literature, and limited evidence available on how to plan for literature searching. I set out below, some examples of next steps, based on the work presented in this thesis and my development through undertaking the work.

Studies indicate that searching beyond bibliographic databases is valuable (95, 459) and that researchers need to develop a clearer understanding of the world beyond bibliographic databases (8). Planning, in this context, could include thinking about the types of studies or evidence required (79, 80) and the format(s) in which this is presented (75). I explore this in Chapter 8, the case study based on the Cochrane review of participation in environmental enhancement activities, where I attempted to link the types of studies needed to where the studies were likely to be identified. I drew on the specific search-method-by-search-method evidence that I identified and summarised in Chapter 5, in the review of supplementary search methods. The idea of linking studies needed, to their location, and the search methods best suited to identifying studies seeks to add to the work by Booth, who looked at 'search styles and tactics' (101), and guidance in empirical studies, such as the work by Bayliss et al. who explore 'where and how to search' for public health interventions (82). Looking beyond bibliographic databases is central to tailored literature searching and

it is linked to the idea of matching models of information retrieval to information needs (131, 132, 185, 286). It requires further research on how to plan approaches to literature searching, especially since the resource implications are unclear, which impedes planning.

Guidance documents can inform planning and the conventional approach (Chapter 3) offers a pragmatic model to literature searching in reviews (7). A challenge remains to test and establish how the conventional approach is used in practice, although it should be remembered that the model has been tacitly tested in numerous systematic reviews already, without formal recognition of the process. That guidance documents exist does not mean that they are followed in practice although it is worth remembering that, in some cases, such as those described in the case study presented in Chapter 8, guidance must be followed.

A balance in planning needs to be found between delivering robust and transparent literature searching in newly effective and efficient ways and not confusing this with shortcuts, since these have been linked to damaging the quality of systematic reviews (469). An information professional must recognise when to use a tailored approach or when there is limited benefit to altering the process of searching.

The experience of the information professional, and the willingness to take on the role of 'decision maker' and potentially set out and defend approaches to literature searching, is important in the planning stages. This is linked not only to investing time to plan literature searching but also setting what are (in essence) proposals for a particular approach. The information professional may benefit from training to prepare for this role. I indicate this in my proposal for future research and training (see 11.5).

9.5   Issues with the conventional approach or reasons to use a tailored search approach

The case studies reported in Chapter 7 and Chapter 8 highlight some of the issues with the use of the conventional approach in systematic reviews. I summarise and discuss these below.

9.5.1   The issue: whether comprehensive literature searches are necessary in all types of systematic review?

As I set out in Chapter 3, the aim of the conventional approach is for 'comprehensive' study identification which is linked to minimising the introduction of bias (5, 7). Put

simply, the aim of the conventional approach is not to miss studies, since it has been demonstrated that missing studies in reviews of intervention effectiveness can alter the pooled estimate of effectiveness (28). The effect of missing studies is less certain in other types of systematic reviews, such as public health reviews of complex interventions, where meta-analysis may not be appropriate or feasible, and qualitative evidence synthesis (112). With the absence of statistical meta-analysis, it is more difficult to demonstrate the 'effect' of missed studies on the synthesis with certainty. I explore this issue in the case study based on a Cochrane review in Chapter 8, and I contribute to the evidence on measuring the effect of missing studies in a qualitative evidence synthesis, through the use of CERQual (See Chapter 8) (433). I set this out in my contributions to knowledge arising from this thesis (see 9.7).

The issue here is that comprehensive literature searches have been adopted as the prevailing marker of search quality (see Chapter 6 (113)). Missing studies in literature searching for a systematic review is considered a serious problem. The conventional approach attempts to guard against the possibility of missing studies through the chronological application of all known and all possible literature search methods, which can lead to research waste. As I report in Chapter 3, and I find in Chapter 6, and is reflected in the literature generally, even comprehensive searching can miss studies (28).

9.5.2 The challenge: looking beyond 'comprehensive' literature searches

The challenge here is two-fold:

i) what constitutes "effectiveness" in the context of systematic literature searching is worthy of further consideration (Chapter 6), since it is possible that different types of review have materially different thresholds for what constitutes effectiveness and comprehensive literature searches are not always possible or even attainable (112); and

ii) the 'value' of studies identified by literature searching should be weighted when considering the true 'effect' of missing studies. I explore this point in the Cochrane review case study (Chapter 8) where, using CERQual (433), I determined that a qualitative evidence synthesis would have been unchanged had we 'missed' two studies identified by the comprehensive literature searches required for a Cochrane systematic review (77).

The view that comprehensive literature searching is optimal in all types of systematic reviews is worthy of further challenge and evaluation. The findings of this thesis would seem to support calls for new methods and ways to measure, demonstrate and explain comprehensiveness in literature searching and its impact on systematic reviews. I contribute to this debate in the systematic review I present in Chapter 6 and the evaluations undertaken in the case studies in Chapter 7 and Chapter 8. Moreover, a clearer understanding of what constitutes 'research waste' in literature searching would be valuable, helping to contextualise this work.

9.5.3 The issue: allocating time to literature search

None of the nine handbooks or supporting studies reviewed in the critical literature review (Chapter 3) reported how much time to allocate to the process of literature searching in systematic reviews. This is perhaps because the time to search depends on the topic of review and the experience of the information professional (325). It is also unclear how much overall searching time should be allocated per search method within the conventional approach, and the impact of literature searching in terms of the number of studies generated for screening.

Allocating time to literature searching is an important issue because the time available to identify relevant literature in a systematic review is limited (325). Moreover, as I demonstrate in the case studies reported in Chapter 7 and Chapter 8, how time is allocated to search methods can have profound effects on the identification of studies.

Studies have attempted to estimate the time to undertake a systematic review (187, 473, 474) and, furthermore, categorise the time to individual tasks (325, 473) and grey literature searching specifically (325). They are unable to generate reliable conclusions, simply because, as I found in the literature review of supplementary search methods (Chapter 5), the underlying data is not reported in studies.

In Chapter 5, I categorised the resource requirements of supplementary search methods, in an attempt to generalise the findings per search method and produce an initial average to help inform the selection of supplementary search methods in future reviews. I was unable to achieve this aim, since studies seldom report the time taken to search. I conclude that reporting the time to search, and indeed undertake other tasks in systematic reviews, can inform decision making.

There is an argument for 'time saving' from using tailored approaches to literature searching. I demonstrate this in Chapter 7, where I argued that the time saved (estimated as 35 days) allowed better engagement with supplementary search methods, which ultimately identified studies missed by bibliographic database searching. Until such time as an accurate record of timing is recorded for review tasks, arguments for time saving should be considered with caution. In the Cochrane review case study (Chapter 8), for instance, whilst a lower number of studies was screened in the tailored protocol, identifying the relevant studies took as long as comprehensive database searching. There is also an argument that the time saved in screening fewer studies for inclusion, owing to the 'technically' more efficient searching, is counterbalanced by the time initially invested in preparing for the tailored literature search. Accordingly, more time invested upfront by the individual information professional potentially saves the whole review team time later in the screening process.

9.5.4   The challenge: researchers should record time

The challenge raised by this issue is for information professionals and researchers to record and report the time taken to both literature search generally and undertake individual search methods specifically. In time, these measurements could more generally be used to inform broader guidance on the time needed to literature search (242), since researchers indicate (as reported in Chapter 5) that different search methods take different amounts of time to undertake (75, 325). This might be developed further to correlate time with cost (e.g. (390)) or, as I discuss below (Section 9.6.1) and in Chapter 7 and Chapter 8, time with value.

9.5.5   The issue: transparent search reporting

A particular strength of the conventional approach is that it assists with transparent search reporting. This is a requirement of all the guidance documents reviewed in Chapter 3, the critical literature review, and it is linked to the reproducibility of literature searches and minimising the introduction of bias. Transparent reporting is also linked to increased confidence in the search results (164) and, therefore, the underlying belief that the conclusions of the review are reliable, since the steps taken can be repeated (178).

The quality of literature search reporting is generally poor (154, 157, 262, 270), despite reporting guidelines being available (c.f. (183, 475)) and studies providing

guidance on how to report literature searches (195). The issue is that reporting bibliographic database searching is easier, since the exporting of search strategies is possible and it is possible to re-create literature searches. Supplementary searches present problems with literature search reporting (199). In web-searching, for instance, users experience through different browsers different versions of what is technically the same internet, so reporting and, thereafter, recreating web-searches is almost impossible (201, 308).

9.5.6 The challenge: developing search reporting guidelines that incorporate supplementary search methods

This issue represents a challenge to both the conventional approach, which after all uses supplementary searches, and the tailored approach, which makes perhaps greater use of supplementary search methods. Improvements in search reporting are required to meet the needs of transparently reported and robust systematic reviews (14). An area for further work is to explore a version of PRESS (183, 475) for supplementary search methods.

9.6 What does effectiveness mean in literature searching?

9.6.1 The issue: defining effectiveness

It was my contention at the start of this thesis that effectiveness in literature searching meant different things to different groups of people. Researchers ideally wanted low total volumes of studies to screen as an output of literature searching, without missing any potentially relevant studies. Information professionals were focused on not missing studies, so favoured high sensitivity in literature searching, referencing the documented need for comprehensive literature searches in the leading handbooks on the process of systematic reviews. This point is perhaps even further confused, since effectiveness means different things in different types of review (as I explore above). I explored these issues and find support for them in the systematic review of methods or metrics to evaluate literature searched effectiveness (Chapter 6).

9.6.2 The challenge

In measuring effectiveness, the challenge is possibly one of communication between the information professional and the research team, since 'effectiveness' is possibly situated in the aims of the review and the information needs. Different researchers

have different information needs and also different understandings of comprehensiveness (219). Developing a team understanding of what constitutes 'effectiveness' will ensure that the correct parameters are recorded in literature searching and the most useful metrics are reported.

There is also the challenge of demonstrating effectiveness. In Chapter 8, through the development of a search filter for a review of hard-to-reach populations, I attempted to demonstrate effectiveness not only in quantitative terms but also qualitatively as 'value' found in the studies as measured by the contribution of studies to the synthesis. My work on value takes on a challenge set by Savoie et al. who suggested that demonstrating value was key to determining effectiveness of search methods (459). Whilst I found a way to demonstrate value, the demonstration of value is context specific, and the findings are limited to one case-study (as it stands). This challenge remains unsatisfied and this is an area for further research.

9.7   Contributions to knowledge arising

Having discussed the key issues above, in Table 26 below, I set out what I believe to be the contributions to knowledge arising from the work conducted for this thesis.

Table 26 Contributions to knowledge

Contribution to knowledge	Description of contribution	Where located in the thesis
1. Identification of the conventional approach	I demonstrate the existence of a shared model of the literature searching process in systematic reviews. This confirms that nine different guidance documents articulate the same process to literature searching for different types systematic reviews. By demonstrating the existence of this shared model, I am able to confirm that it exists and I question if it suitable to meet the needs of systematic reviewers. Defining the key stages in this review helps categorise the scholarship available, and it prioritises areas for development and further study.	Chapter 3. The paper supporting this model has also been published as stand-alone, open-access piece of work (see (7)).
2. Development of a new and alternative model of information retrieval specific to systematic reviews	By setting out the tailored approach to literature searching, I have developed a new model of information retrieval and one that is specifically focused on literature searching in systematic reviews. It offers an alternative processing model to the conventional approach which I identify constitutes recommend best practice in literature searching. I illustrate, demonstrate, and I evaluate the effect of this new model compared to the conventional approach in the published case studies, Chapter 7 and Chapter 8.	Chapter 4.
3. The first published review of supplementary search methods	I have undertaken and published in an open access journal the first review of supplementary search methods. By categorising how supplementary search methods work, how they are applied in systematic reviews, and the consequent advantages, disadvantages and resource implications of each search method, I have produced a leading resource for the systematic review community. It will allow researchers to access one source to make evidence-based decisions on using or better understanding supplementary search methods.	Chapter 5. This paper has been published as an open-access study (see Table 1).
4. The first systematic review to identify methods and metrics used to determine literature search effectiveness	I have undertaken and published the first systematic review of metrics or methods to evaluate the effectiveness of literature searching. This review identifies and it reports detail on six metrics and one method and their corresponding calculations. This review aims to clarify the language used to explore effectiveness evaluation in literature searching and it makes detailed and specific recommendations to improve practice.	Chapter 6. This paper has been published (see Table 1).
5. The first published evaluations of a new model of	I have contributed two published case studies which explore a new model of information retrieval (the tailored approach).	Reported in Chapter 7 and Chapter 8.

information retrieval specific to systematic reviews.	These publications make a number of sub contributions to knowledge, namely:	This paper has been published (see Table 1).
	• They identify that experienced researchers can work together to challenge guidance and yet produce literature searches that meet the needs of decision makers; • The case studies highlight the importance and benefit of preparing to literature search and working as a team to develop literature searches; • The case studies respond to a gap in knowledge to find and examine new ways to determine and report the effectiveness of literature searching; and • The case-studies demonstrate the role of the information specialist as decision maker and the effect they can have on the process of systematic review.	
6. I have responded to calls to simplify the reporting of literature search effectiveness evaluation and I extend the evaluation of effectiveness to articulate and capture the idea of value in literature searching.	In Chapter 6, I set out how I would compare the conventional approach to the tailored approach, to directly address the research aims of this thesis. I base the metrics used on the systematic review that I undertook in Chapter 6 and I situate the simplified reporting in calls from researchers to investigate new ways to report evaluations of literature search effectiveness. Whilst elements of this simplified approach to reporting effectiveness evaluation are potentially unique, the extension of effectiveness evaluation to identify and capture value (by measuring the value of individual studies on the synthesis of studies and linked to this the role and contribution of the information professional) is a unique contribution to knowledge. It moves the dialogue in literature searching on from a purely quantitative understanding of effectiveness to explain the value of found. As information specialists are challenged to identify studies in new ways and establish their role in literature searching, this could be an important contribution since it explains why the studies identified matter and the contribution of the information professional.	Reported in Chapter 6 and evaluated in Chapter 7, and Chapter 8 The use of this simplified approach to effectiveness evaluation has been reported in the two published case studies (see Table 1).
7. The first use of CERQual to determine the value of studies in a qualitative synthesis	Researchers have used statistical meta-analysis to demonstrate the potential impact of 'missing' studies in literature searching for systematic reviews of intervention effectiveness (see Chapter 3). In other types of reviews, those that cannot or do not undertake statistical meta-analysis, how to demonstrate the effect of missing studies is unclear. In my second case-study (Chapter 9), I used CERQual to demonstrate the value of studies identified in a qualitative evidence synthesis and so determine the impact of potentially missing studies. This is a novel methodological contribution and it allowed my analysis to move beyond a purely quantitative finding that one search approach was superior to the other due to the number of studies identified. By including an analysis of value found, I can explain why and why this mattered. This is a significant advance in literature search evaluation since it demonstrates the value of an alternative approach to literature searching.	Reported in Chapter 8. This paper has been published (see Table 1).

10 Conclusions

In Chapter 3, I demonstrated that guidance reported in nine handbooks to the process of systematic review mandated the same pragmatic process of literature searching in different 'types' of systematic reviews. I defined this process as a conceptual model and termed it 'the conventional approach' to literature searching. I conclude that the nine leading guidance documents recommend the same process to literature searching, irrespective of the type of review, research question or studies required.

In Chapter 4, I set out a new and alternative approach to literature searching for use in systematic reviews. This model seeks to tailor the process of searching to reflect the research question, the information need, and the types of studies or study data needed. I test this process, which I call 'the tailored approach', in two case studies (Chapter 7 and Chapter 8) and I conclude that tailored literature searches can reduce the screening burden and, through the strategic allocation of time to preparing to literature search, identify studies missed by the conventional approach.

On the basis of the two case studies (reported in Chapter 7 and Chapter 8), and the three supporting methodological reviews presented in this thesis, I conclude that experienced information professionals and review teams can alter the conventional approach, to produce tailored literature searches.

The use of tailored literature searching represents a challenge to the current orthodoxy of the conventional approach. It suggests that guidance on the process of literature searching can be replaced by more streamlined, relevant and pragmatic approaches in certain contexts. I conclude that tailored approaches to literature searching may be considered more appropriate in resource (or time) limited reviews, reviews of diffuse topic areas, and reviews in which the comprehensive evaluation of effectiveness is not the primary objective of the review.

I conclude that the role of the information professional is key to this challenge. In tailored literature searching, the information professional becomes 'the decision maker,' working with(in) a research team to identify and define a new and unique approach to study identification in each and every review. This requires an understanding of not only when and how to use supplementary search methods but also, and more specifically, when and how not to use them.

The success of tailored search approaches (and the conventional approach too, for which I still see an important role) requires further development of the question 'what does effectiveness mean in literature searching?' I suggest that 'effectiveness' means different things to different researchers and that any understanding of literature search 'effectiveness' should be coupled with an understanding of the 'value' which literature searches can produce in the search for studies.

11 Towards a future research agenda

This thesis offers an original contribution to the science of information retrieval. It also suggests an agenda for future research.

11.1 Linking the research on information retrieval theory with the research on information retrieval practice

To date, broadly speaking, a separation has existed between two bodies of literature on information retrieval; one of theory (e.g. Wilson, Ellis and Kuhlthau) (4, 133, 135) and the other of practice (e.g. Paisley, Booth, Jenkins) (113, 285, 328). My work in this thesis, and the associated publications, locates the pragmatic craft of information retrieval within its theoretical landscape while making an instrumental contribution to the literature on practice.

My thesis represents a novel attempt to demonstrate that these two bodies of work can be connected in order to enhance understanding of the information retrieval process in systematic reviews. It highlights opportunities for additional work and research to further strengthen this connection.

One possible project could use systematic mapping (121) to systematically identify and map the information retrieval models, making a resource for researchers to work from as a basis for further work and practical testing of the theoretical models and their underpinning theories. The map would help by illustrating the models that exist and identify any gaps in knowledge (476). The research question for this work would be: what models of information retrieval theory exist?

11.2 Development of further case studies

I demonstrate that, as an alternative to conventional models of information retrieval, experienced information professionals can develop tailored models to better suit the information needs of systematic reviews and, potentially, reduce 'research waste' linked to the information retrieval.

Further case studies are needed to develop the evidence for alternative approaches to literature searching, and potentially develop new models of information retrieval. It is important that future case studies evaluate and report outcomes in similar ways, so as to permit generalisability between evaluation studies. A generic research question, that can be adapted into other research fields and types of systematic review, is suitable. For example: What is the effectiveness, efficiency, and value of a tailored

literature search approach compared to the conventional approach to literature searching in a qualitative evidence synthesis of diabetes self-management interventions? This question can easily be adapted by researchers and extended into other topic areas of types of systematic review. The example set out here, is under development.

11.3 Further work to update supplementary search methods

In Chapter 3, I summarised the advantages, disadvantages and resources requirements of supplementary search methods, bringing together primary studies that evaluated supplementary search methods into one review for the first time.

A finding of this review is that many of the studies meeting inclusion in the review were not recent. For example, the only study to provide data on the costs involved in handsearching was a 1995 study by Moher et al. (303). With advances in technology, access to e-resources, and inflation, the effectiveness and costs involved in handsearching have likely changed.

For researchers to make informed decisions about the use of supplementary search methods, it is important that they are working from current and reliable data. Updating the evidence available on handsearching, and the other supplementary search methods would be valuable and it represents low-hanging fruit since, in most cases, the evidence needed can be derived through judicious record keeping of day-to-day literature searching activity. This is attractive since the work could be delivered without applying for specific research funding and it might represent a suitable training project for information professionals looking to learn how to plan research projects, undertake research, and write-up research for publication.

The research questions would be straight-forward, taking handsearching as an example:

- What is the effectiveness of handsearching compared to bibliographic database searching?

- What is the efficiency of handsearching compared to bibliographic database searching?

- What are the costs involved in handsearching versus bibliographic database searching?

- And what is the value of handsearching when compared to bibliographic database searching.

I have structured these research questions in-line with my work developed in Chapter 6 and the belief that researcher's need to demonstrate the value of study identification methods in addition to showing that one approach is more effective than another.

11.4 Measuring effectiveness in literature searching

In the systematic review reported in Chapter 6, I identify that consensus is lacking on exactly what constitutes an 'effective result' in literature searching. It is timely to consider the development of this question in the context of abbreviated, accelerated or rapid reviews. I also identified that it is likely that researchers involved in evidence synthesis have differing understandings on what constitutes an effective literature search or what effectiveness means in literature searching.

A potential approach to exploring this problem (in the absence of published study data, as I found in Chapter 6) is to ask researchers what effectiveness in literature searching means to them. I gathered this data (with ethics approval) using the following research question: in literature searching, what does effective (or effectiveness in) literature searching mean to you?

Thematic analysis could be used to identify, analyse and report themes emerging from researcher replies. This work would extend the systematic review reported in Chapter 6 where data to address this question was lacking. This work is pending submission.

11.5 The professional development of information professionals

Information professionals are shown to have a vital role in the quality of the review, and in key decision making in systematic review teams. Many information professionals may be adjusting, testing and developing search methods but not reporting them and so there is little formal development of information science as an academic pursuit. This could mean that a lot of knowledge is tacit and not part of any on-going research agenda. Information professionals need to be encouraged to make this tacit knowledge explicit and show how what they do adapts and deviates from the proposed conventional approaches.

Information professionals may benefit from training, including: training on research methods, planning and undertaking research, and writing for academic audience.

References

1.   Heale R, Twycross A. What is a case study? *Evid Based Nurs*. 2018;21(1):7-8.

2.   Gerring J. What Is a Case Study and What Is It Good for? *Am Polit Sci Rev*. 2004;98(2):341-54.

3.   Centre for Reviews and Dissemination. *Systematic reviews – CRD's guidance for undertaking reviews in healthcare* [Internet]. York: Centre for Reviews and Dissemination, University of York; 2009 [Cited 2017 Dec 7]. Available from: https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf

4.   Kuhlthau C. Inside the search process: Information seeking from the user's perspective. *J Assoc Inf Sci Technol*. 1991;42(5):361-71.

5.   Cooper C, Varley-Campbell J, Booth A, Britten N, Garside R. Systematic review identifies six metrics and one method for assessing literature search effectiveness but no consensus on appropriate use. *J Clin Epidemiol*. 2018;99:53-63.

6.   Cooper C, Booth A, Britten N, Garside R. A comparison of results of empirical studies of supplementary search techniques and recommendations in review methodology handbooks: a methodological review. *Syst Rev*. 2017;6(1):234.

7.   Cooper C, Booth A, Varley-Campbell J, Britten N, Garside R. Defining the process to literature searching in systematic reviews: a literature review of guidance and supporting studies. *BMC Med Res Methodol*. 2018;18(1):85.

8.   Paisley S, Foster MJ. Innovation in information retrieval methods for evidence synthesis studies. *Res Synth Methods*. 2018;9(4):506-9.

9.   Heneghan C, Godlee F. Where next for evidence based healthcare? *BMJ*. 2013;346.

10.   Funk ME, Reid CA. Indexing consistency in MEDLINE. *Bull Med Libr Assoc*. 1983;71(2):176-83.

11.   Garfield E. Citation Analysis as a Tool in Journal Evaluation. *Science* 1972;178(4060):471-9.

12.   Bates M. The design of browsing and berrypicking techniques for the online search interface. *Online Review*. 1989;13(5):407-24.

13.   Chalmers I. The Cochrane collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care. *Ann N Y Acad Sci*. 1993;703(1):156-63

14.   Lefebvre C, Glanville J, Wieland LS, Coles B, Weightman AL. Methodological developments in searching for studies for systematic reviews: past, present and future? *Syst Rev*. 2013;2:78.

15.   Cochrane A. *Effectiveness and Efficiency. Random Reflections on Health Services.* London: Nuffield Provincial Hospitals Trust; 1972.

16.   Cochrane A. 1931-1971: A critical review, with particular reference to the medical profession. In: Teeling-Smith G, Wells N. (eds.). *Medicines for the year 2000*. London: Office of Health Economics; 2000. p. 1-11.

17.   Green S, Higgins JPT, Alderson P, Clarke M, Mulrow CD, Oxman AD. Chapter 1: Introduction. In: Higgins JPT, Green S, (eds.) *Cochrane Handbook for Systematic Reviews of*

Interventions. Version 5.2.0 (updated March 2011) [Internet]. The Cochrane Collaboration; 2011 [Cited 2017 Dec 7]. Available from: www.handbook.cochrane.org

18.   Crowley P, Chalmers I, Keirse MJ. The effects of corticosteroid administration before preterm delivery: an overview of the evidence from controlled trials. *BJOG*. 1990;97(1):11-25.

19.   The Cochrane Collaboration. *'Our logo tells a story'* [Internet]. Oxford: The Cochrane Collaboration; 2017 [Cited 2017 Dec 7]. Available from: https://www.cochrane.org/about-us/difference-we-make

20.   Becker L. The Cochrane Collaboration - Valuable resource for family physicians. *Can Fam Physician*. 1997;43:403-4.

21.   Haugh MC, Boissel JP, Pignon JP, Chalmers I. The Cochrane collaboration: The need for international cooperation. *Therapie*. 1996;51(3):253-6.

22.   Leng GC, Jepson RG, Fowkes FG. Why we need the Cochrane Collaboration. *Int Angiol*. 1996;15(1):81-3.

23.   Chalmers I, Dickersin K, Chalmers TC. Getting to grips with Archie Cochrane's agenda. *BMJ*. 1992;305(6857):786-8.

24.   Lefebvre C. The Cochrane Collaboration: the role of the UK Cochrane Centre in identifying the evidence. *Health Libr Rev*. 1994;11(4):235-42.

25.   Lefebvre C, Manheimer E, Glanville J. Chapter 6: Searching for studies. In: Higgins JPT, Green S, (eds.) *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0 (updated March 2011) [Internet]. The Cochrane Collaboration; 2011 [Cited 2017 Dec 7]. Available from: http://handbook.cochrane.org/

26.   Stewart LA, Parmar MKB. Bias in the analysis and reporting of randomized controlled trials. *Int J Technol Assess Health Care*. 1996;12(2):264-75.

27.   Ahmed I, Sutton AJ, Riley RD. Assessment of publication bias, selection bias, and unavailable data in meta-analyses using individual participant data: a database survey. *BMJ*. 2012;344.

28.   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):1-76.

29.   Moher D, Fortin P, Jadad AR, Juni P, Klassen T, LeLorier 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(8998):363-6.

30.   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(9074):326-9.

31.   Bernstein F. The retrieval of randomized clinical trials in liver diseases from the medical literature: manual versus MEDLARS searches. *Control Clin Trials*. 1988;9(1):23-31.

32.   Poynard T, Conn HO. The retrieval of randomized clinical trials in liver disease from the medical literature. A comparison of MEDLARS and manual methods. *Control Clin Trials*. 1985;6(4):271-9.

33.   Lefebvre C. *Identification of randomized controlled trials using MEDLINE: the situation in 1993*. Bethesda, MD: National Institutes of Health, Office of Medical Applications of Research; 1993.

34.   Hyde CJ. Using the evidence - A need for quantity, not quality? *Int J Technol Assess Health Care*. 1996;12(2):280-7.

35.   Adams CE, Frederick K. An investigation of the adequacy of MEDLINE searches for randomized controlled trials (RCTs) of the effects of mental health care. *Psychol Med*. 1994;24(3):741-8.

36.   Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *BMJ*. 1994;309(6964):1286-91.

37.   Lefebvre C, Eisinga A, McDonald S, Paul N. Enhancing access to reports of randomized trials published world-wide – the contribution of EMBASE records to the Cochrane Central Register of Controlled Trials (CENTRAL) in The Cochrane Library. *Emerg Themes Epidemiol*. 2008;5:13.

38.   Meinert CL. Terminology - a plea for standardization. *Control Clin Trials*. 1980;1(2):97-9.

39.   Wieland S, Dickersin K. Why were they missed? Randomized controlled trials (RCTs) identified through the MEDLINE Retagging Project but not the US National Library of Medicine (NLM). Paper presented at: *The 16th Cochrane Colloquium;* 2008 Oct 3-7; Freiburg, Germany.

40.   Dickersin K, Manheimer E, Wieland S, Robinson KA, Lefebvre C, McDonald S. Development of the Cochrane Collaboration's CENTRAL Register of controlled clinical trials. *Eval Health Prof*. 2002;25(1):38-64.

41.   McDonald SJ, Lefebvre C, Clarke MJ. Identifying reports of controlled trials in the BMJ and the Lancet. *BMJ*. 1996;313(7065):1116-7.

42.   Glanville JM, Lefebvre C, Miles JNV, Camosso-Stefinovic J. How to identify randomized controlled trials in MEDLINE: ten years on. *J Med Libr Assoc*. 2006;94(2):130-6.

43.   Klemenz BE. Cochrane databases - Evaluation of randomized controlled studies as the basis of progressive clinical medicine. *Deutsche Medizinische Wochenschrift*. 1997;122(12):384-5.

44.   Farquhar C, Vandekerckhove P. The Cochrane library. *Baillieres Clin Obstet Gynaecol*. 1996;10(4):569-83.

45.   Jadad AR, Carroll D, Moore A, McQuay H. Developing a database of published reports of randomised clinical trials in pain research. *Pain*. 1996;66(2-3):239-46.

46.   Pignon JP. Interest of the Cochrane collaboration: The point of view of a meta-analysis practitioner. *Therapie*. 1996;51(3):257-60.

47.   Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA*. 1996;276(8):637-9.

48.   Royle P, Milne R. Literature searching for randomized controlled trials used in Cochrane reviews: rapid versus exhaustive searches. *Int J Technol Assess Health Care*. 2003;19(4):591-603.

49.   Clunie G, Ludbrook J, Faris I. The reporting of randomized controlled clinical trials and The Cochrane Collaboration. *Aust N Z J Surg*. 1995;65(7):461-2.

50.   Cooper C, Levay P, Lorenc T, Craig GM. A population search filter for hard-to-reach populations increased search efficiency for a systematic review. *J Clin Epidemiol*. 2014;67(5):554-9.

51.   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(1):150-3.

52.   Haynes RB, Wilczynski N, McKibbon KA, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. *J Am Med Inform Assoc*. 1994;1(6):447-58.

53.   Watson RJ, Richardson PH. Identifying randomized controlled trials of cognitive therapy for depression: comparing the efficiency of Embase, Medline and PsycINFO bibliographic databases. *Br J Med Psychol*. 1999;72 (Pt 4):535-42.

54.   Marson AG, Chadwick DW. How easy are randomized controlled trials in epilepsy to find on Medline? The sensitivity and precision of two Medline searches. *Epilepsia*. 1996;37(4):377-80.

55.   Royle P, Waugh N. A simplified search strategy for identifying randomised controlled trials for systematic reviews of health care interventions: a comparison with more exhaustive strategies. *BMC Med Res Methodol*. 2005;5:23.

56.   Royle P, Bain L, Waugh N. Systematic reviews of epidemiology in diabetes: finding the evidence. *BMC Med Res Methodol*. 2005;5:2.

57.   Bannach-Brown A, Przybyła P, Thomas J, Rice ASC, Ananiadou S, Liao J, et al. The use of text-mining and machine learning algorithms in systematic reviews: reducing workload in preclinical biomedical sciences and reducing human screening error. *bioRxivorg*. 2018.

58.   Gates A, Johnson C, Hartling L. Technology-assisted title and abstract screening for systematic reviews: a retrospective evaluation of the Abstrackr machine learning tool. *Syst Rev*. 2018;7(1):45.

59.   O'Mara-Eves A, Thomas J, McNaught J, Miwa M, Ananiadou S. Using text mining for study identification in systematic reviews: a systematic review of current approaches. *Syst Rev*. 2015;4(5).

60.   Stansfield C, O'Mara-Eves A, Thomas J. Text mining for search term development in systematic reviewing: A discussion of some methods and challenges. *Res Synth Methods*. 2017;8(3):355-65.

61.   Thomas J, McNaught J, Ananiadou S. Applications of text mining within systematic reviews. *Res Synth Methods*. 2011;2(1):1-14.

62.   Marshall IJ, Noel-Storr A, Kuiper J, Thomas J, Wallace BC. Machine learning for identifying Randomized Controlled Trials. *Res Synth Methods*. 2018;9(4):602-14.

63.   Shemilt I, Simon A, Hollands GJ, Marteau TM, Ogilvie D, O'Mara-Eves A, et al. Pinpointing needles in giant haystacks: use of text mining to reduce impractical screening workload in extremely large scoping reviews. *Res Synth Methods*. 2014;5(1):31-49.

64.   Doshi P, Jefferson T. Clinical study reports of randomised controlled trials: an exploratory review of previously confidential industry reports. *BMJ Open*. 2013;3(2).

65.   Jefferson T, Doshi P. Multisystem failure: the story of anti-influenza drugs. *BMJ*. 2014;348.

66.   Jefferson T, Jones M, Doshi P, Spencer EA, Onakpoya I, Heneghan CJ. Oseltamivir for influenza in adults and children: systematic review of clinical study reports and summary of regulatory comments. *BMJ*. 2014;348.

67.   Goldacre B. What the Tamiflu saga tells us about drug trials and big pharma. The Guardian [Internet]. 2014 Apr 10 [Cited 2017 Dec 7]; Business. Available from: https://www.theguardian.com/business/2014/apr/10/tamiflu-saga-drug-trials-big-pharma

68.   Jefferson T, Doshi P. Drug Data Shouldn't Be Secret. New York Times [internet]. 2012 Apr 10 [Cited 2017 Dec 7]; Opinion. Available from: https://www.nytimes.com/2012/04/11/opinion/drug-data-shouldnt-be-secret.html

69.   Hodkinson A, Dietz KC, Lefebvre C, Golder S, Jones M, Doshi P, et al. The use of clinical study reports to enhance the quality of systematic reviews: a survey of systematic review authors. *Syst Rev*. 2018;7(1):117.

70.   Anderson LM, Oliver SR, Michie S, Rehfuess E, Noyes J, Shemilt I. Investigating complexity in systematic reviews of interventions by using a spectrum of methods. *J Clin Epidemiol*. 2013;66(11):1223-9.

71.   Gough D, Thomas J, Oliver S. Clarifying differences between review designs and methods. *Syst Rev*. 2012;1(1):28.

72.   Munn Z, Stern C, Aromataris E, Lockwood C, Jordan Z. What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. *BMC Med Res Methodol*. 2018;18(1):5.

73.   Noyes J, Gough D, Lewin S, Mayhew A, Michie S, Pantoja T, et al. A research and development agenda for systematic reviews that ask complex questions about complex interventions. *J Clin Epidemiol*. 2013;66(11):1262-70.

74.   Petticrew M. Time to rethink the systematic review catechism? Moving from 'what works' to 'what happens'. *Syst Rev*. 2015;4(1):36.

75.   Stansfield C, Brunton G, Rees R. Search wide, dig deep: literature searching for qualitative research. An analysis of the publication formats and information sources used for four systematic reviews in public health. *Res Synth Methods*. 2014;5(2):142-51.

76.   Stansfield C, Kavanagh J, Rees R, Gomersall A, Thomas J. The selection of search sources influences the findings of a systematic review of people's views: a case study in public health. *BMC Med Res Methodol*. 2012;12:55.

77.   Cooper C, Lovell R, Husk K, Booth A, Garside R. Supplementary search methods were more effective and offered better value than bibliographic database searching: A case study from public health and environmental enhancement. *Res Synth Methods*. 2018;9(2):195-223.

78.   Cooper C, Rogers M, Bethel A, Briscoe S, Lowe J. A mapping review of the literature on UK-focused health and social care databases. *Health Info Libr J*. 2015;32(1):5-22.

79.   Levay P, Ainsworth N, Kettle R, Morgan A. Identifying evidence for public health guidance: a comparison of citation searching with Web of Science and Google Scholar. *Res Synth Methods*. 2016;7(1):34-45.

80.   Levay P, Raynor M, Tuvey D. The Contributions of MEDLINE, Other Bibliographic Databases and Various Search Techniques to NICE Public Health Guidance. *2015*. 2015;10(1):19.

81.   Petticrew M, Roberts H. *Systematic Reviews in the Social Sciences: A Practical Guide*. Oxford: Blackwell Publishing Ltd; 2006.

82.   Bayliss SE, Davenport CF, Pennant ME. Where and how to search for information on the effectiveness of public health interventions – a case study for prevention of cardiovascular disease. *Health Info Libr J*. 2014;31(4):303-13.

83.   Geersing GJ, Bouwmeester W, Zuithoff P, Spijker R, Leeflang M, Moons K. Search filters for finding prognostic and diagnostic prediction studies in medline to enhance systematic reviews. *PLoS One*. 2012;7(2).

84.   Leeflang MMG, Scholten RJPM, Rutjes AWS, Reitsma JB, Bossuyt PMM. Use of methodological search filters to identify diagnostic accuracy studies can lead to the omission of relevant studies. *J Clin Epidemiol*.59(3):234-40.

85.   Mann R, Gilbody SM. Should methodological filters for diagnostic test accuracy studies be used in systematic reviews of psychometric instruments? a case study involving screening for postnatal depression. *Syst Rev*. 2012;1(1):9.

86.   Whiting P, Westwood M, Beynon R, Burke M, Sterne JAC, Glanville J. Inclusion of methodological filters in searches for diagnostic test accuracy studies misses relevant studies. *J Clin Epidemiol*. 2011;64(6):602-7.

87.   Ritchie G, Glanville J, Lefebvre C. Do published search filters to identify diagnostic test accuracy studies perform adequately? *Health Info Libr J*. 2007;24(3):188-92.

88.   Adams J, Hillier-Brown FC, Moore HJ, Lake AA, Araujo-Soares V, White M, et al. Searching and synthesising 'grey literature' and 'grey information' in public health: critical reflections on three case studies. *Syst Rev*. 2016;5(1):164.

89.   Bayliss HR, Beyer FR. Information retrieval for ecological syntheses. *Res Synth Methods*. 2015;6(2):136-48.

90.   Bayliss SE, Davenport CF, Pennant ME. Where and how to search for information on the effectiveness of public health interventions--a case study for prevention of cardiovascular disease. *Health Info Libr J*. 2014;31(4):303-13.

91.   Beahler CC, Sundheim JJ, Trapp NI. Information retrieval in systematic reviews: challenges in the public health arena. *Am J Prev Med*. 2000;18(4 Suppl):6-10.

92.   Betrán AP, Say L, Gülmezoglu AM, Allen T, Hampson L. Effectiveness of different databases in identifying studies for systematic reviews: Experience from the WHO systematic review of maternal morbidity and mortality. *BMC Med Res Methodol*. 2005;5.

93.   Mahood Q, Van Eerd D, Irvin E. Searching for grey literature for systematic reviews: challenges and benefits. *Res Synth Methods*. 2014;5(3):221-34.

94.   Ogilvie D, Hamilton V, Egan M, Petticrew M. Systematic reviews of health effects of social interventions: 1. Finding the evidence: how far should you go? *J Epidemiol Community Health*. 2005;59(9):804-8.

95.   Papaioannou D, Sutton A, Carroll C, Booth A, Wong R. Literature searching for social science systematic reviews: consideration of a range of search techniques. *Health Info Libr J*. 2010;27(2):114-22.

96.   Rogers M, Bethel A, Abbott R. Locating qualitative studies in dementia on MEDLINE, EMBASE, CINAHL, and PsycINFO: A comparison of search strategies. *Res Synth Methods*. 2018;9(4).

97.   Baeza-Yates R. Information retrieval in the Web: beyond current search engines. *Int J Approx Reason*. 2003;34(2):97-104.

98.   Beckles Z, Glover S, Ashe J, Stockton S, Boynton J, Lai R, et al. Searching CINAHL did not add value to clinical questions posed in NICE guidelines. *J Clin Epidemiol*. 2013;66(9):1051-7.

99.   Beyer FR, Wright K. Can we prioritise which databases to search? A case study using a systematic review of frozen shoulder management. *Health Info Libr J*. 2013;30(1):49-58.

100.   Blakeman K. Finding research information on the web: how to make the most of Google and other free search tools. *Sci Prog*. 2013;96(Pt 1):61-84.

101.   Booth A. Unpacking your literature search toolbox: on search styles and tactics. *Health Info Libr J*. 2008;25(4):313-7.

102.   Choong MK, Tsafnat G. Role of Citation Tracking in Updating of Systematic Reviews. *AMIA Jt Summits Transl Sci Proc* 2014;2014:18.

103.   Falagas ME, Pitsouni EI, Malietzis GA, Pappas G. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *Faseb j*. 2008;22(2):338-42.

104.   Glanville J, Cikalo M, Crawford F, Dozier M, McIntosh H. Handsearching did not yield additional unique FDG-PET diagnostic test accuracy studies compared with electronic searches: a preliminary investigation. *Res Synth Methods*. 2012;3(3):202-13.

105.   Kloda LA. Use Google Scholar, Scopus and Web of Science for Comprehensive Citation Tracking. *Evid Based Libr Inf Pract*. 2007;2(3):87-90.

106.   O'Mara-Eves A, Brunton G, McDaid D, Kavanagh J, Oliver S, Thomas J. Techniques for identifying cross-disciplinary and 'hard-to-detect' evidence for systematic review. *Res Synth Methods*. 2014;5(1):50-9.

107.   Wright K, Golder S, Lewis-Light K. What value is the CINAHL database when searching for systematic reviews of qualitative studies? *Syst Rev*. 2015;4:104.

108.   Young T, Hopewell S. Methods for obtaining unpublished data. *Cochrane Database Syst Rev*. 2011;9(11).

109.   Younger P, Boddy K. When is a search not a search? A comparison of searching the AMED complementary health database via EBSCOhost, OVID and DIALOG. *Health Info Libr J*. 2009;26(2):126-35.

110.   Booth A. State of the Art, State of the Science? *Res Synth Methods*. 2018;9(4):615-8.

111.   Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *BMJ*. 2005;331.

112.   Lorenc T, Pearson M, Jamal F, Cooper C, Garside R. The role of systematic reviews of qualitative evidence in evaluating interventions: a case study. *Res Synth Methods*. 2012;3(1):1-10.

113. Booth A. How much searching is enough? Comprehensive versus optimal retrieval for technology assessments. *Int J Technol Assess Health Care*. 2010;26(4):431-5.

114. Booth A. Searching for qualitative research for inclusion in systematic reviews: a structured methodological review. *Syst Rev*. 2016;5:74.

115. Collaboration for Environmental Evidence. *Guidelines for Systematic Review and Evidence Synthesis in Environmental Management* [Internet]. Bangor: Environmental Evidence; 2013 [Cited 2017 Dec 7]. Available from: http://www.environmentalevidence.org/wp-content/uploads/2017/01/Review-guidelines-version-4.2-final-update.pdf

116. The Joanna Briggs Institute. *Joanna Briggs Institute Reviewers' Manual: 2014 edition* [Internet]. Adelaide: The Joanna Briggs Institute; 2014 [Cited 2017 Dec 7]. Available from: https://joannabriggs.org/assets/docs/sumari/ReviewersManual-2014.pdf

117. Institute for Quality and Efficiency in Health Care (IQWiG). *IQWiG Methods Resources. 7 Information retrieval* [Internet]. Berlin: Institute for Quality and Efficiency in Health Care (IQWiG); 2014 [Cited 2017 Dec 7]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK385787/

118. EUnetHTA: European Network for Health Technology Assessment. *Process of information retrieval for systematic reviews and health technology assessments on clinical effectiveness* [Internet]. Netherlands: EUnetHTA: European Network for Health Technology Assessment; 2016 [Cited 2017 Dec 7]. Available from: http://www.eunethta.eu/sites/default/files/Guideline_Information_Retrieval_V1-1.pdf

119. Kugley S WA, Thomas J, Mahood Q, Jørgensen AMK, Hammerstrøm K, Sathe N. *Searching for studies: a guide to information retrieval for Campbell systematic reviews* [Internet]. Oslo: Campbell Collaboration; 2017 [Cited 2017 Dec 7]. Available from: https://www.campbellcollaboration.org/library/searching-for-studies-information-retrieval-guide-campbell-reviews.html

120. National Institute for Health Care Excellence (NICE). *Developing NICE guidelines: the manual* [Internet]. London: National Institute for Health Care Excellence (NICE); 2014 [Cited 2017 Dec 7]. Available from: https://www.nice.org.uk/media/default/about/what-we-do/our-programmes/developing-nice-guidelines-the-manual.pdf

121. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Info Libr J*. 2009;26(2):91-108.

122. Sampson M, McGowan J, Lefebvre C, Moher D, Grimshaw J. *PRESS: Peer Review of Electronic Search Strategies* [Internet]. Toronto, ON: Canadian Agency for Drugs and Technologies in Health; 2008 [Cited 2017 Dec 7]. Available from: https://www.cadth.ca/media/pdf/477_PRESS-Peer-Review-Electronic-Search-Strategies_tr_e.pdf

123. Kneale D, Thomas J, Harris K. Developing and Optimising the Use of Logic Models in Systematic Reviews: Exploring Practice and Good Practice in the Use of Programme Theory in Reviews. *PLoS One*. 2015;10(11):e0142187.

124. Kaltenthaler E, Tappenden P, Paisley S, Squires H. *NICE DSU Technical Support Document 13: Identifying and reviewing evidence to inform the conceptualisation and population of cost-effectiveness models* [Internet]. London: NICE Decision Support Unit; 2011 [Cited 2017 Dec 7]. Available from: http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2016/03/TSD-13-model-parameters.pdf

125. Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M. Conceptualizing a Model: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force–2. *Med Decis Making*. 2012;32(5):678-89.

126. Koffel JB. Use of Recommended Search Strategies in Systematic Reviews and the Impact of Librarian Involvement: A Cross-Sectional Survey of Recent Authors. *PLoS One*. 2015;10(5):e0125931.

127. Cole C. A theory of information need for information retrieval that connects information to knowledge. *J Assoc Inf Sci Technol*. 2011;62(7):1216-31.

128. Rautenberg T, Hulme C, Edlin R. Methods to construct a step-by-step beginner's guide to decision analytic cost-effectiveness modeling. *Clinicoecon Outcomes Res*. 2016;8:573-81.

129. Sonnenberg FA, Roberts MS, Tsevat J, Wong JB, Barry M, Kent DL. Toward a Peer Review Process for Medical Decision Analysis Models. *Medical Care*. 1994;32(7):JS52-JS64.

130. Taylor RS. The process of asking questions. *American Documentation*. 1962;13(4):391-6.

131. Belkin N. Anomalous states of knowledge as a basis for information retrieval. *J Can Health Libr Assoc*. 1980;5:133-43.

132. Belkin N, Oddy R, Brooks H. ASK for information retrieval: part 1. Background and theory. *J Doc*. 1982;38(2).

133. Wilson TD. On user studies and information needs. *J Doc*. 1981;37(1):3-15.

134. Dervin B. An overview of sense-making research: concepts, methods, and results to date. Paper presented at: *The International Communication Association Annual Meeting;* 1983; Dallas, Texas.

135. Ellis D. A behavioural approach to information retrieval system design. *J Doc*. 1989;45(3):171-212.

136. Ellis D. Modeling the information-seeking patterns of academic researchers - a grounded theory approach. *Libr Q*. 1993;63(4):469-86.

137. Bystrom K, Hansen P. Work tasks as units for analysis in information seeking and retrieval studies. Paper presented at: *Colis4: Emerging Frameworks and Methods;* 2002 Jul 21-25; Seattle, Washington, USA.

138. Du T. The information journey of marketing professionals: Incorporating work task-driven information seeking, information judgments, information use, and information sharing. *J Assoc Inf Sci Technol*. 2014;65(9):1850-69.

139. Beverley CA, Booth A, Bath PA. The role of the information specialist in the systematic review process: a health information case study. *Health Info Libr J*. 2003;20(2):65-74.

140. Cooper ID, Crum JA. New activities and changing roles of health sciences librarians: a systematic review, 1990-2012. *J Med Libr Assoc*. 2013;101(4):268-77.

141. Dudden RF, Protzko SL. The systematic review team: contributions of the health sciences librarian. *Med Ref Serv Q*. 2011;30(3):301-15.

142.   Foster M. An overview of the role of librarians in systematic reviews: from expert search to project manager. *J Eur Assoc Health Inf Libr*. 2015;11(3):3-7.

143.   Harris MR. The librarian's roles in the systematic review process: a case study. *J Med Libr Assoc*. 2005;93(1):81-7.

144.   Lawson L. *Operating outside library walls* [Internet]. Winchester: Winchester and Eastleigh Healthcare NHS Trust; 2004 [Cited 2017 Dec 7]. Available from: http://www.hantshealthcarelibrary.nhs.uk/assets/files/CL%20EVALUATION%20REPORT%202004.pdf

145.   Li L, Tian J, Tian H, Moher D, Liang F, Jiang T, et al. Network meta-analyses could be improved by searching more sources and by involving a librarian. *J Clin Epidemiol*. 2014;67(9):1001-7.

146.   McGowan J, Sampson M. Systematic reviews need systematic searchers. *J Med Libr Assoc*. 2005;93(1):74-80.

147.   Rethlefsen ML, Farrell AM, Osterhaus Trzasko LC, Brigham TJ. Librarian co-authors correlated with higher quality reported search strategies in general internal medicine systematic reviews. *J Clin Epidemiol*. 2015;68(6):617-26.

148.   Swinkels A, Briddon J, Hall J. Two physiotherapists, one librarian and a systematic literature review: collaboration in action. *Health Info Libr J*. 2006;23(4):248-56.

149.   Townsend WA, Anderson PF, Ginier EC, MacEachern MP, Saylor KM, Shipman BL, et al. A competency framework for librarians involved in systematic reviews. *J Med Libr Assoc*. 2017;105(3):268-75.

150.   Vassar M, Yerokhin V, Sinnett PM, Weiher M, Muckelrath H, Carr B, et al. Database selection in systematic reviews: an insight through clinical neurology. *Health Info Libr J*. 2017;34(2):156-64.

151.   Weller AC. Mounting evidence that librarians are essential for comprehensive literature searches for meta-analyses and Cochrane reports. *J Med Libr Assoc*. 2004;92(2):163-4.

152.   Crum JA, Cooper ID. Emerging roles for biomedical librarians: a survey of current practice, challenges, and changes. *J Med Libr Assoc*. 2013;101(4):278-86.

153.   Golder S, Loke Y, McIntosh HM. Poor reporting and inadequate searches were apparent in systematic reviews of adverse effects. *J Clin Epidemiol*. 2008;61(5):440-8.

154.   Maggio LA, Tannery NH, Kanter SL. Reproducibility of literature search reporting in medical education reviews. *Acad Med*. 2011;86(8):1049-54.

155.   Meert D, Torabi N, Costella J. Impact of librarians on reporting of the literature searching component of pediatric systematic reviews. *J Med Libr Assoc*. 2016;104(4):267-77.

156.   Morris M, Boruff JT, Gore GC. Scoping reviews: establishing the role of the librarian. *J Med Libr Assoc*. 2016;104(4):346-54.

157.   Koffel JB, Rethlefsen ML. Reproducibility of Search Strategies Is Poor in Systematic Reviews Published in High-Impact Pediatrics, Cardiology and Surgery Journals: A Cross-Sectional Study. *PLoS One.* 2016;11(9):e0163309.

158.   Brettle A. The librarian within research evidence-based practice. In: Brettle A, Urquhart C, (eds.). *Changing roles and contexts for health library and information professionals*. London: Facet; 2012. p. 135-60.

159.   Gray M. *Evidence-based Healthcare*. London: Churchill-Livingstone; 2001.

160.   Andrew B. Clear and present questions: formulating questions for evidence based practice. *Library Hi Tech*. 2006;24(3):355-68.

161.   Grewal A, Kataria H, Dhawan I. Literature search for research planning and identification of research problem. *Indian J Anaesth*. 2016;60(9):635-9.

162.   Craven J, & Levay, P. Recording Database Searches for Systematic Reviews - What is the Value of Adding a Narrative to Peer-Review Checklists? A Case Study of NICE Interventional Procedures Guidance. *Evid Based Libr Inf Pract*. 2011;6(4):72-87.

163.   Cooper C, Dawson S, Peters J, Varley-Campbell J, Cockcroft E, Hendon J, et al. Revisiting the need for a literature search narrative: A brief methodological note. *Res Synth Methods*. 2018;9(3):361-5.

164.   Fehrmann P, Thomas J. Comprehensive computer searches and reporting in systematic reviews. *Res Synth Methods*. 2011;2(1):15-32.

165.   Tricco AC, Tetzlaff J, Sampson M, Fergusson D, Cogo E, Horsley T, et al. Few systematic reviews exist documenting the extent of bias: a systematic review. *J Clin Epidemiol*. 2008;61(5):422-34.

166.   Felson DT. Bias in meta-analytic research. *J Clin Epidemiol*. 1992;45(8):885-92.

167.   Franco A, Malhotra N, Simonovits G. Publication bias in the social sciences: Unlocking the file drawer. *Science*. 2014;345(6203):1502-5.

168.   Hartling L, Featherstone R, Nuspl M, Shave K, Dryden DM, Vandermeer B. Grey literature in systematic reviews: a cross-sectional study of the contribution of non-English reports, unpublished studies and dissertations to the results of meta-analyses in child-relevant reviews. *BMC Med Res Methodol*. 2017;17(1):64.

169.   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(41):1-90.

170.   Schmucker CM, Blümle A, Schell LK, Schwarzer G, Oeller P, Cabrera L, et al. Systematic review finds that study data not published in full text articles have unclear impact on meta-analyses results in medical research. *PLoS One*. 2017;12(4):e0176210.

171.   Mills EJ, Kanters S, Thorlund K, Chaimani A, Veroniki A-A, Ioannidis JPA. The effects of excluding treatments from network meta-analyses: survey. *BMJ*. 2013;347.

172.   Hartling L, Featherstone R, Nuspl M, Shave K, Dryden DM, Vandermeer B. The contribution of databases to the results of systematic reviews: a cross-sectional study. *BMC Med Res Methodol*. 2016;16(1):127.

173.   van Driel ML, De Sutter A, De Maeseneer J, Christiaens T. Searching for unpublished trials in Cochrane reviews may not be worth the effort. *J Clin Epidemiol*. 2009;62(8):838-44.e3.

174.   Buchberger B, Krabbe L, Lux B, Mattivi JT. Evidence mapping for decision making: feasibility versus accuracy - when to abandon high sensitivity in electronic searches. *German medical science*. 2016;14:Doc09.

175.   Barroso J, Gollop CJ, Sandelowski M, Meynell J, Pearce PF, Collins LJ. The challenges of searching for and retrieving qualitative studies. *West J Nurs Res*. 2003;25(2):153-78.

176.   Booth A, Carroll C. Systematic searching for theory to inform systematic reviews: is it feasible? Is it desirable? *Health Info Libr J*. 2015;32(3):220-35.

177.   Britten N, Garside R, Pope C, Frost J, Cooper C. Asking More of Qualitative Synthesis: A Response to Sally Thorne. *Qual Health Res*. 2017;27(9):1370-6.

178.   Gough D. Weight of Evidence: a framework for the appraisal of the quality and relevance of evidence. *Res Pap Educ*. 2007;22(2):213-28.

179.   Kwon Y, Powelson SE, Wong H, Ghali WA, Conly JM. An assessment of the efficacy of searching in biomedical databases beyond MEDLINE in identifying studies for a systematic review on ward closures as an infection control intervention to control outbreaks. *Syst Rev*. 2014;3:135.

180.   Nussbaumer-Streit B, Klerings I, Wagner G, Heise TL, Dobrescu AI, Armijo-Olivo S, et al. Abbreviated literature searches were viable alternatives to comprehensive searches: a meta-epidemiological study. *J Clin Epidemiol*. 2018;102:1-11.

181.   Wagner G, Nussbaumer-Streit B, Greimel J, Ciapponi A, Gartlehner G. Trading certainty for speed - how much uncertainty are decisionmakers and guideline developers willing to accept when using rapid reviews: an international survey. *BMC Med Res Methodol*. 2017;17(1):121.

182.   Pearson M, Moxham T, Ashton K. Effectiveness of Search Strategies for Qualitative Research About Barriers and Facilitators of Program Delivery. *Eval Health Prof*. 2011;34(3):297-308.

183.   McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. *J Clin Epidemiol*. 2016;75:40-6.

184.   National Institute for Health Care Excellence (NICE). *Methods for the development of NICE public health guidance* [Internet]. London: National Institute for Health Care Excellence (NICE); 2009 [Cited 2017 Dec 7]. Available from: https://www.nice.org.uk/process/pmg4/chapter/introduction

185.   Cooper C, Gomersall A. WS68: Database selection bias and its affect on systematic reviews: A United Kingdom perspective. Paper presented at: *Joint Cochrane and Campbell Colloquium;* 2010 Oct 18-22; Keystone, Colorado, USA.

186.   Haddaway NR, Westgate M. Predicting the time needed to conduct an environmental systematic review or systematic map: analysis and decision support tool. *bioRxivorg*. 2018.

187.   Bullers K, Howard AM, Hanson A, Kearns WD, Orriola JJ, Polo RL, et al. It takes longer than you think: librarian time spent on systematic review tasks. *J Med Libr Assoc*. 2018;106(2):10.

188.   Bethel A. Search summary tables for systematic reviews: results and findings. Paper presented at: *Chartered Institute of Library and Information Professionals (CILIP) Health Libraries Group Conference;* 2016 Sep 15-16; Scarborough, UK.

189.   Flach P, Kull M. Precision-Recall-Gain Curves: PR Analysis Done Right. Paper presented at: *The Advances in Neural Information Processing Systems 28 (NIPS 2015);* 2015 Dec 7-12; Montreal, Canada.

190.   Sampson M, McGowan J. Inquisitio validus Index Medicus: A simple method of validating MEDLINE systematic review searches. *Res Synth Methods.* 2011;2(2):103-9.

191.   Booth A. "Brimful of STARLITE": toward standards for reporting literature searches. *J Med Libr Assoc.* 2006;94(4):421-e205.

192.   Beale S, Duffy S, Glanville J, Lefebvre C, Wright D, McCool R, et al. Choosing and using methodological search filters: searchers' views. *Health Info Libr J.* 2014;31(2):133-47.

193.   Lefebvre C, Glanville J, Beale S, Boachie C, Duffy S, Fraser C, et al. Assessing the performance of methodological search filters to improve the efficiency of evidence information retrieval: five literature reviews and a qualitative study. *Health Technol Assess.* 2017;21(69):1-148.

194.   Atkinson KM, Koenka AC, Sanchez CE, Moshontz H, Cooper H. Reporting standards for literature searches and report inclusion criteria: making research syntheses more transparent and easy to replicate. *Res Synth Methods.* 2015;6(1):87-95.

195.   Rader T, Mann M, Stansfield C, Cooper C, Sampson M. Methods for documenting systematic review searches: a discussion of common issues. *Res Synth Methods.* 2014;5(2):98-115.

196.   Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *J Clin Epidemiol.* 2009;62:1006-12.

197.   Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and Reporting Characteristics of Systematic Reviews. *PLoS Med.* 2007;4(3):e78.

198.   Delaney A, Bagshaw SM, Ferland A, Manns B, Laupland KB, Doig CJ. A systematic evaluation of the quality of meta-analyses in the critical care literature. *Crit Care.* 2005;9(5):R575.

199.   Witkowski MA, Aldhouse N. Transparency and Reproducibility of Supplementary Search Methods in Nice Single Technology Appraisal Manufacturer Submissions. *Value in Health.* 2015;18(7):A721-A2.

200.   Briscoe S. Web searching for systematic reviews: a case study of reporting standards in the UK Health Technology Assessment programme. *BMC Res Notes.* 2015;8:153.

201.   Briscoe S. A review of the reporting of web searching to identify studies for Cochrane systematic reviews. *Res Synth Methods.* 2018;9(1):89-99.

202.   Haddaway NR, Collins AM, Coughlin D, Kirk S. A rapid method to increase transparency and efficiency in web-based searches. *Environmental Evidence.* 2017;6(1):1.

203.   Briscoe S, Cooper C, Glanville J, Lefebvre C. The loss of the NHS EED and DARE databases and the effect on evidence synthesis and evaluation. *Res Synth Methods.* 2017;8(3):256-7.

204.   Stansfield C, Thomas J, Kavanagh J. 'Clustering' documents automatically to support scoping reviews of research: a case study. *Res Synth Methods.* 2013;4(3):230-41.

205.   Petrova M, Sutcliffe P, Fulford KW, Dale J. Search terms and a validated brief search filter to retrieve publications on health-related values in Medline: a word frequency analysis study. *J Am Med Inform Assoc*. 2012;19(3):479-88.

206.   Colquhoun HL, Levac D, O'Brien KK, Straus S, Tricco AC, Perrier L, et al. Scoping reviews: time for clarity in definition, methods, and reporting. *J Clin Epidemiol*. 2014;67(12):1291-4.

207.   Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol*. 2005;8(1):19-32.

208.   Armstrong R, Hall BJ, Doyle J, Waters E. Cochrane Update. 'Scoping the scope' of a cochrane review. *J Public Health (Oxf)*. 2011;33(1):147-50.

209.   Cooper C. *Planning and Budgeting for Systematic Reviews* [Lecture]. Freiburg: The German Cochrane Centre; 2017 [Cited 2017 Dec 7]. Available from: www.cochrane.de/lit_vortrag_planning_and_budgeting

210.   Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci*. 2010;5(1):69.

211.   O'Brien KK, Colquhoun H, Levac D, Baxter L, Tricco AC, Straus S, et al. Advancing scoping study methodology: a web-based survey and consultation of perceptions on terminology, definition and methodological steps. *BMC Health Serv Res*. 2016;16(1):305.

212.   Anderson S, Allen P, Peckham S, Goodwin N. Asking the right questions: scoping studies in the commissioning of research on the organisation and delivery of health services. *Health Res Policy Syst*. 2008;6(7):6-7.

213.   Booth A. Australian supermodel?—A practical example of evidence-based library and information practice (EBLIP). *Health Info Libr J*. 2006;23(1):69-72.

214.   Cooke A, Smith D, Booth A. Beyond PICO: the SPIDER tool for qualitative evidence synthesis. *Qual Health Res*. 2012;22(10):1435-43.

215.   Methley AM, Campbell S, Chew-Graham C, McNally R, Cheraghi-Sohi S. PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC Health Serv Res*. 2014;14:579.

216.   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(36):1-154.

217.   Hausner E, Waffenschmidt S, Kaiser T, Simon M. Routine development of objectively derived search strategies. *Syst Rev*. 2012;1(1):19.

218.   Hausner E, Guddat C, Hermanns T, Lampert U, Waffenschmidt S. Prospective comparison of search strategies for systematic reviews: an objective approach yielded higher sensitivity than a conceptual one. *J Clin Epidemiol*. 2016;77:118-24.

219.   Brettle AJ, Long AF, Grant MJ, Greenhalgh J. Searching for information on outcomes: do you need to be comprehensive? *Qual Health Care*. 1998;7(3):163-7.

220.   Duffy S, de Kock S, Misso K, Noake C, Ross J, Stirk L. Supplementary searches of PubMed to improve currency of MEDLINE and MEDLINE In-Process searches via Ovid. *J Med Libr Assoc*. 2016;104(4):309-12.

221.   Katchamart W, Faulkner A, Feldman B, Tomlinson G, Bombardier C. PubMed had a higher sensitivity than Ovid-MEDLINE in the search for systematic reviews. *J Clin Epidemiol*. 2011;64(7):805-7.

222.   Lam MT, McDiarmid M. Increasing number of databases searched in systematic reviews and meta-analyses between 1994 and 2014. *J Med Libr Assoc*. 2016;104(4):284-9.

223.   Aagaard T, Lund H, Juhl C. Optimizing literature search in systematic reviews - are MEDLINE, EMBASE and CENTRAL enough for identifying effect studies within the area of musculoskeletal disorders? *BMC Med Res Methodol*. 2016;16(1):161.

224.   Kelly L, St Pierre-Hansen N. So many databases, such little clarity: Searching the literature for the topic aboriginal. *Can Fam Physician*. 2008;54(11):1572-3.

225.   Lawrence DW. What is lost when searching only one literature database for articles relevant to injury prevention and safety promotion? *Inj Prev*. 2008;14(6):401-4.

226.   Lemeshow AR, Blum RE, Berlin JA, Stoto MA, Colditz GA. Searching one or two databases was insufficient for meta-analysis of observational studies. *J Clin Epidemiol*. 2005;58(9):867-73.

227.   Sampson M, Barrowman NJ, Moher D, Klassen TP, Pham B, Platt R, et al. Should meta-analysts search Embase in addition to Medline? *J Clin Epidemiol*. 2003;56(10):943-55.

228.   Stevinson C, Lawlor DA. Searching multiple databases for systematic reviews: added value or diminishing returns? *Complement Ther Med*. 2004;12(4):228-32.

229.   Suarez-Almazor ME, Belseck E, Homik J, Dorgan M, Ramos-Remus C. Identifying clinical trials in the medical literature with electronic databases: MEDLINE alone is not enough. *Control Clin Trials*. 2000;21(5):476-87.

230.   Taylor B, Wylie E, Dempster M, Donnelly M. Systematically retrieving research: a case study evaluating seven databases. *Res Soc Work Pract*. 2007;17(6):697-706.

231.   Hopewell S, Clarke M, Lusher A, Lefebvre C, Westby M. A comparison of handsearching versus MEDLINE searching to identify reports of randomized controlled trials. *Stat Med*. 2002;21(11):1625-34.

232.   Marson AG, Chadwick DW. How easy are randomized controlled trials in epilepsy to find on Medline? The sensitivity and precision of two Medline searches. *Epilepsia*. 1996;37(4):377-80.

233.   Cooper C, Briscoe S, Crathorne L. Should RCT search filters account for the phases of clinical trials in addition to study design? Paper presented at: *Health Technology Assesment International (HTAi) Conference;* 2016 May 10-14; Tokyo, Japan.

234.   Dixon-Woods M, Fitzpatrick R, Roberts K. Including qualitative research in systematic reviews: opportunities and problems. *J Eval Clin Pract*. 2001;7(2):125-33.

235.   Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open*. 2016;6(11).

236.   Noel-Storr A, Thomas J, Mavergames C, Elliott J. Cochrane Crowd: Using citizen science to meet the challenge of information overload in evidence production. Paper presented at: *The European Association for Health Information and Libraries (EAHIL) Conference;* 2016 Jun 12-16; Dublin, Ireland.

237.   Marshall IJ, Noel-Storr A, Kuiper J, Thomas J, Wallace BC. Machine learning for identifying Randomized Controlled Trials: An evaluation and practitioner's guide. *Res Synth Methods*. 2018;9(4):602-14.

238.   Hinde S, Spackman E. Bidirectional citation searching to completion: an exploration of literature searching methods. *Pharmacoeconomics*. 2015;33(1):5-11.

239.   Matthews EJ, Edwards AG, Barker J, Bloor M, Covey J, Hood K, et al. Efficient literature searching in diffuse topics: lessons from a systematic review of research on communicating risk to patients in primary care. *Health Libr Rev*. 1999;16(2):112-20.

240.   McManus RJ, Wilson S, Delaney BC, Fitzmaurice DA, Hyde CJ, Tobias RS, et al. Review of the usefulness of contacting other experts when conducting a literature search for systematic reviews. *BMJ*. 1998;317(7172):1562-3.

241.   Westphal A, Kriston L, Holzel LP, Harter M, von Wolff A. Efficiency and contribution of strategies for finding randomized controlled trials: a case study from a systematic review on therapeutic interventions of chronic depression. *J Public Health Res*. 2014;3(2):177.

242.   Jadad AR, McQuay HJ. A high-yield strategy to identify randomized controlled trials for systematic reviews. *Online J Curr Clin Trials*. 1993;33.

243.   Bethel A. *Endnote Training* [YouTube Video]. Exeter: Alison Bethel; 2014 [Cited 2017 Dec 7]. Available from: http://medicine.exeter.ac.uk/esmi/workstreams/informationscience/is_resources,_guidance_&_advice/

244.   Bramer WM, Giustini D, de Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. *J Med Libr Assoc*. 2016;104(3):240-3.

245.   Bramer WM, Milic J, Mast F. Reviewing retrieved references for inclusion in systematic reviews using EndNote. *J Med Libr Assoc*. 2017;105(1):84-7.

246.   Gall C, Brahmi FA. Retrieval comparison of EndNote to search MEDLINE (Ovid and PubMed) versus searching them directly. *Med Ref Serv Q*. 2004;23(3):25-32.

247.   Ahmed KK, Al Dhubaib BE. Zotero: A bibliographic assistant to researcher. *J Pharmacol Pharmacother*. 2011;2(4):303-5.

248.   Coar JT, Sewell JP. Zotero: harnessing the power of a personal bibliographic manager. *Nurse Educ*. 2010;35(5):205-7.

249.   Lorenzetti DL, Ghali WA. Reference management software for systematic reviews and meta-analyses: an exploration of usage and usability. *BMC Med Res Methodol*. 2013;13(1):141.

250.   Nilashi M, Ibrahim O, Sohaei S, Ahmadi H, Almaee A. Features Influencing Researchers' Selection of Reference Management Software. *Journal of Information & Knowledge Management*. 2016;15(03):1650032.

251.   Kali A. Reference management: A critical element of scientific writing. *J Adv Pharm Technol Res*. 2016;7(1):27-9.

252.   Melles A, Unsworth K. Examining the Reference Management Practices of Humanities and Social Science Postgraduate Students and Academics. *Australian Academic & Research Libraries*. 2015;46(4):250-76.

253.   Salem J, Fehrmann P. Bibliographic Management Software: A Focus Group Study of the Preferences and Practices of Undergraduate Students. *Public Services Quarterly*. 2013;9(2):110-20.

254.   Mead TL, Berryman DR. Reference and PDF-Manager Software: Complexities, Support and Workflow. *Med Ref Serv Q*. 2010;29(4):388-93.

255.   Clarivate Analytics. EndNote X8 [Computer program]. Philadelphia: USA; 2018.

256.   Mourad Ouzzani, Hossam Hammady, Zbys Fedorowicz, Elmagarmid A. Rayyan — a web and mobile app for systematic reviews. *Syst Rev*. 2016;5(210).

257.   The Nordic Cochrane Centre. Review Manager (RevMan) 5.3 [Computer program]. Copenhagen: Denmark; 2014.

258.   Zhang Y. Comparison of Select Reference Management Tools. *Med Ref Serv Q*. 2012;31(1):45-60.

259.   Ovadia S. Managing Citations With Cost-Free Tools. *Behav Soc Sci Librar*. 2011;30(2):107-11.

260.   Emanuel J. Users and citation management tools: Use and support. *Ref Serv Rev*. 2013;41(4):639-59.

261.   Fitzgibbons M, Meert D. Are Bibliographic Management Software Search Interfaces Reliable?: A Comparison between Search Results Obtained Using Database Interfaces and the EndNote Online Search Function. *The Journal of Academic Librarianship*. 2010;36(2):144-50.

262.   Sampson M, McGowan J, Tetzlaff J, Cogo E, Moher D. No consensus exists on search reporting methods for systematic reviews. *J Clin Epidemiol*. 2008;61(8):748-54.

263.   Toews LC. Compliance of systematic reviews in veterinary journals with Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) literature search reporting guidelines. *J Med Libr Assoc*. 2017;105(3):233-9.

264.   Akhigbe T, Zolnourian A, Bulters D. Compliance of systematic reviews articles in brain arteriovenous malformation with PRISMA statement guidelines: Review of literature. *J Clin Neurosci*. 2017;39:45-8.

265.   Bigna JJ, Um LN, Nansseu JR. A comparison of quality of abstracts of systematic reviews including meta-analysis of randomized controlled trials in high-impact general medicine journals before and after the publication of PRISMA extension for abstracts: a systematic review and meta-analysis. *Syst Rev*. 2016;5(1):174.

266.   Faggion CM, Jr., Wu YC, Tu YK, Wasiak J. Quality of search strategies reported in systematic reviews published in stereotactic radiosurgery. *Br J Radiol*. 2016;89(1062):20150878.

267.   Mullins MM, DeLuca JB, Crepaz N, Lyles CM. Reporting quality of search methods in systematic reviews of HIV behavioral interventions (2000–2010): are the searches clearly explained, systematic and reproducible? *Res Synth Methods*. 2014;5(2):116-30.

268.   Tao KM, Li XQ, Zhou QH, Moher D, Ling CQ, Yu WF. From QUOROM to PRISMA: a survey of high-impact medical journals' instructions to authors and a review of systematic reviews in anesthesia literature. *PLoS One*. 2011;6(11):e27611.

269.   Wasiak J, Tyack Z, Ware R, Goodwin N, Faggion CM, Jr. Poor methodological quality and reporting standards of systematic reviews in burn care management. *Int Wound J*. 2016;14(5):754-63.

270.   Yoshii A, Plaut DA, McGraw KA, Anderson MJ, Wellik KE. Analysis of the reporting of search strategies in Cochrane systematic reviews. *J Med Libr Assoc*. 2009;97(1):21-9.

271.   The Cochrane Collobration *Cochrane peer review policy* [Internet]. Oxford: The Cochrane Collaboration; 2017 [Cited 2017 Dec 7]. Available from: https://community.cochrane.org/editorial-and-publishing-policy-resource/cochrane-review-management/cochrane-peer-review-policy

272.   Higgins JPT, Lasserson T, Chandler J, Tovey D, Churchill R. *Methodological Expectations of Cochrane Intervention Reviews* [Internet]. London: Cochrane Methods Group; 2017 [Cited 2017 Dec 7]. Available from: http://methods.cochrane.org/mecir

273.   Smith TO, Sexton D, Mann C, Donell S. Sutures versus staples for skin closure in orthopaedic surgery: meta-analysis. *BMJ*. 2010;340.

274.   Vochteloo A. Fundamental limitations of "Sutures versus staples for skin closure in orthopaedic surgery: meta-analysis". *BMJ*. 2010;340.

275.   Groves T, Loder E. Prepublication histories and open peer review at The BMJ. *BMJ*. 2014;349.

276.   Smith T. In response to Mr Ockendon. *BMJ*. 2010;340.

277.   Sampson M, Zhang L, Morrison A, Barrowman NJ, Clifford TJ, Platt RW, et al. An alternative to the hand searching gold standard: validating methodological search filters using relative recall. *BMC Med Res Methodol*. 2006;6(33).

278.   Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*. 2017;358.

279.   Whiting P, Savovic J, Higgins JP, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol*. 2016;69:225-34.

280.   Roundtree AK, Kallen MA, Lopez-Olivo MA, Kimmel B, Skidmore B, Ortiz Z, et al. Poor reporting of search strategy and conflict of interest in over 250 narrative and systematic reviews of two biologic agents in arthritis: a systematic review. *J Clin Epidemiol*. 2009;62(2):128-37.

281.   Relevo R, Balshem H. Finding evidence for comparing medical interventions: AHRQ and the Effective Health Care Program. *J Clin Epidemiol*. 2011;64(11):1168-77.

282.   The Institute of Medicine. *Finding What Works in Health Care: Standards for Systematic Reviews*. Washington DC: The National Academies; 2011.

283.   Canadian Agency for Drugs and Technologies in Health (CADTH). *Resources* [Internet]. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2017 [Cited 2017 Dec 7]. Available from: https://www.cadth.ca/resources

284.   Wilson TD. Information behaviour: An interdisciplinary perspective. *Inf Process Manag*. 1997;33(4):551-72.

285.   Paisley S. Identifying evidence to inform decision-analytic models of cost-effectiveness a qualitative study of information seeking processes and behaviour [dissertation]. Sheffield: University of Sheffield; 2012.

286.   Delaney A, Tamas PA. Searching for evidence or approval? A commentary on database search in systematic reviews and alternative information retrieval methodologies. *Res Synth Methods*. 2018;9(1):124-31.

287.   Armstrong R, Jackson N, Doyle J, Waters E, Howes F. It's in your hands: the value of handsearching in conducting systematic reviews of public health interventions. *J Public Health (Oxf)*. 2005;27(4):388-91.

288.   Bakkalbasi N, Bauer K, Glover J, Wang L. Three options for citation tracking: Google Scholar, Scopus and Web of Science. *Biomed Digit Libr*. 2006;3(7):3-7.

289.   Blumle A, Antes G. [Handsearching for randomized controlled clinical trials in German medical journals]. *Dtsch Med Wochenschr*. 2008;133(6):230-4.

290.   Bramer WM, Giustini D, Kramer BMR, Anderson PF. The comparative recall of Google Scholar versus PubMed in identical searches for biomedical systematic reviews: a review of searches used in systematic reviews. *Syst Rev*. 2013;2:115.

291.   Croft AM, Vassallo DJ, Rowe M. Handsearching the Journal of the Royal Army Medical Corps for trials. *J R Army Med Corps*. 1999;145(2):86-8.

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

293.   Gibson CA, Bailey BW, Carper MJ, LeCheminant JD, Kirk EP, Huang G, et al. Author contacts for retrieval of data for a meta-analysis on exercise and diet restriction. *Int J Technol Assess Health Care*. 2006;22(2):267-70.

294.   Glanville JM, Duffy S, McCool R, Varley D. Searching ClinicalTrials.gov and the International Clinical Trials Registry Platform to inform systematic reviews: what are the optimal search approaches? *J Med Libr Assoc*. 2014;102(3):177-83.

295.   Godin K, Stapleton J, Kirkpatrick SI, Hanning RM, Leatherdale ST. Applying systematic review search methods to the grey literature: a case study examining guidelines for school-based breakfast programs in Canada. *Syst Rev*. 2015;4(1):138.

296.   Hay PJ, Adams CE, Lefebvre C. The efficiency of searches for randomized controlled trials in the International Journal of Eating Disorders: a comparison of handsearching, EMBASE and PsycLIT. *Health Libr Rev*. 1996;13(2):91-6.

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

298.   Janssens AC, Gwinn M. Novel citation-based search method for scientific literature: application to meta-analyses. *BMC Med Res Methodol*. 2015;15:84.

299.   Jones CW, Keil LG, Weaver MA, Platts-Mills TF. Clinical trials registries are under-utilized in the conduct of systematic reviews: A cross-sectional analysis. *Syst Rev*. 2014;3 (126).

300.   Langham J, Thompson E, Rowan K. Identification of randomized controlled trials from the emergency medicine literature: comparison of hand searching versus MEDLINE searching. *Ann Emerg Med*. 1999;34(1):25-34.

301.   Mattioli S, Farioli A, Cooke RMT, Baldasseroni A, Ruotsalainen J, Placidi D, et al. Hidden effectiveness? Results of hand-searching Italian language journals for occupational health interventions. *Occup Environ Med*. 2012;69(7):522-4.

302.   Milne R, Thorogood M. Hand searching the Journal of Epidemiology and Community Health as part of th Cochrane Collaboration. *J Epidemiol Community Health*. 1996;50(2):178-81.

303.   Moher D, The OCG. How good are volunteers at searching for published randomized controlled trials? *Fundam Clin Pharmacol*. 1995;9(4):384-6.

304.   O'Leary F. Is email a reliable means of contacting authors of previously published papers? A study of the Emergency Medicine Journal for 2001. *Emerg Med J*. 2003;20(4):352-3.

305.   Reveiz L, Cardona AF, Ospina EG, de Agular S. An e-mail survey identified unpublished studies for systematic reviews. *J Clin Epidemiol*. 2006;59(7):755-8.

306.   Robinson KA, Dunn AG, Tsafnat G, Glasziou P. Citation networks of related trials are often disconnected: implications for bidirectional citation searches. *J Clin Epidemiol*. 2014;67(7):793-9.

307.   Selph SS, Ginsburg AD, Chou R. Impact of contacting study authors to obtain additional data for systematic reviews: diagnostic accuracy studies for hepatic fibrosis. *Syst Rev*. 2014;3(1):107.

308.   Stansfield C, Dickson K, Bangpan M. Exploring issues in the conduct of website searching and other online sources for systematic reviews: how can we be systematic? *Syst Rev*. 2016;5(1):191.

309.   van Enst WA, Scholten RJ, Hooft L. Identification of additional trials in prospective trial registers for Cochrane systematic reviews. *PLoS One*. 2012;7(8):e42812.

310.   Wright K, Golder S, Rodriguez-Lopez R. Citation searching: a systematic review case study of multiple risk behaviour interventions. *BMC Med Res Methodol*. 2014;14(1):73.

311.   Higgins JPT, Deeks JJ. Chapter 7: Selecting studies and collecting data. In: Higgins JPT, Green S, (eds.) *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0 (updated March 2011) [Internet]. The Cochrane Collaboration; 2011 [Cited 2017 Dec 7]. Available from: www.handbook.cochrane.org

312.   McGrath J, Davies G, Soares K. Writing to authors of systematic reviews elicited further data in 17% of cases. *BMJ*. 1998;316(7131):631.

313.   Mullan RJ, Flynn DN, Carlberg B, Tleyjeh IM, Kamath CC, LaBella ML, et al. Systematic reviewers commonly contact study authors but do so with limited rigor. *J Clin Epidemiol*. 2009;62(2):138-42.

314.   Ogilvie D, Egan M, Hamilton V, Petticrew M. Systematic reviews of health effects of social interventions: 2. Best available evidence: how low should you go? *J Epidemiol Community Health*. 2005;59(10):886-92.

315.   Sterne JAC, Egger M, Moher D. Chapter 10: Addressing reporting biases. In: Higgins JPT, Green S, (eds.) *Cochrane Handbook for Systematic Reviews of Interventions*. Version

5.1.0 (updated March 2011) [Internet]. The Cochrane Collaboration; 2011 [Cited 2017 Dec 7]. Available from: www.handbook.cochrane.org

316.   Kostoff RN. The use and misuse of citation analysis in research evaluation. *Scientometrics*. 1998;43(1):27-43.

317.   Seglen PO. [Use of citation analysis and other bibliometric methods in evaluation of the quality of research]. *Tidsskr Nor Laegeforen*. 1989;109(31):3229-4.

318.   Kuper H, Nicholson A, Hemingway H. Searching for observational studies: what does citation tracking add to PubMed? A case study in depression and coronary heart disease. *BMC Med Res Methodol*. 2006;6(4).

319.   de Winter JCF, Zadpoor AA, Dodou D. The expansion of Google Scholar versus Web of Science: a longitudinal study. *Scientometrics*. 2014;98(2):1547-65.

320.   Lowe J, Peters J, Shields B, Cooper C. Methods to update systematic literature searches: full update searching vs. forward citation chasing: A case study from a systematic review of diagnostic test accuracy. Paper presented at: *The InterTASC ISSG Conference;* 2014 Jul 9; Exeter, UK.

321.   Hausner E. *PubMed Commons (Elke Hausner)* [Internet]. Bethesda, Maryland: PubMed Commons; 2017 [Cited 2017 Dec 7]. Available from: https://www.ncbi.nlm.nih.gov/pubmed/22916163#comments

322.   De Angelis C, Drazen JM, Frizelle FAP, Haug C, Hoey J, Horton R, et al. Clinical Trial Registration: A Statement from the International Committee of Medical Journal Editors. *N Engl J Med*. 2004;351(12):1250-1.

323.   Hartung DM, Zarin DA, Guise JM, McDonagh M, Paynter R, Helfand M. Reporting discrepancies between the ClinicalTrials.gov results database and peer-reviewed publications. *Ann Intern Med* 2014;160(7):477-83.

324.   Ogino D, Takahashi K, Sato H. Characteristics of clinical trial websites: information distribution between ClinicalTrials.gov and 13 primary registries in the WHO registry network. *Trials*. 2014;15:428.

325.   Saleh AA, Ratajeski MA, Bertolet M. Grey Literature Searching for Health Sciences Systematic Reviews: A Prospective Study of Time Spent and Resources Utilized. *Evid Based Libr Inf Pract*. 2014;9(3):28-50.

326.   Mhaskar R, Emmanuel P, Mishra S, Patel S, Naik E, Kumar A. Critical appraisal skills are essential to informed decision-making. *Indian J Sex Transm Dis*. 2009;30(2):112-9.

327.   Kent A, Berry MM, Luehrs FU, Perry JW. Machine literature searching VIII. Operational criteria for designing information retrieval systems. *American Documentation*. 1955;6(2):93-101.

328.   Jenkins M. Evaluation of methodological search filters--a review. *Health Info Libr J*. 2004;21(3):148-63.

329.   Harbour J, Fraser C, Lefebvre C, Glanville J, Beale S, Boachie C, et al. Reporting methodological search filter performance comparisons: a literature review. *Health Info Libr J*. 2014;31(3):176-94.

330.   Astin MP, Brazzelli MG, Fraser CM, Counsell CE, Needham G, Grimshaw JM. Developing a sensitive search strategy in MEDLINE to retrieve studies on assessment of the diagnostic performance of imaging techniques. *Radiology*. 2008;247(2):365-73.

331.   Bachmann LM, Coray R, Estermann P, Ter Riet G. Identifying diagnostic studies in MEDLINE: reducing the number needed to read. *J Am Med Inform Assoc*. 2002;9(6):653-8.

332.   Blanc X, Collet TH, Auer R, Iriarte P, Krause J, Legare F, et al. Retrieval of publications addressing shared decision making: an evaluation of full-text searches on medical journal websites. *JMIR Res Protoc*. 2015;4(2):e38.

333.   Cathey J, Al Hajeri AA, Fedorowicz Z. A comparison of handsearching versus EMBASE searching of the Annals of Saudi Medicine to identify reports of randomized controlled trials. *Ann Saudi Med*. 2006;26(1):49-51.

334.   Dumbrigue HB, Esquivel JF, Jones JS. Assessment of MEDLINE search strategies for randomized controlled trials in prosthodontics. *Journal of Prosthodontics*. 2000;9(1):8-13.

335.   Gehanno JF, Rollin L, Le Jean T, Louvel A, Darmoni S, Shaw W. Precision and recall of search strategies for identifying studies on return-to-work in Medline. *J Occup Rehabil*. 2009;19(3):223-30.

336.   Haynes RB, Wilczynski NL. Optimal search strategies for retrieving scientifically strong studies of diagnosis from Medline: analytical survey. *BMJ*. 2004;328(7447):1040.

337.   Haynes RB, Kastner M, Wilczynski NL. Developing optimal search strategies for detecting clinically sound and relevant causation studies in EMBASE. *BMC Med Inform Decis Mak*. 2005;5 (8).

338.   Haynes RB, McKibbon KA, Wilczynski NL, Walter SD, Werre SR. Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. *BMJ*. 2005;330(7501):1179.

339.   Hildebrand AM, Iansavichus AV, Haynes RB, Wilczynski NL, Mehta RL, Parikh CR, et al. High-performance information search filters for acute kidney injury content in PubMed, Ovid Medline and Embase. *Nephrol Dial Transplant*. 2014;29(4):823-32.

340.   Holland JL, Wilczynski NL, Haynes RB. Optimal search strategies for identifying sound clinical prediction studies in EMBASE. *BMC Med Inform Decis Mak*. 2005;5(11).

341.   Ingui BJ, Rogers MA. Searching for clinical prediction rules in MEDLINE. *J Am Med Inform Assoc*. 2001;8(4):391-7.

342.   Jenuwine ES, Floyd JA. Comparison of Medical Subject Headings and text-word searches in MEDLINE to retrieve studies on sleep in healthy individuals. *J Med Libr Assoc*. 2004;92(3):349-53.

343.   Kassaï B, Sonié 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(7):710-4.

344.   Layton DM, Clarke M. Search Strategy to Identify Dental Survival Analysis Articles Indexed in MEDLINE. *Int J Prosthodont*. 2016;29(1):20-7.

345.   Linder SK, Kamath GR, Pratt GF, Saraykar SS, Volk RJ. Citation searches are more sensitive than keyword searches to identify studies using specific measurement instruments. *J Clin Epidemiol*. 2015;68(4):412-7.

346.   McKibbon KA, Wilczynski NL, Haynes RB. Developing optimal search strategies for retrieving qualitative studies in PsycINFO. *Eval Health Prof*. 2006;29(4):440-54.

347.   McKibbon KA, Wilczynski NL, Haynes RB. Retrieving randomized controlled trials from medline: A comparison of 38 published search filters. *Health Info Libr J*. 2009;26(3):187-202.

348.   McKinlay RJ, Wilczynski NL, Haynes RB. Optimal search strategies for detecting cost and economic studies in EMBASE. *BMC Health Serv Res*. 2006;6 (67).

349.   Montori VM, Wilczynski NL, Morgan D, Haynes RB. Optimal search strategies for retrieving systematic reviews from Medline: Analytical survey. *BMJ*. 2005;330(7482):68-71.

350.   Nasser M, Al Hajeri A. A comparison of handsearching versus embase searching of the archives of Iranian medicine to identify reports of randomized controlled trials. *Arch Iran Med*. 2006;9(3):192-5.

351.   Rogerson TE, Ladhani M, Mitchell R, Craig JC, Webster AC. Efficient strategies to find diagnostic test accuracy studies in kidney journals. *Nephrology (Carlton)*. 2015;20(8):513-8.

352.   Spoor P, Airey M, Bennett C, Greensill J, Williams R. Use of the capture-recapture technique to evaluate the completeness of systematic literature searches. *BMJ*. 1996;313(7053):342-3.

353.   Taljaard M, McGowan J, Grimshaw JM, Brehaut JC, McRae A, Eccles MP, et al. Electronic search strategies to identify reports of cluster randomized trials in MEDLINE: low precision will improve with adherence to reporting standards. *BMC Med Res Methodol*. 2010;10(15).

354.   Ugolini D, Neri M, Casilli C, Bonassi S. Development of search filters for retrieval of literature on the molecular epidemiology of cancer. *Mutat Res Genet Toxicol Environ Mutagen*. 2010;701(2):107-10.

355.   van de Glind EM, van Munster BC, Spijker R, Scholten RJ, Hooft L. Search filters to identify geriatric medicine in Medline. *J Am Med Inform Assoc*. 2012;19(3):468-72.

356.   Walters LA, Wilczynski NL, Haynes RB. Developing optimal search strategies for retrieving clinically relevant qualitative studies in EMBASE. *Qual Health Res*. 2006;16(1):162-8.

357.   Wilczynski NL, Walker CJ, McKibbon KA, Haynes RB. Assessment of methodologic search filters in MEDLINE. Paper presented at: *The Annual Symposium on Computer Applications in Medical Care;* 1993 Oct 30; Washington DC, USA.

358.   Wilczynski NL, Walker CJ, McKibbon KA, Haynes RB. Quantitative comparison of pre-explosions and subheadings with methodologic search terms in MEDLINE. Paper presented at: *The Annual Symposium on Computer Applications in Medical Care;* 1994 Nov 5-9; Washington DC, USA.

359.   Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound causation studies in MEDLINE. Paper presented at: *The AMIA Symposium;* 2003 Nov 8-12; Washington DC, USA.

360.   Wilczynski NL, Haynes RB, Eady A, Haynes B, Marks S, McKibbon A, et al. Developing optimal search strategies for detecting clinically sound prognostic studies in MEDLINE: An analytic survey. *BMC Med*. 2004;2 (23).

361.   Wilczynski NL, Haynes RB, Lavis JN, Ramkissoonsingh R, Arnold-Oatley AE. Optimal search strategies for detecting health services research studies in MEDLINE. *CMAJ*. 2004;171(10):1179-85.

362.   Wilczynski NL, Haynes RB. Optimal search strategies for detecting clinically sound prognostic studies in EMBASE: an analytic survey. *J Am Med Inform Assoc*. 2005;12(4):481-5.

363.   Wilczynski NL, Haynes RB, Hedges T. Optimal search strategies for identifying mental health content in MEDLINE: An analytic survey. *Ann Gen Psychiatry*. 2006;5 (4).

364.   Wilczynski NL, Haynes RB. EMBASE search strategies achieved high sensitivity and specificity for retrieving methodologically sound systematic reviews. *J Clin Epidemiol*. 2007;60(1):29-33.

365.   Wilczynski NL, Haynes RB, Team QIH. Optimal search filters for detecting quality improvement studies in Medline. *Qual Saf Health Care*. 2010;19(6):e31.

366.   Wilczynski NL, McKibbon KA, Haynes RB. Search filter precision can be improved by NOTing out irrelevant content. Paper presented at: *The AMIA Symposium;* 2011 Oct 22-26; Washington DC, USA.

367.   Wong SS, Wilczynski NL, Haynes RB, Ramkissoonsingh R. Developing optimal search strategies for detecting sound clinical prediction studies in MEDLINE. Paper presented at: *The AMIA Symposium;* 2003 Nov 8-12; Washington DC, USA.

368.   Wong SS, Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically relevant qualitative studies in MEDLINE. *Medinfo*. 2004;11(1):311-6.

369.   Wong SS, Wilczynski NL, Haynes RB. Optimal CINAHL search strategies for identifying therapy studies and review articles. *J Nurs Scholarsh*. 2006;38(2):194-9.

370.   Wong SS, Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound treatment studies in EMBASE. *J Med Libr Assoc*. 2006;94(1):41-7.

371.   Kok R, Verbeek JAHM, Faber B, van Dijk FJH, Hoving JL. A search strategy to identify studies on the prognosis of work disability: a diagnostic test framework. *BMJ Open*. 2015;5(5).

372.   Lee E, Dobbins M, Decorby K, McRae L, Tirilis D, Husson H. An optimal search filter for retrieving systematic reviews and meta-analyses. *BMC Med Res Methodol*. 2012;12:51.

373.   Waffenschmidt S, Guddat C. Searches for randomized controlled trials of drugs in MEDLINE and EMBASE using only generic drug names compared with searches applied in current practice in systematic reviews. *Res Synth Methods*. 2015;6(2):188-94.

374.   Doust JA, Pietrzak E, Sanders S, Glasziou PP. Identifying studies for systematic reviews of diagnostic tests was difficult due to the poor sensitivity and precision of methodologic filters and the lack of information in the abstract. *J Clin Epidemiol*. 2005;58(5):444-9.

375.   Vlachos A, Craven M. Biomedical event extraction from abstracts and full papers using search-based structured prediction. *BMC Bioinformatics*. 2012;13(Suppl 11):S5.

376.   Guns R, Lioma C, Larsen B. The tipping point: F-score as a function of the number of retrieved items. *Inf Process Manag*. 2012;48(6):1171-80.

377.   Hripcsak G, Rothschild AS. Agreement, the F-Measure, and Reliability in Information Retrieval. *J Am Med Inform Assoc*. 2005;12(3):296-8.

378.   Liu Z, Tan M, Jiang F. Regularized F-measure maximization for feature selection and classification. *J Biomed Biotechnol*. 2009;2009:617946.

379.	Patrick TB, Demiris G, Folk LC, Moxley DE, Mitchell JA, Tao D. Evidence-based retrieval in evidence-based medicine. *J Med Libr Assoc*. 2004;92(2):196-9.

380.	Hausner E, Guddat C, Hermanns T, Lampert U, Waffenschmidt S. Development of search strategies for systematic reviews: validation showed the noninferiority of the objective approach. *J Clin Epidemiol*. 2015;68(2):191-9.

381.	Glanville J, Kaunelis D, Mensinkai S. How well do search filters perform in identifying economic evaluations in MEDLINE and EMBASE. *Int J Technol Assess Health Care*. 2009;25(4):522-9.

382.	Kastner M, Straus SE, McKibbon KA, Goldsmith CH. The capture-mark-recapture technique can be used as a stopping rule when searching in systematic reviews. *J Clin Epidemiol*. 2009;62(2):149-57.

383.	Lane D, Dykeman J, Ferri M, Goldsmith CH, Stelfox HT. Capture-mark-recapture as a tool for estimating the number of articles available for systematic reviews in critical care medicine. *J Crit Care*. 2013;28(4):469-75.

384.	Siva S. Optimal strategies for literature search. *Indian J Urol*. 2009;25(2):246-50.

385.	Hildebrand AM, Iansavichus AV, Lee CW, Haynes RB, Wilczynski NL, McKibbon KA, et al. Glomerular disease search filters for Pubmed, Ovid Medline, and Embase: a development and validation study. *BMC Med Inform Decis Mak*. 2012;12(49).

386.	Webster AJ, Kemp R. Estimating Omissions From Searches. *Am Stat*. 2013;67(2):82-9.

387.	Ross-White A, Godfrey C. Is there an optimum number needed to retrieve to justify inclusion of a database in a systematic review search? *Health Info Libr J*. 2017;34(3):217-24.

388.	Egan M, MacLean A, Sweeting H, Hunt K. Comparing the effectiveness of using generic and specific search terms in electronic databases to identify health outcomes for a systematic review: a prospective comparative study of literature search methods. *BMJ Open*. 2012;2(3).

389.	Booth A. The number needed to retrieve: a practically useful measure of information retrieval? *Health Info Libr J*. 2006;23(3):229-32.

390.	Shemilt I, Khan N, Park S, Thomas J. Use of cost-effectiveness analysis to compare the efficiency of study identification methods in systematic reviews. *Syst Rev*. 2016;5(1):140.

391.	Bambra C. Real world reviews: a beginner's guide to undertaking systematic reviews of public health policy interventions. *J Epidemiol Community Health*. 2011;65(1):14-9.

392.	Briscoe S, Cooper C. The British Nursing Index and CINAHL: a comparison of journal title coverage and the implications for information professionals. *Health Info Libr J*. 2014;31(3):195-203.

393.	Flemming K, Briggs M. Electronic searching to locate qualitative research: evaluation of three strategies. *J Adv Nurs*. 2007;57(1):95-100.

394.	National Institute for Health and Care Excellence (NICE). *Guide to the methods of technology appraisal* [Internet]. London: National Institute for Health and Care Excellence (NICE); 2013 [Cited 2017 Dec 7]. Available from: https://www.nice.org.uk/process/pmg9

395.   National Institute for Health and Care Excellence (NICE)*Tuberculosis evidence review 1 [PH37]: review of barriers and facilitators.* London: National Institute for Health and Care Excellence (NICE); 2010.

396.   National Institute for Health and Care Excellence (NICE)*Identifying and managing tuberculosis among hard-to-reach groups [PH 37].* London: National Institute for Health and Care Excellence (NICE); 2012.

397.   National Institute for Health and Care Excellence (NICE)*Tuberculosis in hard to reach groups [PH37]: final scope.* London: National Institute for Health and Care Excellence (NICE); 2010.

398.   Royle PL, Waugh NR. Making literature searches easier: a rapid and sensitive search filter for retrieving randomized controlled trials from PubMed. *Diabet Med*. 2007;24(3):308-11.

399.   Kelly M, Morgan A, Ellis S, Younger T, Huntley J, Swann C. Evidence based public health: A review of the experience of the National Institute of Health and Clinical Excellence (NICE) of developing public health guidance in England. *Soc Sci Med*. 2010;71(6):1056-62.

400.   Alpi KM. Expert searching in public health. *J Med Libr Assoc*. 2005;93(1):97-103.

401.   Grant MJ. How does your searching grow? A survey of search preferences and the use of optimal search strategies in the identification of qualitative research. *Health Info Libr J*. 2004;21(1):21-32.

402.   Shaw RL, Booth A, Sutton AJ, Miller T, Smith JA, Young B, et al. Finding qualitative research: an evaluation of search strategies. *BMC Med Res Methodol*. 2004;4(1):5.

403.   Flanagan SM, Hancock B. 'Reaching the hard to reach' - lessons learned from the VCS (voluntary and community Sector). A qualitative study. *BMC Health Serv Res*. 2010;10(92).

404.   Dixon-Woods M, Cavers D, Agarwal S, Annandale E, Arthur A, Harvey J, et al. Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. *BMC Med Res Methodol*. 2006;6(1):35.

405.   Health Protection Agency. *Tuberculosis in the UK: annual report on tuberculosis surveillance in the UK*. London: Health Protection Agency; 2009.

406.   Lonnroth K, Jaramillo E, Williams BG, Dye C, Raviglione M. Drivers of tuberculosis epidemics: the role of risk factors and social determinants. *Soc Sci Med*. 2009;68(12):2240-6.

407.   Story A, Murad S, Roberts W, Verheyen M, Hayward AC. Tuberculosis in London: the importance of homelessness, problem drug use and prison. *Thorax*. 2007;62(8):667-71.

408.   French CE, Antoine D, Gelb D, Jones JA, Gilbert RL, Watson JM. Tuberculosis in non-UK-born persons, England and Wales, 2001-2003. *Int J Tuberc Lung Dis*. 2007;11(5):577-84.

409.   Gershon RR, Vlahov D, Escamilla-Cejudo JA, Badawi M, McDiarmid M, Karkashian C, et al. Tuberculosis risk in funeral home employees. *J Occup Environ Med*. 1998;40(5):497-503.

410.   Nist A. Tuberculosis in farmworkers. *J Int Assoc Physicians AIDS Care (Chic)*. 1996;2(8):15.

411.   Zakopaĭlo GG, Kaliberskaia SG, Makarchuk LF. Status of knowledge regarding various aspects of tuberculosis among industrial workers. *Probl Tuberk Bolezn Legk*. 1990(10):22-3.

412.   Mikitka D, Mills SE, Dazey SE, Gabriel ME. Tuberculosis infection in U.S. Air Force dentists. *Am J Dent* 1995;8(1):33-6.

413.   Mangura BT, Napolitano EC, Passannate MR, McDonald RJ, Reichman LB. Mycobacterium tuberculosis miniepidemic in a church gospel choir. *Chest*. 1998;113(1):234-7.

414.   MacDonald ME, Rigillo N, Brassard P. Urban Aboriginal understandings and experiences of tuberculosis in Montreal, Quebec, Canada. *Qual Health Res*. 2010;20(4):506-23.

415.   Rodríguez-Reimann DI, Nicassio P, Reimann JO, Gallegos PI, Olmedo EL. Acculturation and health beliefs of Mexican Americans regarding tuberculosis prevention. *J Immigr Health*. 2004;6(2):51-62.

416.   Husk K, Lovell R, Cooper C, Stahl-Timmins W, Garside R. Participation in environmental enhancement and conservation activities for health and well-being in adults: a review of quantitative and qualitative evidence. *Cochrane Database Syst Rev*. 2016(5).

417.   Dicks LV, Hodge I, Randall NP, Scharlemann JPW, Siriwardena GM, Smith HG, et al. A Transparent Process for "Evidence-Informed" Policy Making. *Conserv Lett*. 2014;7(2):119-25.

418.   Bilotta GS, Milner AM, Boyd I. On the use of systematic reviews to inform environmental policies. *Environ Sci Policy*. 2014;42:67-77.

419.   Bilotta GS, Milner AM, Boyd IL. Quality assessment tools for evidence from environmental science. *Environmental Evidence*. 2014;3(1):14.

420.   Fazey I, Salisbury JG, Lindenmayer DB, Maindonald J, Douglas R. Can methods applied in medicine be used to summarize and disseminate conservation research? *Environ Conserv*. 2004;31(3):190-8.

421.   Haddaway NR, Bayliss HR. Shades of grey: Two forms of grey literature important for reviews in conservation. *Biol Conserv*. 2015;191(Supplement C):827-9.

422.   Pullin AS, Knight TM. Effectiveness in Conservation Practice: Pointers from Medicine and Public Health [Efectividad de la Conservación Práctica: Indicadores de Medicina y Salud Pública]. *Conserv Biol*. 2001;15(1):50-4.

423.   Stewart GB, Coles CF, Pullin AS. Applying evidence-based practice in conservation management: Lessons from the first systematic review and dissemination projects. *Biol Conserv*. 2005;126(2):270-8.

424.   Kareiva P, Marvier M, West S, Hornisher J. Slow-moving journals hinder conservation efforts. *Nature*. 2002;420(6911):15-.

425.   Husk K, Lovell R, Cooper C, Garside R. Participation in environmental enhancement and conservation activities for health and well-being in adults. *Cochrane Database Syst Rev*. 2013(2).

426.   Armstrong R, Waters E, Doyle J. Chapter 21: Reviews in health promotion and public health. In: Higgins JPT, Green S, (eds.). *Cochrane Handbook for Systematic Reviews of*

Interventions. Version 5.1.0 (updated March 2011) [Internet]. The Cochrane Collaboration; 2011 [Cited 2017 Dec 7]. Available from: www.handbook.cochrane.org

427.   Grayson L, Gomersall A. *A difficult business: finding the evidence for social science reviews* [Internet]. London: ESRC UK Centre for Evidence Based Policy and Practice; 2003 [Cited 2017 Dec 7]. Available from: https://www.kcl.ac.uk/sspp/departments/politicaleconomy/research/cep/pubs/papers/assets/wp19.pdf

428.   Garside R. Should we appraise the quality of qualitative research reports for systematic reviews, and if so, how? *Innovation (Abingdon)*. 2014;27(1):67-79.

429.   Armijo-Olivo S, Stiles CR, Hagen NA, Biondo PD, Cummings GG. Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. *J Eval Clin Pract*. 2012;18(1):12-8.

430.   Effective Public Health Practice Project. *Effective Public Health Practice Project: Quality Assessment Tool for Quantitative Studies* [Internet]. Hamilton: Effective Public Health Practice Project; 2009 [Cited 2017 Dec 7]. Available from: https://merst.ca/ephpp/

431.   Wallace A, Croucher K, Quilgars D, Baldwin S. Meeting the challenge: developing systematic reviewing in social policy. *Policy Polit*. 2004;32(4):455-70.

432.   Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med Res Methodol*. 2008;8(1):45.

433.   Lewin S, Glenton C, Munthe-Kaas H, Carlsen B, Colvin CJ, Gulmezoglu M, 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 Med*. 2015;12(10):e1001895.

434.   Burls A. People and green spaces: promoting public health and mental health well-being though ecotherapy. *J Public Ment Health*. 2007;6(3):24-39.

435.   Gooch M. Voices of the volunteers: an exploration of the experiences of catchment volunteers in coastal Queensland, Australia. *Local Environment*. 2005;10(1):1-10.

436.   Brooker J, Brooker M. *Comparative exercise values of green gym and conventional gym: a personal evaluation*. Oxford: Wallingford green gym; 2008.

437.   Brooker J, Brooker M. *Comparative heart rates following green gym, other outdoor exercise and conventional gym: a personal evaluation* [Internet]. Oxford: Wallingford green gym; 2008 [Cited 2017 Dec 7]. Available from: https://www.wallingfordgreengym.org.uk/files/2011/02/post_exercise_evaluation.doc

438.   British Trust for Conservation Volunteers (BTCV). *Wellbeing comes naturally: year one report*. Doncaster: British Trust for Conservation Volunteers (BTCV); 2010.

439.   Christie J. Volunteer attitudes and motivations: research findings and their application for sustainable community involvement programs in natural resource management. Paper presented at: *Effective Sustainability Education: What Works? Why? Where Next? Linking Research and Practice;* 2004 Feb 18-20; Sydney, Australia.

440.   Eastaugh K, Tudge K, Lawes K. *Wye wood evaluation: 2006-2009*. Telford: Small Woods Association; 2010.

441.   Halpenny EA, Caissie LT. Volunteering on nature conservation projects: volunteer experience, attitudes and values. *Tourism Recreation Research*. 2003;28(3):25-33.

442.   Small Woods Association. *Amazon Woman Hereford SF36 Analysis*. Telford: Small Woods Association; 2011.

443.   Wilson N. *Branching Out. Greenspace and conservation on referral* [Internet]. Glasgow: Forestry Commission Scotland, NHSGGC, Glsgow Centre for Population Health, Glasgow Clyde Valley Green Newtork Partnership; 2009 [Cited 2017 Dec 7]. Available from: http://www.forestry.gov.uk/pdf/Branching_Out_Report.pdf/$FILE/Branching_Out_Report.pdf

444.   Yerrell P. *National Evaluation of BTCV's Green Gym* [Internet]. Oxford: School of Health and Social Care, Oxford Brookes University; 2008 [Cited 2017 Dec 7]. Available from: https://www.tcv.org.uk/sites/default/files/green-gym-evaluation-full.pdf

445.   Townsend M, Marsh R. *Exploration of the health and well-being benefits of membership of Truganina explosives reserve preservation society*. Burwood, Australia: School of Health and Social Development, Deakin University; 2004.

446.   Moore M, Townsend M, Oldroyd J. Linking Human and Ecosystem Health: The Benefits of Community Involvement in Conservation Groups. *EcoHealth*. 2006;3(4):255-61.

447.   Barton J. The effects of green exercise on psychological health and well-being [dissertation]. Colchester: University of Essex; 2009.

448.   Birch M. Cultivating wildness : three conservation volunteers' experiences of participation in the Green Gym scheme. *Br J Occup Ther*. 2005;68(6):244-52.

449.   Carter C. *Offenders and nature schemes: using conservation and forest management in rehabilitation* [Internet]. Farnham: Forest Research; 2008 [Cited 2017 Dec 7]. Available from: https://www.forestresearch.gov.uk/documents/1428/SERG_Offenders_and_nature_research_summary.pdf

450.   O'Brien L, Townsend M, Ebden M. *'I like to think when I'm gone I will have left this a better place' Environmental volunteering: motivations, barriers and benefits* [Internet]. Glasgow: Scottish Forestry Trust and Forestry Commission; 2008 [Cited 2017 Dec 7]. Available from: https://www.forestresearch.gov.uk/documents/4819/Env_Volunteering_Summary_Report.pdf

451.   Pillemer K. Environmental volunteering and health outcomes over a 20-year period. *Gerontologist*. 2010;50(5):594-602. 2010.

452.   Reynolds V. *The Green Gym: Evaluation of a pilot project in Sonning Common, Oxfordshire* [Internet]. Oxford: Oxford Centre for Health Care Research and Development (OCHRAD); Oxford Brookes; 1999 [Cited 2017 Dec 7]. Available from: http://hdl.handle.net/10068/398307

453.   Townsend M. Feel blue? Touch green! Participation in forest/woodland management as a treatment for depression. *Urban For Urban Green*. 2006;5(3):111-20.

454.   Townsend M, Moore M. *Research into the health, wellbeing & social capital benefits of community involvement in the management of land for conservation: final report* [Internet]. Geelong, Victoria: Deakin University, Trust for Nature; 2005 [Cited 2017 Dec 7]. Available from: http://dro.deakin.edu.au/view/DU:30010377

455. Lovell R, Husk K, Cooper C, Stahl-Timmins W, Garside R. Understanding how environmental enhancement and conservation activities may benefit health and wellbeing: a systematic review. *BMC Public Health*. 2015;15(1):864.

456. O'Brien L, Burls A, Townsend M, Ebden M. Volunteering in nature as a way of enabling people to re-integrate into society. *Perspect Public Health*. 2010;131(2):71-81.

457. Golder S, Loke YK. The contribution of different information sources for adverse effects data. *Int J Technol Assess Health Care*. 2012;28(2):133-7.

458. Helmer D, Savoie I, Green C, Kazanjian A. Evidence-based practice: extending the search to find material for the systematic review. *Bull Med Libr Assoc*. 2001;89(4):346-52.

459. Savoie I, Helmer D, Green CJ, Kazanjian A. Beyond Medline: reducing bias through extended systematic review search. *Int J Technol Assess Health Care*. 2003;19(1):168-78.

460. Lasserre K. Expert searching in health librarianship: a literature review to identify international issues and Australian concerns. *Health Info Libr J*. 2012;29(1):3-15.

461. Gore GC, Jones J. Systematic Reviews and Librarians: A Primer for Managers. *Partnership*. 2015;10(1).

462. Janke R, Rush KL. The academic librarian as co-investigator on an interprofessional primary research team: a case study. *Health Info Libr J*. 2014;31(2):116-22.

463. Lessick S. The Medical Library Association Essential Guide to Becoming an Expert Searcher: Proven Techniques, Strategies, and Tips for Finding Health Information. *J Med Libr Assoc*. 2009;97(2):153-5.

464. Tan MC, Maggio LA. Expert searcher, teacher, content manager, and patient advocate: an exploratory study of clinical librarian roles. *J Med Libr Assoc*. 2013;101(1):63-72.

465. Whipple EC, McGowan JJ, Dixon BE, Zafar A. The selection of high-impact health informatics literature: a comparison of results between the content expert and the expert searcher. *J Med Libr Assoc*. 2009;97(3):212-8.

466. Wilkinson A, Papaioannou D, Keen C, Booth A. The role of the information specialist in supporting knowledge transfer: a public health information case study. *Health Info Libr J*. 2009;26(2):118-25.

467. Hsieh-Yee I. Effects of search experience and subject knowledge on the search tactics of novice and experienced searchers. *J Assoc Inf Sci Technol*. 1993;44(3):161-74.

468. Shelda D. Wrong way: go back! An exploration of novice search behaviours while conducting an information search. *The Electronic Library*. 2001;19(6):371-82.

469. Pham MT, Waddell L, Rajić A, Sargeant JM, Papadopoulos A, McEwen SA. Implications of applying methodological shortcuts to expedite systematic reviews: three case studies using systematic reviews from agri-food public health. *Res Synth Methods*. 2016;7(4):433-46.

470. Uttley L, Montgomery P. The influence of the team in conducting a systematic review. *Syst Rev*. 2017;6(1):149.

471. Frost J, Garside R, Cooper C, Britten N. A qualitative synthesis of diabetes self-management strategies for long term medical outcomes and quality of life in the UK. *BMC Health Serv Res*. 2014;14:348.

472.   Thorne S. Metasynthetic Madness:What Kind of Monster Have We Created? *Qual Health Res*. 2017;27(1):3-12.

473.   Borah R, Brown AW, Capers PL, Kaiser KA. Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. *BMJ Open*. 2017;7(2).

474.   Livas C, Pandis N, Ren Y. Time relevance, citation of reporting guidelines, and breadth of literature search in systematic reviews in orthodontics. *Eur J Orthod*. 2015;37(2):183-7.

475.   Sampson M, McGowan J, Cogo E, Grimshaw J, Moher D, Lefebvre C. An evidence-based practice guideline for the peer review of electronic search strategies. *J Clin Epidemiol*. 2009;62(9):944-52.

476.   Miake-Lye IM, Hempel S, Shanman R, Shekelle PG. What is an evidence map? A systematic review of published evidence maps and their definitions, methods, and products. *Syst Rev*. 2016;5(1):28.

Improving literature searching in systematic reviews:

the application of tailored literature searching compared to 'the conventional approach'

Submitted by Chris Cooper to the University of Exeter

as a thesis for the degree of

Doctor of Philosophy

In December 2017

Volume 2 of 2

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

In this volume (2 of 2) of my PhD thesis, I set out the appendix to the main thesis and I report the supplementary materials to each of the published studies reported as chapters in volume one.

Table of Contents

TABLES:

FIGURES:

Annex 1. Models on information retrieval and information behaviour

In Table Four (volume one) of this thesis, I report nine models of information retrieval or behaviour which I identified and that I used to explore the development of the tailored approach.

For my own development, I illustrated these models as reported in the underlying papers identified. I also situated them alongside the conventional approach to illustrate any commonalities between models. I set this work out below.

Figure 1 Taylor 1962

The process of asking question & question negotiation (Taylor, 1962)

'one person tries to describe for another person not something he knows but rather something he doe not know' (Taylor, 1968)

Visceral	Conscious	Formalized	Compromised
the actual but unexpressed need for information	the conscious within brain description of the need	the formal statement of the question	the question as presented to the information system

'information needs may be unconscious and inexpressible' (Chang, 2013)

question negotiation

'An inquiry…[is an] area of doubt in which the question is open-ended, negotiable and dynamic' (Tay

Key stage one: who should literature search?

Key stage two: aims and purpose of literature searching

Key stage three: preparation

Key stage four: the search strategy

Key stage five: searching bibliographic databases

Key stage six: supple-mentary searching

Key stage seven: managing references

Key stage eight report the sea proce

Figure 2 Belkin 1980 and 1982

Anomalous state of knowledge: ASK (Belkin, 1980) + (Belkin et al. 1982)

Belkin's idea adapted by me

A

B

informs the information need

A = what the user knows

B = what exists but is unknown to the user

GENERATOR'S IMAGE OF THE WORLD

TEXT

REQUEST

USER['S] IMAG[E] THE [WORLD]

linguistic, pragmatic transformations

CONCEPTUAL STATE OF KNOWLEDGE

INFORMATION

ANOMALOUS STATE OF KNOWLEDGE

CONCEPT[UAL] OF KNOW[LEDGE]

belief, intent, knowledge of user transformations

realization of need

| Key stage one: who should literature search? | Key stage two: aims and purpose of literature searching | Key stage three: preparation | Key stage four: the search strategy | Key stage five: searching bibliographic databases | Key stage six: supple-mentary searching | Key stage seven: managing references | Key stage eight: reporting the search process |

'The driving force of the IR situation is the user's problem' (Belkin et. al 1982)

Figure 3 Wilson 1981

Model of information behaviour (Wilson, 1981)

Information User

"Need"

Satisfaction or Non-satisfaction

Information-seeking Behaviour

Information Use

Information Exchange

Demands on Information Systems

Demands on other Information Sources

Other People

Success

Failure

Information Transfer

information-seeking behavior results from the recognition [of] need, perceived by the user

individual looking for facts, advice or opinions, and may receive any of these either in writing or orally

the channel may be of overriding significance… orally given advice may be *preferred*

information-seeking behaviour may make demands upon formal [systems] such as libraries, on-line services, or information centres
OR
upon systems which may perform information functions in additi[on to] primary, non-information function

information transfer may be a fact, opinion or advice transferred orally or a physical document 'containing' fact, opinion or advice

1 2 3

| Key stage one: who should literature search? | Key stage two: aims and purpose of literature searching | Key stage three: preparation | Key stage four: the search strategy | Key stage five: searching bibliographic databases | Key stage six: supple-mentary searching | Key stage seven: managing references | Key stage eight: reporting the search process |

1. The result of a more truly user-oriented, innovative, experimental information profession should be a reduction in the *marginality* of information service
2. Wilson explores the idea of the 'information need,' although he argues that 'satisfaction of needs' might be advisable
3. 'attention has been given in the past chiefly to the demands people make upon formal information systems. Curiously, information use (which ought to point most directly to the needs experienced by people) is one of the most neglected areas' (Wilson, 1981)

4

Figure 4 Dervin 1983

Sense-Making triangle (Dervin, 1983)

Situation

Gap   Outcome

Dervin's original 'sense-making' triangle. Dervin now considers this model to be out of date but, for the purpose of testing the the conventional approach, it helps conceptually define the context of process in information behaviour

Dervin's 1983 model re-drawn by Wilson

Bridge

_____time/space_____

Situation   Gap   Outcome

where the situation or need for 'sense' arises

the output or 'outcome': the consequ 'sense-making'

Wilson argues that 'it may be preferable to use the bridge metaphor more directly,' hence his re-drawing of Dervin's original model. Wilson argues that the 'bridge' as a metaphor presents a way to question the 'nature of the problem,' raising the question how can information help 'bridge the uncertainty.' This can 'influence' information service delivery (Dervin, 1983, Wilson 1999).

| Key stage one: who should literature search? | Key stage two: aims and purpose of literature searching | Key stage three: preparation | Key stage four: the search strategy | Key stage five: searching bibliographic databases | Key stage six: supple-mentary searching | Key stage seven: managing references | Key stage eight: reporting the search process |

I think: The conventional approach represents the bridge: or how the information required to address to the particular problem is identified

Figure 5 Ellis 1989

Ellis (1989) and Ellis, Cox & Hall (1993)

- Starting: the means employed by the user to begin seeking information, for example, asking some knowledgeable colleague;
- Chaining: following footnotes and citations in known material or 'forward' chaining from known items through citation indexes;
- Browsing: 'semi-directed or semi-structured searching' (Ellis, 1989: 187)
- Differentiating: using known differences in information sources as a way of filtering the amount of information obtained;
- Monitoring: keeping up-to-date or current awareness searching;
- Extracting: selectively identifying relevant material in an information source;
- Verifying: checking the accuracy of information;
- Ending: which may be defined as 'tying up loose ends' through a final search.

The way Ellis describes 'starting', the process is broadly analogous systematic literature searchers would recognise as scoping prior to systematic review; Namely 'Starting refers to characteristics of the information seeking patterns of researchers who are commencing w on a new topic or in a new area (Ellis, 1989)'

Chaining is citation searching and the rationale, as a pra information retrieval technique in 1989, remains valid in 2

'browsing is a recognised, if sometimes disparaged form of inf seeking activity' (Ellis, 1989)

The key word here is 'filtering.' This does not happen in the conventional

This stage exists in order to ensure the currency of information dur the project

| Key stage one: who should literature search? | Key stage two: aims and purpose of literature searching | Key stage three: preparation | Key stage four: the search strategy | Key stage five: searching bibliographic databases | Key stage six: supple-mentary searching | Key stage seven: managing references | Key stage eight: reporting the search process |

Figure 6 Bates 1989

Bates: Berry Picking (1989)

Bates rejects 'the classic model' in favour of 'Berry Picking'

methods used in Berry Picking

Bates defines the figure above as 'the classic information retrieval model'

Bates' 'classic model' broadly corresponds with the conventional approach. The information need is key stage 2 and the documents are identified at key stage 5 and 6 with the match being determined at the end of the process.

Q = query variation
T = thought
E = exit
= documents, informat
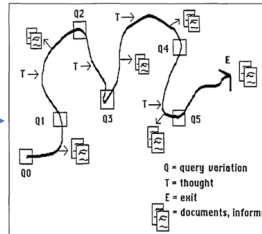
1. footnote chasing;
2. citation chasing;
3. journal run (aka handsearching);
4. area scanning (aka browsing);
5. subject searches in bibliographic databas
6. author searching
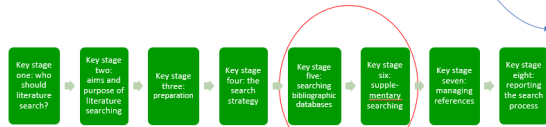
'real searches…use all of the above techni
and more, in endless variation' (Bates, 19

| Key stage one: who should literature search? | Key stage two: aims and purpose of literature searching | Key stage three: preparation | Key stage four: the search strategy | Key stage five: searching bibliographic databases | Key stage six: supplementary searching | Key stage seven: managing references | Key stage eight: reporting the search process |

Figure 7 Kuhlthau 1991

Kuhlthau: The Information Search Process (ISP) (1991)

1. Initiation
'A person first becomes aware of a lack of knowledge'. The task is to recognise the 'information need'. Actions involve: discussing possible topics & approaches

2. Selection/exploration
Selecting the general topic or the approach to be pursued. 'Thoughts center on…assignment requirements, information available, and time allocated' and actions 'involve locating information about the general topic, reading to become informed, and relating new information to what is already known'

3. Formulation
'The turning point' of the model! The task is to form a focus from the information encountered

4. Collection
The interaction between user and the information system. The task is to gather information related to the focused topic.

5. Presenta
The task is to con the search and prepare to prese findings

| Key stage one: who should literature search? | Key stage two: aims and purpose of literature searching | Key stage three: preparation | Key stage four: the search strategy | Key stage five: searching bibliographic databases | Key stage six: supplementary searching | Key stage seven: managing references | Key s eig repo the s pro |

Figure 8 Bystrom 1999

Byström (1999): Information needs, seeking and use (INSU)

Figure 9 Du 2014

Du: the information journey model (2014)

Annex 2. Supplementary material to Publication One, reported in Chapter 3

Below I report the appendix tables to support the literature review presented in Chapter 3 and the search strategy used.

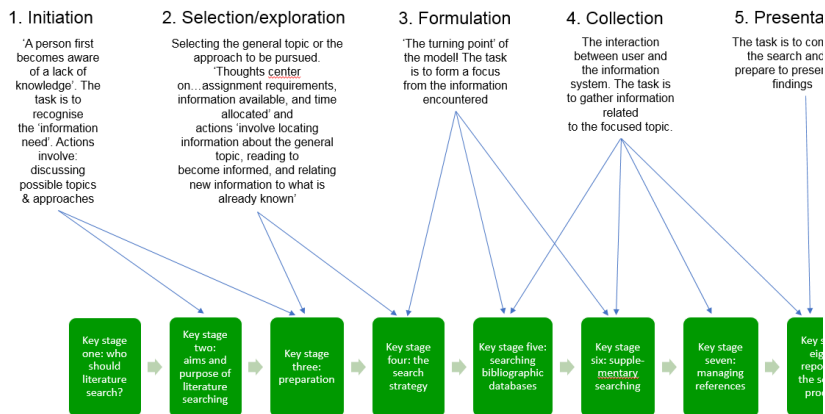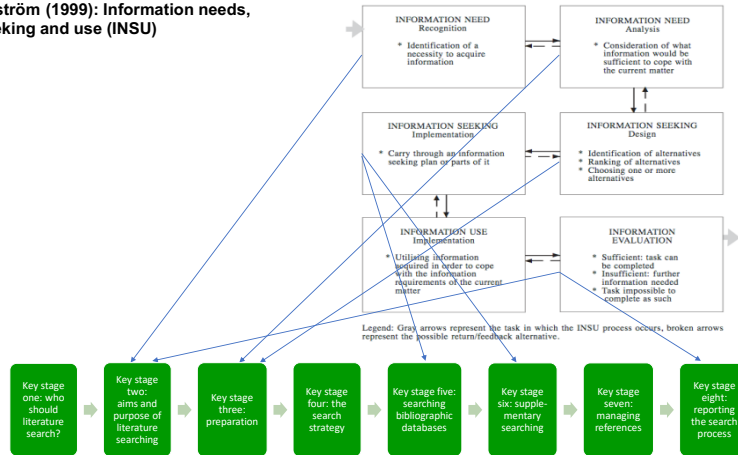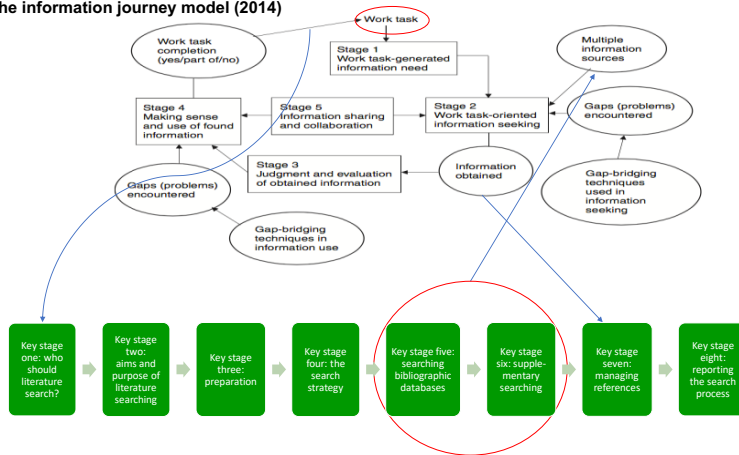| Key stage 1 | Beverley CA, Booth A, Bath PA. The role of the information specialist in the systematic review process: a health information case study. Health information and libraries journal. 2003;20(2):65-74. |
	McGowan J, Sampson M. Systematic reviews need systematic searchers. Journal of the Medical Library Association. 2005;93(1):74-80.
	Meert D, Torabi N, Costella J. Impact of librarians on reporting of the literature searching component of pediatric systematic reviews. Journal of the Medical Library Association : JMLA. 2016;104(4):267-77.
Key stage 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 technology assessment (Winchester, England). 2003;7(1):1-76.
	Fehrmann P, Thomas J. Comprehensive computer searches and reporting in systematic reviews. Research Synthesis Methods. 2011;2(1):15-32.
Key stage 3	No studies identified
Key Sage 4	Methley AM, Campbell S, Chew-Graham C, McNally R, Cheraghi-Sohi S. PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. BMC Health Services Research. 2014;14:579.
	Andrew Booth. Clear and present questions: formulating questions for evidence based practice. Library Hi Tech. 2006;24(3):355-68.
	Cooke A, Smith D, Booth A. Beyond PICO: the SPIDER tool for qualitative evidence synthesis. Qual Health Res. 2012;22(10):1435-43.
Key stage 5	Cooper C, Booth A, Britten N, Garside R. A comparison of results of empirical studies of supplementary search techniques and recommendations in review methodology handbooks: A methodological review. BMC Systematic Reviews (In Press)
Key stage 6	Booth A. Unpacking your literature search toolbox: on search styles and tactics. Health Information & Libraries Journal. 2008;25(4):313-7.
Key stage 7	Bramer WM, Giustini D, de Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. J Med Libr Assoc. 2016;104(3):240-3.
Key stage 8	Rader T, Mann M, Stansfield C, Cooper C, Sampson M. Methods for documenting systematic review searches: a discussion of common issues. Research Synthesis Methods. 2014;5(2):98-115.

| | McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. J Clin Epidemiol. 2016;75:40-6.

Sampson M, McGowan J, Cogo E, Grimshaw J, Moher D, Lefebvre C. An evidence-based practice guideline for the peer review of electronic search strategies. J Clin Epidemiol. 2009;62(9):944-52.

Sampson M MJ, Lefebvre C, Moher D, Ramshaw J. PRESS: Peer Review of Electronic Search Strategies. 2008. |
|---|---|

	The CRD Handbook	The Cochrane Handbook	Collaboration for environmental evidence	Joanna Briggs reviewers manual	IQWiG Methods Resources	Systematic reviews in the social sciences : a practical guide	Eunethta[a]	Campbell Handbook	Dev. NICE guide man.
Purpose of guidance	Reviews of healthcare interventions*	Reviews of interventions*	Reviews of environmental evidence	Reviews of qualitative studies*	Reviews of healthcare interventions	Reviews in social topics	Reviews of health technologies	Reviews of social sciences	Dev. NIC
Purpose of literature searching	Conducting a **thorough** search to identify relevant studies is a key factor in minimizing bias in the review process. The search process should be as **transparent** as possible and documented in a way that enables it to be evaluated and reproduced.	Systematic reviews of interventions require a **thorough,** objective and **reproducible** search of a range of sources to identify as many relevant studies as possible (within resource limits).	Systematic and **comprehensive** searching for relevant studies is essential to minimise publication bias in a SR and to assess the strength of the evidence case	The search strategy aims to find both published and unpublished studies.	A systematic literature search aims to identify **all publications relevant** to the particular research question (i.e. publications that contribute to a gain in knowledge on the topic). The search for primary literature is normally orientated towards the aim of achieving high sensitivity.	The aim of the literature search is **not to retrieve everything. It is to retrieve everything of relevance**, while leaving behind the irrelevant.	Information retrieval for systematic reviews needs to be performed in a **thorough, transparent and reproducible manner.** The aim is to identify all relevant studies and study results on the question of interest (within resource limits) [3]. This requires both searches in several information sources and the use of comprehensive search strategies [3-5]. This approach is a key factor in minimizing	Systematic reviews of interventions require a **thorough, objective and reproducible** search of a range of sources to identify as many relevant studies as possible (within resource limits).	The **iden best evid** addr parti ques prod unm volu resu

1

	Pg 4 1.1.1 The review team	Section 6.1 Introduction	Pg 28	Pg 57	No data reported	Pg 85	Pg 14	Pg 8	No Data reported
							bias in the review process [5].		
Who should undertake literature searching	'Ideally these should include expertise in systematic review methods, information retrieval,'	'Cochrane Review Groups (CRGs) are responsible for providing review authors with references to studies that are possibly relevant to their review. The majority of CRGs employ a dedicated Trials Search Co-ordinator to provide this service' 6.1.1.1 Role of the Trials Search Co-ordinator	'enlisting an information specialist in the review team is recommended so that an efficient search strategy can be established.'	'if possible, authors should always seek the advice of a research librarian in the construction of a search strategy.'		It will be clear from the above description that there are particular skills involved in carrying out sensitive and specific searches, and these tend not to be part of the core training of social researchers. This is why systematic reviewers usually seek expert help from an information scientist, as these specialists are trained in searching electronic and other sources. They will know which databases are available and how to access them through the various interfaces and providers, as well as differences between them with respect to search terms and	Information specialists should form an integral part of the project team of a systematic review from the beginning of the project.	This guide's fundamental premise is that information retrieval is an essential component of the systematic review process, analogous to the data collection phase of a primary research study, and requires the expertise of TSC, an information specialist (IS) or a librarian.	

		'The Trials Search Co-ordinator for each CRG is responsible for providing assistance to authors with searching for studies for inclusion in their reviews. The range of assistance varies according to the resources available to individual CRGs but may include some or all of the following: providing relevant studies from the CRG's Specialized Register (see Section 6.3.2.4 for more detail), designing search strategies for the main bibliographic databases, running these searches in				indexing. They are also skilled in locating gray literature.			

		databases available to the CRG, saving search results and sending them to authors, advising authors on how to run searches in other databases and how to download results into their reference management software' section 6.3.1 provides guidance on how involve trials search co-ordinators							
Preparation/ Scoping	Pg 3 Determining if there are already existing or ongoing review or if a new review is	6.1.2 Cochrane review authors should seek advice from the Trials Search Co-ordinator of their Cochrane	Pg 26-27 A thorough scope should entail:	Pg 11 In order to avoid duplication, reviewers are advised to register their review title as	Not reported	Pg 80 'There are many sources of information to consider, but a search of electronic databases is often the main starting	Section 2.3.2 Conducting preliminary searches 'At the start of a project – before the development of the actual search	Not reported	Not reported

justified. Guidance on where to scope is provided.	Review Group (CRG) *before* starting a search.	The development and testing of a search strategy. An estimate of the volume of relevant literature. Critical appraisal of study quality and data extraction of a small subset of relevant papers. An estimate of resources required based on the above. The expected output from a scoping exercise is an	mentioned previously. It is also recommended that reviewers search major electronic databases to determine that there have been no recently published systematic reviews on the same topic prior to registration of a review title. A search of the Joanna Briggs Institute Library of Systematic Review Protocols, Joanna Briggs Institute Library of Systematic Reviews, Cochrane Library, MEDLINE, PROSPERO		point. However such databases are not the only source of literature, and sometimes they are not even the most useful.' 'The type of information being sought will depend on both the review question and the inclusion criteria.' Guidance on how to broaden these 'scoping' searches is given, including: searching unpublished reports, conference proceedings and gray literature. The advice that searches should include book chapters is unique.	strategy – a preliminary search (also known as a scoping search) should be conducted. This preliminary search has various goals.' 'Firstly, to help prepare the overall project [32], i.e. understanding the key questions [4], identifying existing systematic reviews [5,50,51], identifying a first set of potentially relevant primary studies [52], and estimating the resources necessary to perform the systematic review [50]. Secondly, the results of the preliminary search can be used in the development of the search strategy, for instance, by generating a list of search terms from the analysis of		

			estimate of the quantity and quality of	and DARE databases will assist to establish			identified relevant articles [4,53-55] and subsequently used in the		
			evidence, and a characterisation of the likely evidence base, pertaining to the question	whether or not a recent review report exists on the topic of interest.			development of the search strategy.'		
			(see Box 3 for example). The extent of investment in review scoping is a matter of				'The most effective way of conducting a preliminary search is first to search for systematic reviews.'		
			judgement and will differ with each review.						
The search strategy	Pg 19

PICOs

Pg 12 | 6.4.1

'The eligibility criteria for studies to be included in the review will inform how the | Pg 28

'This may include considering synonyms, alternative spellings, and | Pg 30

'Limiting the search by date may be used where the focus of the review is on a | 7.1.7

'The languages of publication are usually restricted to those of Western | Pg 81

'For an intervention, this involves listing the different ways in which it can be defined, perhaps by drawing up a list of | 2.3.3

PICOs

'The research question is | Pg 23

'The structure of a search strategy should be based on the main concepts | 5.4

'Review questions can be broken down into different concepts, which can be |

	'Whenever feasible, all relevant studies should be included regardless of language. However, realistically this is not always possible due to a lack of time, resources and facilities for translation.'	search is conducted. The eligibility criteria will specify the types of designs, types of participants, types of intervention (experimental and comparator) and, in some cases, the types of outcomes to be addressed.'	non-English language terms within the search strategy. An initial list of search terms may be compiled with the help of the commissioning organisation and stakeholders.' Pg 29 'A final step in the development of the search terms is to test the strategy with a set of known relevant articles (these may often be provide	more recent intervention or innovation. However, potentially relevant studies as well as seminal, early studies in the field may be missed if the limit set is too recent thus date limits should be used in an informed way, based on knowledge of key papers relevant to the review question' p30 'If limiting by language is required, it is preferable to	Europe. However, other foreign-language publications may also be included if the available information on these publications indicates that additional and relevant information for answering the research question is to be expected.'	synonyms. Then the population is specified, and the outcome of interest – again, including relevant synonyms.' 'For social systematic reviews of effectiveness the range of eligible study designs may be wide. Searching for controlled trials alone may either uncover few studies or may not identify other relevant evaluative research, and is likely to exclude studies reporting on process and implementation issues. Similarly, evaluations of the effects of social policies may involve randomized and non- randomized controlled studies, but a range of other study designs and search terms will be relevant depending	commonly broken into concepts, and only the most important ones are used to develop the search strategy' ' For more complex review questions, it may be necessary to use several combinations of search concepts to capture a review topic' 'If search strategies are limited, for example, by language or publication year, this should be justified in the methods section of the systematic review. However, such limits should be used with caution, as they may introduce bias [3,4,10]. Moreover,	being examined in a review.' P24 'Generally speaking, a search strategy to identify intervention studies will typically have three sets of terms: 1) the condition of interest, i.e., the population; 2) the intervention(s) evaluated; and 3) the outcomes (optional). Limiting commands may be used to further narrow the results by study design (or document	combined to devise a search strategy. For example, the PICO (population, intervention, comparator and outcome) or the SPICE (setting, perspective, intervention, comparison, evaluation; Booth 2004) framework can be used to structure a search strategy.' 'When the relevant literature for a question is less well defined or indexed, a multi-stranded approach to searching may be more efficient. This involves developing several shorter search

			'd by review commissioners or subject experts or have been selected for the trial critical appraisal'	search inclusively, and keep a record of numbers of studies per language group. This allows the reader to identify how many studies have been identified, but are not included, therefore promoting transparency in the process'		on the study question.'	they should only be considered if they can be reliably applied'	type), dates, language, etc'.	strategies (strands) with an emphasis on precision.'
Database searching	Pg 17 'The selection of electronic databases to search will depend upon the review topic.'	6.2.1.1 'Searches of health-related bibliographic databases are generally the easiest and least time-consuming way to identify an	Pg 36 'Different databases and catalogues sample different subsets of the literature, and so multiple	Pg 29 'An initial limited search of MEDLINE and CINAHL will be undertaken followed by analysis of the text words	7.1.2 'The selection of databases for each product is generally based on the focus (i.e. regarding content,	Pg 101 'The number of databases or other sources that one needs to search varies from topic to topic, and depends on the time and resources available. It also clearly	2.3.4 'The production of a systematic review requires a systematic search in several bibliographic databases. For example, previous research has	3.1. pg 11 'Social science-related subject databases are generally the best way to identify an initial set of relevant	5.3 'The selection of sources will vary according to the requirements of the review question. For reviews of the effectiveness of pharmacological

17

	'Due to the diversity of questions addressed by systematic reviews, there can be no agreed standard for what constitutes an acceptable search in terms of the number of databases searched. For example, if the review is on a cross-cutting public health topic such as housing and health it is advisable to search a wider range of databases than if the review is of a pharmaceutical intervention	initial set of relevant reports of studies.' 'A key advantage of these databases is that they can be searched electronically both for words in the title or abstract and by using the standardized indexing terms, or controlled vocabulary, assigned to each record.' Further guidance is provided on national and regional databases, subject-specific databases, citation indexes and grey	sources should be accessed to ensure the search is comprehensive and unbiased, but avoids unnecessary duplication.' P37 'Different Review Teams often have access to different resources, and so the list of resources searched for each review will vary, but checking bibliographies and contact with authors should help to test if relevant	contained in the title and abstract, and of the index terms used to describe article. A second search using all identified keywords and index terms will then be undertaken across all included databases.'	methods, and region) of the bibliographic databases. At least 2 large biomedical databases (e.g. MEDLINE and EMBASE) are always selected. For the preparation of health information a search for qualitative studies is additionally conducted in CINAHL and PsycInfo.'	depends on one's tolerance to the risk of missing a relevant study or studies, and one's assessment of the cost of missing it.'	shown that searching MEDLINE alone is insufficient to identify all published relevant studies on the topic of interest and may produce biased results [64-67]. This is due to the fact that journal inclusion rates differ between databases [68,69]. Furthermore, the time and quality of indexing differs [65,69-71], meaning that a reference might be more difficult to find or be found with delay in some databases, but not in others.' 'insufficient empirical evidence is available so far on how many and which databases should be regularly searched.'	reports of studies within a specific field.' ' Decisions related to which subject-specific databases are to be searched, in addition to the main field-related database, will be influenced by the topic of the review, access to specific databases, and budget considerations.'	interventions, the Cochrane Central Register of Controlled Trials (CENTRAL), EMBASE and MEDLINE should be prioritised for searching. For other questions, it might be as or more important to search other sources.'

	for a known health condition'	literature databases.	articles are retrieved.'						
Supplementary searching	P17 'In addition to searching electronic databases, published and unpublished research may also be obtained by using one or more of the following methods.'	6.2.4 'Conference abstracts and other grey literature can be an important source of studies for inclusion in reviews.' 'Efforts should be made to identify unpublished studies.'	General guidance on methods reported in table 2 is reported. No specific guidance on aims or purpose.	No specific guidance on aims or purpose.	7.1.6 'Besides bibliographical database searches, it can be useful (depending on the research question) to conduct a handsearch in selected scientific journals and proceedings of abstracts from scientific meetings. This is decided on a case-by-case basis.'	4.2 'databases are not the only source of literature, and sometimes they are not even the most useful. In many research areas, particularly in the social sciences, the bulk of the relevant evidence may not appear in journals, but will be located in reports in the "gray literature," much of which may not be indexed in electronic databases. There is thus a real risk that electronic searches alone will fail to locate a good deal of relevant information.'	General guidance on methods reported in table 2 is reported. No specific guidance on aims or purpose.	General guidance on methods reported in table 2 is reported. No specific guidance on aims or purpose.	General guidance on methods reported in table 2 is reported. No specific guidance on aims or purpose.

Managing references	1.3.1.8 advocates the use of bibliographic management software	6.5 No guidance reported	4.1.6 guidance on bibliographic software, on downloading data where a direct export option in not available, and saving the results	No guidance reported	4.16 guidance on bibliographic software is given	No guidance reported	No guidance reported	6.1 guidance on bibliographic software.	5.8 guidance on bibliographic software.
Documenting the search	1.3.1.10 No specific guidance was recommend. The following criteria were documented for repotting: The write up of the search should include information	6.6 PRISMA guidance is necessary. 'the full search strategies for each database will need to be included in an Appendix of the review. The search strategies will need to be copied and pasted exactly	4.1.5 no specific reporting guidelines mentioned but a review by Mant et al was referenced. The following criteria was documented for reporting:	No guidance reported	No guidance reported	7.1.8 no specific guidance was recommend. The following criteria were documented: All steps in the search in bibliographic databases are documented. This especially includes:	No guidance reported	8.1.2 no specific guidance was recommend The following criteria were documented for reporting: Reporting the search process in the review:	5.9 specific criteria were documented for reporting. the following information should be documented: date(s) on which the searches were carried out, including

	about the databases and interfaces searched (including the dates covered), full detailed search strategies (including any justifications for date or language restrictions) and the number of records retrieved.	as run and included in full, together with the search set numbers and the number of records retrieved. The search strategies should not be re-typed as this can introduce errors.' 'save locally or file print copies of any information found on the internet, such as information about ongoing trials.'	Specifically, for each source searched a record should be made of: the dates of individual searches; the full list of search terms employed and how these were combined; any changes to the default search settings of the source used; the nature of the search (e.g. keywords, topics, or full texts) and other search options (e.g. lemmatization); the			the search strategy for the databases selected the search date the user interface the number of hits after perusal of all hits: documentation of the publications judged relevant to the research question posed (citations) after perusal of the full texts: documentation of the citations not judged relevant; alternatively, documentation of the topic-related publications that were, however,		List all databases searched; Note the dates of the last search for each database AND the period searched; Note any language or publication status restrictions; List grey literature sources; List individuals or organizations contacted;	the date(s) of any re-run searches (see section 5.10) names of the databases, database host systems and database coverage dates names of any other sources searched search strategies for all sources, annotated to explain any decisions on included and excluded terms which are not self

			removal of duplicates if automatically carried out when downloading results; and all the results returned by each search.			irrelevant for the report (in each case providing a reason for exclusion) All other steps in the information retrieval procedure are also documented (e.g. correspondence with authors, queries to manufacturers).		List any journals and conference proceedings specifically handsearched for the review; List any other sources searched (e.g. reference lists, the internet).	-explanatory details of any supplementary searching undertaken, including the rationale any limits or search filters applied to the search (for example, language, date, study design).

Key: * These handbooks focus on specific guidance (as per their topic) in the first instance. Subsequent chapters within the handbooks focus on other review topics.

These include: reviews of diagnostic and prognostic studies, review of public health topics and reviews of qualitative studies. These chapters develop the primary

guidance accounting for peculiarities specific to the topic. I have not focused on these additional topics here.

PubMed search strategy

Database: PubMed

Host: National library for medicine (NLM) via:
https://www.ncbi.nlm.nih.gov/pubmed/

Date searched: August 30th 2017

Searcher: Chris Cooper

Checked by: Jo Varley-Campbell

Search strategy: (literature search*[Title/Abstract]) AND sysrev_methods[sb]

Results: 586

Annex 3. Supplementary material to Publication Two, reported in Chapter 6

Below I report the supplementary material for the systematic review presented in Chapter 6.

Search strategies

Database: Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily, Ovid MEDLINE and Versions(R)

Host: OVID

Data Parameters: 1946-Current

Search date: 23/02/2017

Hits: 6085

Search strategy:

#	Searches	Results
1	(((sensitivity and specificity) or precision or recall or number needed or NNR or capture-recapture or (capture adj3 recapture) or (mark adj3 capture)) and (search* or literature or stud* or systematic or database* or published or unpublished)).ti,kw.	1877
2	(search* and (assess* or identify or evaluat* or performance or valid* or reliability or optimal* or accuracy or precision or comprehen* or coverage or recall or effect* or efficien* or inclusion or compar* or performance or stop* or estimat* or predict* or miss* or omission*)).ti.	4288
3	1 or 2	6085

Notes: N/A

Saved: MEDLINE6085.txt

Database: EMBASE

Host: OVID

Data Parameters: 1974 to 2017 February 22

Search date: 23/02/2017

Hits: 8194

Search strategy:

Search Strategy:

#	Searches	Results
1	(((sensitivity and specificity) or precision or recall or number needed or NNR or capture-recapture or (capture adj3 recapture) or (mark adj3 capture)) and (search* or literature or stud* or systematic or database* or published or unpublished)).ti,kw.	3597
2	(search* and (assess* or identify or evaluat* or performance or valid* or reliability or optimal* or accuracy or precision or comprehen* or coverage or recall or effect* or efficien* or inclusion or compar* or performance or stop* or estimat* or predict* or miss* or omission*)).ti.	4685
3	1 or 2	8194

F Score searches

As reported in section 6.3.5.4., in peer review, I was asked why F Score had not been included. Simply, it was because no relevant studies were identified using F Score. I set out the searches used to respond to the reviewer on this point.

Database: Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily, Ovid MEDLINE and Versions(R)

Host: OVID

Data Parameters: 1946-Current

Search date: 31/01/2018

Hits: 18

Search strategy:

| 1 | ("f score" or "fscore" or "f-score" or "f 1score" or "f1score" or "f measure" or "f measure").ti,kw. | 18 |

Notes: N/A

Saved:
MEDLINE18.txt

Database: EMBASE

Host: OVID

Data Parameters: 1974 to 2017 February 22

Search date: 31/01/2018

Hits: 30

Search strategy:

1 | ("f score" or "fscore" or "f-score" or "f 1score" or "f1score" or "f measure" or "f measure").ti,kw. | 30

Notes: N/A

Saved: Embase30.txt

Database: Library, Information Science & Technology Abstracts (LISTA)

Host: EBSCOhost

Data Parameters: 1960-Current

Search date: 31/01/2018

Hits: 4

Search strategy:

S1 TI (("f score" or "fscore" or "f-score" or "f 1score" or "f1score" or "f measure" or "f measure")) OR KW (("f score" or "fscore" or "f-score" or "f 1score" or "f1score" or "f measure" or "f measure"))
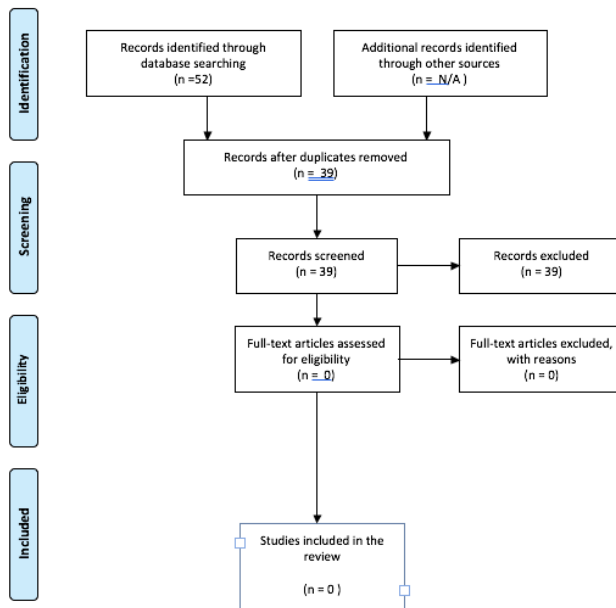
Notes: N/A

Saved: LISTA.txt

PRISMA flow chart for the F-Score searching

27

PRISMA 2009 Flow Diagram

Records identified through database searching
(n =52)

Additional records identified through other sources
(n = N/A)

Records after duplicates removed
(n = 39)

Records screened
(n = 39)

Records excluded
(n = 39)

Full-text articles assessed for eligibility
(n = 0)

Full-text articles excluded, with reasons
(n = 0)

Studies included in the review
(n = 0)

Identification

Screening

Eligibility

Included

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement.* PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit www.prisma-statement.org.

Studies excluded at full-text screening

1.   Agoritsas T, Merglen A, Courvoisier DS, Combescure C, Garin N, Perrier A, et al. Sensitivity and predictive value of 15 PubMed search strategies to answer clinical questions rated against full systematic reviews. Journal of Medical Internet Research. 2012;14(3):e85.

2.   Ahmadvand A, Hadjibabaie M, Gholami K. "Precision" or "relevance"? Reporting considerations for articles focusing on evaluation of search strategies. Journal of Clinical Epidemiology. 2013;66(3):351-2.

3.   Akaho E, Bandai A, Miyake A. Evaluation and analysis of descriptor-oriented search for eight topics using Excerpta Medica. Journal of Social and Administrative Pharmacy. 1986;4(1):10-5.

4.   Al Hajeri A, Z F. A comparison of handsearching versus EMBASE searching of the Bahrain Medical Bulletin to identify reports of randomized controlled trials. Bahrain Medical Bulletin. 2005;27(4):160-3.

5.   Allison JJ, Kiefe CI, Weissman NW, Carter J, Centor RM. The art and science of searching MEDLINE to answer clinical questions: Finding the right number of articles. International Journal of Technology Assessment in Health Care. 1999;15(2):281-96.

6.   Alton V. To make an optimal search in PubMed: Strategies to find clinically relevant information of sufficient quality. [Swedish]. Lakartidningen. 2005;102(45):3312-7.

7.   Bayliss SE, Davenport CF, Pennant ME. Where and how to search for information on the effectiveness of public health interventions--a case study for prevention of cardiovascular disease. Health Information & Libraries Journal. 2014;31(4):303-13.

8.   Behm K, Baier M. Evaluating online literature searches. Nurse Educator. 1998;23(2):9-10.

9.   Bennett DA, Latham NK, Stretton C, Anderson CS. Capture-recapture is a potentially useful method for assessing publication bias. Journal of Clinical Epidemiology. 2004;57(4):349-57.

10.   Beynon R, Leeflang MM, McDonald S, Eisinga A, Mitchell RL, Whiting P, et al. Search strategies to identify diagnostic accuracy studies in MEDLINE and EMBASE. Cochrane Database of Systematic Reviews. 2013(9):MR000022.

11.   Boeker M, Vach W, Motschall E. Google Scholar as replacement for systematic literature searches: good relative recall and precision are not enough. BMC Medical Research Methodology. 2013;13:131.

12.   Bramer WM, Giustini D, Kramer BM. Comparing the coverage, recall, and precision of searches for 120 systematic reviews in Embase, MEDLINE, and Google Scholar: a prospective study. Systematic Reviews. 2016;5:39.

13.   Bramer WM, Giustini D, Kramer BM, Anderson P. The comparative recall of Google Scholar versus PubMed in identical searches for biomedical systematic reviews: a review of searches used in systematic reviews. Systematic Reviews. 2013;2:115.

14.   Buchberger B, Krabbe L, Lux B, Mattivi JT. Evidence mapping for decision making: feasibility versus accuracy - when to abandon high sensitivity in electronic searches. German Medical Science. 2016;14:Doc09.

15.   Chapman AL, Morgan LC, Gartlehner G. Semi-automating the manual literature search for systematic reviews increases efficiency. Health Information & Libraries Journal. 2010;27(1):22-7.

16.   Chow TK, To E, Goodchild CS, McNeil JJ. A simple, fast, easy method to identify the evidence base in pain-relief research: validation of a computer search strategy used alone to identify quality randomized controlled trials. Anesthesia & Analgesia. 2004;98(6):1557-65, table of contents.

17.   Cooper C, Levay P, Lorenc T, Craig GM. A population search filter for hard-to-reach populations increased search efficiency for a systematic review. Journal of Clinical Epidemiology. 2014;67(5):554-9.

18.   Craven J, Jefferies J, Kendrick J, Nicholls D, Boynton J, Frankish R. A comparison of searching the Cochrane library databases via CRD, Ovid and Wiley: implications for systematic searching and information services. Health Information & Libraries Journal. 2014;31(1):54-63.

19.   Curnutt GL, Curnutt KL. Reactive chemical hedges: a search tool for comprehensive retrieval of chemical safety data. Journal of Chemical Information & Computer Sciences. 1991;31(1):116-9.

20.   Day D, Furlan A, Irvin E, Bombardier C. Simplified search strategies were effective in identifying clinical trials of pharmaceuticals and physical modalities. Journal of Clinical Epidemiology. 2005;58(9):874-81.

21.   DeNeef P. The comprehensiveness of computer-assisted searches of the medical literature. Journal of Family Practice. 1988;27(4):404-8.

22.   Doust JA, Pietrzak E, Sanders S, Glasziou PP. Identifying studies for systematic reviews of diagnostic tests was difficult due to the poor sensitivity

and precision of methodologic filters and the lack of information in the abstract. Journal of Clinical Epidemiology. 2005;58(5):444-9.

23.   Durao S, Kredo T, Volmink J. Validation of a search strategy to identify nutrition trials in PubMed using the relative recall method. Journal of Clinical Epidemiology. 2015;68(6):610-6.

24.   Dykeman J, Faris P, Dykeman M, Jette N, Wiebe S. Evaluating the use of capture-mark-recapture for determining the number of missing articles in systematic reviews. American Journal of Epidemiology. 2012;175:S20.

25.   Egan M, Maclean A, Sweeting H, Hunt K. Comparing the effectiveness of using generic and specific search terms in electronic databases to identify health outcomes for a systematic review: a prospective comparative study of literature search methods. BMJ Open. 2012;2(3).

26.   Eisinga A, Siegfried N, Clarke M. The sensitivity and precision of search terms in Phases I, II and III of the Cochrane Highly Sensitive Search Strategy for identifying reports of randomized trials in medline in a specific area of health care--HIV/AIDS prevention and treatment interventions. Health Information & Libraries Journal. 2007;24(2):103-9.

27.   Elchesen DR. Cost-effectiveness comparison of manual and on-line retrospective bibliographic searching. Journal of the American Society for Information Science. 1978;29(2):56-66.

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

29.   Farrah K, Mierzwinski-Urban M, Cimon K. Effectiveness of adverse effects search filters: drugs versus medical devices. Journal of the Medical Library Association. 2016;104(3):221-5.

30.   Fehrmann P, Thomas J. Comprehensive computer searches and reporting in systematic reviews. Research Synthesis Methods. 2011;2(1):15-32.

31.   Ferrante di Ruffano L, Davenport C, Eisinga A, Hyde C, Deeks JJ. A capture-recapture analysis demonstrated that randomized controlled trials

evaluating the impact of diagnostic tests on patient outcomes are rare. Journal of Clinical Epidemiology.

32. Fisher HL, Elchesen DR. Effectiveness of combining title words and index terms in machine retrieval searches. Nature. 1972;238(5359):109-10.

33. Flemming K, Briggs M. Electronic searching to locate qualitative research: evaluation of three strategies. Journal of Advanced Nursing. 2007;57(1):95-100.

34. Forsetlund L, Kirkehei I, Harboe I, Odgaard-Jensen J. A comparison of two search methods for determining the scope of systematic reviews and health technology assessments. International Journal of Technology Assessment in Health Care. 2012;28(1):59-64.

35. Fortier KJ, Kiss N, Tongbram V. What is the optimal search engine for results from Embase and medline: Ovid or embase.com? Value in Health. 2013;16 (3):A25.

36. Furlan AD, Irvin E, Bombardier C. Limited search strategies were effective in finding relevant nonrandomized studies. Journal of Clinical Epidemiology. 2006;59(12):1303-11.

37. Gall C, Brahmi FA. Retrieval comparison of EndNote to search MEDLINE (Ovid and PubMed) versus searching them directly. Medical Reference Services Quarterly. 2004;23(3):25-32.

38. Glanville J, Kaunelis D, Mensinkai S. How well do search filters perform in identifying economic evaluations in MEDLINE and EMBASE. International Journal of Technology Assessment in Health Care. 2009;25(4):522-9.

39. Glanville J, Kaunelis D, Mensinkai S. Identifying economic evaluations in medline and embase: How well do published search filters perform? Value in Health. 2009;12 (7):A391.

40. Glanville J, Paisley S. Identifying economic evaluations for health technology assessment. International Journal of Technology Assessment in Health Care. 2010;26(4):436-40.

41.   Golder S, Loke Y. Search strategies to identify information on adverse effects: a systematic review. Journal of the Medical Library Association. 2009;97(2):84-92.

42.   Golder S, Loke YK. The performance of adverse effects search filters in MEDLINE and EMBASE. Health Information & Libraries Journal. 2012;29(2):141-51.

43.   Golder S, Loke YK. Failure or success of electronic search strategies to identify adverse effects data. Journal of the Medical Library Association. 2012;100(2):130-4.

44.   Golder S, Loke YK. Sensitivity and precision of adverse effects search filters in MEDLINE and EMBASE: a case study of fractures with thiazolidinediones. Health Information & Libraries Journal. 2012;29(1):28-38.

45.   Golder S, Loke YK, Zorzela L. Comparison of search strategies in systematic reviews of adverse effects to other systematic reviews. Health Information & Libraries Journal. 2014;31(2):92-105.

46.   Golder S, McIntosh HM, Duffy S, Glanville J. Developing efficient search strategies to identify reports of adverse effects in MEDLINE and EMBASE. Health information and libraries journal. 2006;23(1):3-12.

47.   Gorecki CA, Brown JM, Briggs M, Nixon J. Evaluation of five search strategies in retrieving qualitative patient-reported electronic data on the impact of pressure ulcers on quality of life. Journal of Advanced Nursing. 2010;66(3):645-52.

48.   Gosch F, Lux B, Buchberger B. Development and validation of a search filter to retrieve publications on eHealth-project description. European Journal of Epidemiology. 2016;31:S68.

49.   Gotzsche PC, Lange B. Comparison of search strategies for recalling double-blind trials from MEDLINE. Danish Medical Bulletin. 1991;38(6):476-8.

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

33

51.   Griffon N, Schuers M, Dhombres F, Merabti T, Kerdelhue G, Rollin L, et al. Searching for rare diseases in PubMed: a blind comparison of Orphanet expert query and query based on terminological knowledge. BMC Medical Informatics & Decision Making. 2016;16:101.

52.   Haase A, Follmann M, Skipka G, Kirchner H. Developing search strategies for clinical practice guidelines in SUMSearch and Google Scholar and assessing their retrieval performance. BMC Medical Research Methodology. 2007;7:28.

53.   Harbour J, Fraser C, Lefebvre C, Glanville J, Beale S, Boachie C, et al. Reporting methodological search filter performance comparisons: a literature review. Health Information & Libraries Journal. 2014;31(3):176-94.

54.   Hariri N, Emami Z, Malek M. The precision and recall of general search engines in retreival of images related to endocrine diseases. [Persian]. Iranian Journal of Endocrinology and Metabolism. 2015;17(2):97-104.

55.   Hariri N, Ravandi SN. Comparing the Precision of Information Retrieval of MeSH-Controlled Vocabulary Search Method and a Visual Method in the Medline Medical Database. Electronic Physician [Electronic Resource]. 2014;6(2):832-7.

56.   Harrison J. Designing a search strategy to identify and retrieve articles on evidence-based health care using MEDLINE. Health Libraries Review. 1997;14(1):33-42.

57.   Hausner E, Guddat C, Hermanns T, Lampert U, Waffenschmidt S. Development of search strategies for systematic reviews: validation showed the noninferiority of the objective approach. Journal of Clinical Epidemiology. 2015;68(2):191-9.

58.   Hausner E, Guddat C, Hermanns T, Lampert U, Waffenschmidt S. Prospective comparison of search strategies for systematic reviews: an objective approach yielded higher sensitivity than a conceptual one. Journal of Clinical Epidemiology. 2016;77:118-24.

59.   Hawkins N, Scott DA, Woods B. How far do you go? Efficient searching for indirect evidence. Medical Decision Making. 2009;29(3):273-81.

60.   Haynes RB, McKibbon KA, Walker CJ. Computer searching of the medical literature. An evaluation of MEDLINE searching systems. Annals of Internal Medicine. 1985;103(5):812-6.

61.   Hempel S, Rubenstein LV, Shanman RM, Foy R, Golder S, Danz M, et al. Identifying quality improvement intervention publications--a comparison of electronic search strategies. Implementation Science. 2011;6:85.

62.   Hersh WR, Hickam DH. A comparison of retrieval effectiveness for three methods of indexing medical literature. American Journal of the Medical Sciences. 1992;303(5):292-300.

63.   Hildebrand AM, Iansavichus AV, Lee CW, Haynes RB, Wilczynski NL, McKibbon KA, et al. Glomerular disease search filters for Pubmed, Ovid Medline, and Embase: a development and validation study. BMC Medical Informatics & Decision Making. 2012;12:49.

64.   Hoogendam A, de Vries Robbe PF, Stalenhoef AF, Overbeke AJ. Evaluation of PubMed filters used for evidence-based searching: validation using relative recall. Journal of the Medical Library Association. 2009;97(3):186-93.

65.   Hooijmans CR, Tillema A, Leenaars M, Ritskes-Hoitinga M. Enhancing search efficiency by means of a search filter for finding all studies on animal experimentation in PubMed. Laboratory Animals. 2010;44(3):170-5.

66.   Hopewell S, Clarke M, Lefebvre C, Scherer R. Handsearching versus electronic searching to identify reports of randomized trials. Cochrane Database of Systematic Reviews. 2007(2):MR000001.

67.   Hsu D, Brieva J, Nardone B, Silverberg JI. Validation of database search strategies for the epidemiological study of pemphigus and pemphigoid. British Journal of Dermatology. 2016;174(3):645-8.

68.   Hsu DY, Nardone B, West D, Silverberg JI. Validation of database search strategies for the epidemiological study of eczema herpeticum. British Journal of Dermatology. 2016;175(1):220-2.

69.   Huang Y, Yang Z, Wang J, Zhuo L, Li Z, Zhan S. Performance of search strategies to retrieve systematic reviews of diagnostic test accuracy from the Cochrane Library. Journal of Evidence based Medicine. 2016:06.

70.   Iansavichus AV, Haynes RB, Shariff SZ, Weir M, Wilczynski NL, McKibbon A, et al. Optimal search filters for renal information in EMBASE. American Journal of Kidney Diseases. 2010;56(1):14-22.

71.   Iansavichus AV, Hildebrand AM, Haynes RB, Wilczynski NL, Levin A, Hemmelgarn BR, et al. High-performance information search filters for CKD content in pubmed, ovid MEDLINE, and EMBASE. American Journal of Kidney Diseases. 2014;65(1):26-32.

72.   Jenkins M. Evaluation of methodological search filters--a review. Health Information & Libraries Journal. 2004;21(3):148-63.

73.   Johnson N, Tongbram V, Ndirangu K, Ogden K, Bay C. Development of a systematic review search filter to identify medication adherence studies. Value in Health. 2016;19 (3):A72.

74.   Kastner M, Haynes RB, Wilczynski NL. Inclusion of methodological filters in searches for diagnostic test accuracy studies misses relevant studies. Journal of Clinical Epidemiology. 2012;65(1):116-7; author reply 7-8.

75.   Kastner M, Straus SE, McKibbon KA, Goldsmith CH. The capture-mark-recapture technique can be used as a stopping rule when searching in systematic reviews. Journal of Clinical Epidemiology. 2009;62(2):149-57.

76.   Kastner M, Wilczynski NL, McKibbon AK, Garg AX, Haynes RB. Diagnostic test systematic reviews: bibliographic search filters ("Clinical Queries") for diagnostic accuracy studies perform well. Journal of Clinical Epidemiology. 2009;62(9):974-81.

77.   Keeling JW, Turner AM, Allen EE, Rowe SA, Merrill JA, Liddy ED, et al. Development and evaluation of a prototype search engine to meet public health information needs. AMIA Annual Symposium Proceedings/AMIA Symposium. 2011;2011:693-700.

78.   Kleijnen J, Knipschild P. The comprehensiveness of Medline and Embase computer searches. Searches for controlled trials of homoeopathy,

ascorbic acid for common cold and ginkgo biloba for cerebral insufficiency and intermittent claudication. Pharmaceutisch Weekblad - Scientific Edition. 1992;14(5):316-20.

79.   Kok R, Verbeek JA, Faber B, van Dijk FJ, Hoving JL. A search strategy to identify studies on the prognosis of work disability: a diagnostic test framework. BMJ Open. 2015;5(5):e006315.

80.   Kwak N, Swan JT, Thompson-Moore N, Liebl MG. Validity and Reliability of a Systematic Database Search Strategy to Identify Publications Resulting From Pharmacy Residency Research Projects. Journal of Pharmacy Practice. 2016;29(4):406-14.

81.   Ladhani M, Rogerson T, Mitchel R, Craig JC, Webster AC. Search strategies for locating diagnostic test accuracy (DTA) studies in nephrology journals. Nephrology. 2012;17:52.

82.   Lane D, Dykeman J, Ferri M, Goldsmith CH, Stelfox HT. Capture-mark-recapture as a tool for estimating the number of articles available for systematic reviews in critical care medicine. Journal of Critical Care. 2013;28(4):469-75.

83.   Langham J, Thompson E, Rowan K. Identification of randomized controlled trials from the emergency medicine literature: comparison of hand searching versus MEDLINE searching. Annals of Emergency Medicine. 1999;34(1):25-34.

84.   Leclercq E, Leeflang MM, van Dalen EC, Kremer LC. Validation of search filters for identifying pediatric studies in PubMed. Journal of Pediatrics. 2013;162(3):629-34.e2.

85.   Lee E, Dobbins M, Decorby K, McRae L, Tirilis D, Husson H. An optimal search filter for retrieving systematic reviews and meta-analyses. BMC Medical Research Methodology. 2012;12:51.

86.   Leeflang MM, Scholten RJ, Rutjes AW, Reitsma JB, Bossuyt PM. Use of methodological search filters to identify diagnostic accuracy studies can lead to the omission of relevant studies. Journal of Clinical Epidemiology. 2006;59(3):234-40.

87.   Levay P, Ainsworth N, Kettle R, Morgan A. Identifying evidence for public health guidance: a comparison of citation searching with Web of Science and Google Scholar. Res Synth Methods. 2016;7(1):34-45.

88.   Lin J. Is searching full text more effective than searching abstracts? BMC bioinformatics. 2009;10:46.

89.   Lunny C, McKenzie JE, McDonald S. Retrieval of overviews of systematic reviews in MEDLINE was improved by the development of an objectively derived and validated search strategy. Journal of Clinical Epidemiology. 2016;74:107-18.

90.   Martinez Garcia L, Sanabria AJ, Araya I, Lawson J, Sola I, Vernooij RW, et al. Efficiency of pragmatic search strategies to update clinical guidelines recommendations. BMC Medical Research Methodology. 2015;15:57.

91.   Matthews EJ, Edwards AG, Barker J, Bloor M, Covey J, Hood K, et al. Efficient literature searching in diffuse topics: lessons from a systematic review of research on communicating risk to patients in primary care. Health Libraries Review. 1999;16(2):112-20.

92.   McDonald S, Crumley E, Eisinga A, Villanueva E. Search strategies to identify reports of randomized trials in MEDLINE. Cochrane Database of Systematic Reviews. 2007;(2) (no pagination)(MR000018).

93.   McKeever L, Nguyen V, Peterson SJ, Gomez-Perez S, Braunschweig C. Demystifying the Search Button: A Comprehensive PubMed Search Strategy for Performing an Exhaustive Literature Review. Jpen: Journal of Parenteral & Enteral Nutrition. 2015;39(6):622-35.

94.   McKibbon KA, Haynes RB, Dilks CJ, Ramsden MF, Ryan NC, Baker L, et al. How good are clinical MEDLINE searches? A comparative study of clinical end-user and librarian searches. Computers and biomedical research, an international journal. 1990;23(6):583-93.

95.   Merz RB, Cimino C, Barnett GO, Blewett DR, Gnassi JA, Grundmeier R, et al. A pre-search estimation algorithm for MEDLINE strategies with qualifiers. Proceedings - the Annual Symposium on Computer Applications in Medical Care. 1994:910-4.

96.   Mesgarpour B, Muller M, Herkner H. Development of search strategies to identify reports of off-label drug use in EMBASE. Drug Safety. 2011;34 (10):963.

97.   Mesgarpour B, Muller M, Herkner H. Search strategies to identify reports on "off-label" drug use in EMBASE. BMC Medical Research Methodology. 2012;12:190.

98.   Methley AM, Campbell S, Chew-Graham C, McNally R, Cheraghi-Sohi S. PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. BMC Health Services Research. 2014;14:579.

99.   Mijnhout GS, Hooft L, van Tulder MW, Deville WL, Teule GJ, Hoekstra OS. How to perform a comprehensive search for FDG-PET literature. European Journal of Nuclear Medicine. 2000;27(1):91-7.

100.   Moskovitch R, Martins SB, Behiri E, Weiss A, Shahar Y. A comparative evaluation of full-text, concept-based, and context-sensitive search. Journal of the American Medical Informatics Association. 2007;14(2):164-74.

101.   Murphy SA. Applying methodological search filters to CAB abstracts to identify research for evidence-based veterinary medicine.[Erratum appears in J Med Libr Assoc. 2003 Apr;91(2):275]. Journal of the Medical Library Association. 2002;90(4):406-10.

102.   Murphy SA. Research methodology search filters: are they effective for locating research for evidence-based veterinary medicine in PubMed? Journal of the Medical Library Association. 2003;91(4):484-9.

103.   Ogden K, Thompson JC, Halfpenny NJ, Scott DA. A Conceptual Search Filter to Identify Real-World Evidence. Value in Health. 2015;18(7):A728.

104.   Parkhill AF, Clavisi O, Pattuwage L, Chau M, Turner T, Bragge P, et al. Searches for evidence mapping: effective, shorter, cheaper. Journal of the Medical Library Association. 2011;99(2):157-60.

105.   Patel MR, Schardt CM, Sanders LL, Keitz SA. Randomized trial for answers to clinical questions: evaluating a pre-appraised versus a MEDLINE search protocol. Journal of the Medical Library Association. 2006;94(4):382-7.

106.   Petersen H, Poon J, Poon SK, Loy C. Increased workload for systematic review literature searches of diagnostic tests compared with treatments: challenges and opportunities. JMIR Medical Informatics. 2014;2(1):e11.

107.   Petrova M, Sutcliffe P, Fulford KW, Dale J. Search terms and a validated brief search filter to retrieve publications on health-related values in Medline: a word frequency analysis study. Journal of the American Medical Informatics Association. 2012;19(3):479-88.

108.   Pienaar E, Grobler L, Busgeeth K, Eisinga A, Siegfried N. Developing a geographic search filter to identify randomised controlled trials in Africa: finding the optimal balance between sensitivity and precision. Health Information & Libraries Journal. 2011;28(3):210-5.

109.   Pillastrini P, Vanti C, Curti S, Mattioli S, Ferrari S, Violante FS, et al. Using PubMed search strings for efficient retrieval of manual therapy research literature. Journal of Manipulative & Physiological Therapeutics. 2015;38(2):159-66.

110.   Pols DH, Bramer WM, Bindels PJ, van de Laar FA, Bohnen AM. Development and Validation of Search Filters to Identify Articles on Family Medicine in Online Medical Databases. Annals of Family Medicine. 2015;13(4):364-6.

111.   Porter D, Wigton RS, Reidelbach MA, Bleich HL, Slack WV. Self-service computerized bibliographic retrieval: a comparison of Colleague and PaperChase, programs that search the MEDLINE data base. Computers & Biomedical Research. 1988;21(5):488-501.

112.   Rana GK, Bradley DR, Hamstra SJ, Ross PT, Schumacher RE, Frohna JG, et al. A validated search assessment tool: assessing practice-based learning and improvement in a residency program. Journal of the Medical Library Association. 2011;99(1):77-81.

113.   Refinetti R. Retrieval performance in online search in thermal physiology. Computers and biomedical research, an international journal. 1990;23(1):32-6.

114.   Rice DB, Kloda LA, Levis B, Qi B, Kingsland E, Thombs BD. Are MEDLINE searches sufficient for systematic reviews and meta-analyses of the

diagnostic accuracy of depression screening tools? A review of meta-analyses. Journal of Psychosomatic Research. 2016;87:7-13.

115.   Ritchie G, Glanville J, Lefebvre C. Do published search filters to identify diagnostic test accuracy studies perform adequately? Health Information & Libraries Journal. 2007;24(3):188-92.

116.   Royle P, Waugh N. 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. Health Technology Assessment (Winchester, England). 2003;7(34):iii, ix-x, 1-51.

117.   Royle P, Waugh N. A simplified search strategy for identifying randomised controlled trials for systematic reviews of health care interventions: a comparison with more exhaustive strategies. BMC Medical Research Methodology. 2005;5:23.

118.   Saez Gomez JM, Aguinaga Ontoso E, Negro Alvarez JM, Guillen-Grima F, Ivancevich JC, Bozzola CM. Efficient bibliographic searches on allergology using PubMed. Allergologia et Immunopathologia. 2007;35(6):264-75.

119.   Saez Gomez JM, Annan JW, Negro Alvarez JM, Guillen-Grima F, Bozzola CM, Ivancevich JC, et al. Efficient bibliographic searches on allergy using ISI databases. Allergologia et Immunopathologia. 2008;36(2):101-9.

120.   Samadzadeh GR, Rigi T, Ganjali AR. Comparison of Four Search Engines and their efficacy With Emphasis on Literature Research in Addiction (Prevention and Treatment). International Journal of High Risk Behaviors & Addiction. 2013;1(4):166-71.

121.   Sampson M, McGowan J. Inquisitio validus Index Medicus: A simple method of validating MEDLINE systematic review searches. Research Synthesis Methods. 2011;2(2):103-9.

122.   Sampson M, Tetzlaff J, Urquhart C. Precision of healthcare systematic review searches in a cross-sectional sample. Research Synthesis Methods. 2011;2(2):119-25.

123.   Sampson M, Zhang L, Morrison A, Barrowman NJ, Clifford TJ, Platt RW, et al. An alternative to the hand searching gold standard: validating

methodological search filters using relative recall. BMC Medical Research Methodology. 2006;6:33.

124.   Sassi F, Archard L, McDaid D. Searching literature databases for health care economic evaluations: how systematic can we afford to be? Medical Care. 2002;40(5):387-94.

125.   Saxton JD, Owen DJ. Developing optimal search strategies for finding information on herbs and other medicinal plants in MEDLINE. Journal of Alternative & Complementary Medicine. 2005;11(4):725-31.

126.   Shariff SZ, Bejaimal SA, Sontrop JM, Iansavichus AV, Haynes RB, Weir MA, et al. Retrieving clinical evidence: a comparison of PubMed and Google Scholar for quick clinical searches. J Med Internet Res. 2013;15(8):e164.

127.   Shaw RL, Booth A, Sutton AJ, Miller T, Smith JA, Young B, et al. Finding qualitative research: an evaluation of search strategies. BMC Medical Research Methodology. 2004;4:5.

128.   Shojania KG, Bero LA. Taking advantage of the explosion of systematic reviews: an efficient MEDLINE search strategy. Effective Clinical Practice. 2001;4(4):157-62.

129.   Simmons M, Singhal A, Lu Z. Text Mining for Precision Medicine: Bringing Structure to EHRs and Biomedical Literature to Understand Genes and Health. Advances in Experimental Medicine & Biology. 2016;939:139-66.

130.   Sjogren P, Halling A. Medline search validity for randomised controlled trials in different areas of dental research. British Dental Journal. 2002;192(2):97-9.

131.   Souney PF, Miller C, Braun L. On-line literature searching: comparison of PaperChase and National Library of Medicine data bases. Hospital formulary. 1987;22(2):189-94.

132.   Spiegel I, Crager J. Comparison of SUNY and MEDLINE Searches. Bulletin of the Medical Library Association. 1973;61(2):205-9.

133.   Steeves S, Buchanan-Hughes AM, Evans J, Leonard S. Hunting For Randomised Controlled Trials (Rcts): A Comparison Of Search Filters Designed To Identify Rcts. Value in Health. 2015;18(7):A692.

134.   Stelfox HT, Foster G, Niven D, Kirkpatrick AW, Goldsmith CH. Capture-mark-recapture to estimate the number of missed articles for systematic reviews in surgery. American Journal of Surgery. 2013;206(3):439-40.

135.   Topfer LA, Parada A, Menon D, Noorani H, Perras C. Comparison of literature searches on quality and costs for health technology assessment using the MEDLINE and EMBASE databases. International Journal of Technology Assessment in Health Care. 1999;15(2):297-303.

136.   Van Walraven C, Bennett C, Forster AJ. Derivation and validation of a MEDLINE search strategy for research studies that use administrative data. Health Services Research. 2010;45(6 Pt 1):1836-45.

137.   Varela-Lema L, Punal-Rioboo J, Accion BC, Ruano-Ravina A, Garcia ML. Making processes reliable: a validated pubmed search strategy for identifying new or emerging technologies. International Journal of Technology Assessment in Health Care. 2012;28(4):452-9.

138.   Waffenschmidt S, Guddat C. Searches for randomized controlled trials of drugs in MEDLINE and EMBASE using only generic drug names compared with searches applied in current practice in systematic reviews. Research Synthesis Methods. 2015;6(2):188-94.

139.   Waffenschmidt S, Janzen T, Hausner E, Kaiser T. Simple search techniques in PubMed are potentially suitable for evaluating the completeness of systematic reviews. Journal of Clinical Epidemiology. 2013;66(6):660-5.

140.   Wanke LA, Hewison NS. Comparative usefulness of MEDLINE searches performed by a drug information pharmacist and by medical librarians. American Journal of Hospital Pharmacy. 1988;45(12):2507-10.

141.   Wessels M, Hielkema L, van der Weijden T. How to identify existing literature on patients' knowledge, views, and values: the development of a validated search filter. Journal of the Medical Library Association. 2016;104(4):320-4.

43

142.   Whipple EC, McGowan JJ, Dixon BE, Zafar A. The selection of high-impact health informatics literature: a comparison of results between the content expert and the expert searcher. Journal of the Medical Library Association. 2009;97(3):212-8.

143.   Whiting P, Westwood M, Burke M, Sterne J, Glanville J. Systematic reviews of test accuracy should search a range of databases to identify primary studies. Journal of Clinical Epidemiology. 2008;61(4):357-64.

144.   Wieland S, Dickersin K. Selective exposure reporting and Medline indexing limited the search sensitivity for observational studies of the adverse effects of oral contraceptives. Journal of Clinical Epidemiology. 2005;58(6):560-7.

145.   Wong SS, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. Journal of the Medical Library Association. 2006;94(4):451-5.

146.   Wong SSL, Wilczynski NL, Haynes RB. Optimal CINAHL search strategies for identifying therapy studies and review articles: Health policy and systems. Journal of Nursing Scholarship. 2006;38(2):194-9.

147.   Xue J, Chen W, Chen L, Gaudet L, Moher D, Walker M, et al. Significant discrepancies were found in pooled estimates of searching with Chinese indexes versus searching with English indexes. Journal of Clinical Epidemiology. 2016;70:246-53.

148.   Yaylali IE, Alacam T. Critical Assessment of Search Strategies in Systematic Reviews in Endodontics. Journal of Endodontics. 2016;42(6):854-60.

149.   Yousefi-Nooraie R, Irani S, Mortaz-Hedjri S, Shakiba B. Comparison of the efficacy of three PubMed search filters in finding randomized controlled trials to answer clinical questions. Journal of Evaluation in Clinical Practice. 2013;19(5):723-6.

150.   Zhang L, Ajiferuke I, Sampson M. Optimizing search strategies to identify randomized controlled trials in MEDLINE. BMC Medical Research Methodology. 2006;6:23.

Table 1 Further details on included studies

		Reference Standard		Index	Metric or Method							Sum (s)/ form (f)	Threshold or cut-off defined/ CI reported
		Hand searching	Other		Sensitivity	Specificity	Precision	Accuracy	NNR	Yield	Other		
1	Adams 1994	X[RV]		Adequacy of identifying RCTs in MEDLINE[R]	X		X					s	n/y
2	Astin 2008	X[R]		Sensitive Search Strategy in MEDLINE to Retrieve Studies on Assessment of the Diagnostic Performance of Imaging Techniques[R]	X	X	X					s	n/y
3	Bachmann 2002	X[R]		Search terms to identify diagnostic studies in MEDLINE[V]	X		X		X			s	n/y
4	Blanc 2015		PubMed searches	Full-text searches of journals (web-searching)[R]						X[1]		s	n/y
5	Cathey 2006	X[R]		Comparison of handsearching and searching EMBASE	X							s	n/n
6	Dumbridge 2000	X[R]		Assessment of MEDLINE Search Strategies for Randomized Controlled Trials in Prosthodontics[R]	X	X	X					s	n/n
7	Geersing 2012	X[R]		Search Filters for Finding Prognostic and Diagnostic Prediction Studies in Medline[R]	X	X			X			s	n/y
8	Gehanno 2009	X[R]	Validation set made up of: PubMed searches and handsearching	Search Strategies for Identifying Studies on Return-To-Work in Medline[R]	X		X		X			s	y/n
9	Glanville 2012	X[RV]		Handsearching versus electronic searching in a review of diagnostic test accuracy[R]						X	Time	s	n/y

		Reference Standard		Index	Metric or Method							Sum(s)/ form (f)	Threshold or cut-off defined/ CI reported
		Hand searching	Other		Sensitivity	Specificity	Precision	Accuracy	NNR	Yield	Other		
10	Glanville 2014		Searching CDSR/ MEDLINE	Optimal search approaches for searching trials registries[R]	X		X			X		s	n/n
11	Haynes 1994	X[RV]		Development of an optimal search strategy for identification of clinically sound studies in MEDLINE[R]	X	X	X	X				s	n/y
12	Haynes 2004	X[RV]		Optimal search strategies for retrieving scientifically strong studies of diagnosis from Medline: analytical survey[R]	X	X	X	X				s	y/y
13	Haynes 2005[a]	X[RV]		Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey[R]	X	X	X	X				s	n^1/y
14	Haynes 2005[b]	X[RV]		Development of optimal search strategies for detecting clinically sound and relevant causation studies in EMBASE[R]	X	X	X	X				s	y/y
15	Hilderbrand 2014	X[V]		Search filters for PubMed, Medline and Embase[R]	X	X	X	X				s	n/n
16	Holland 2005	X[V]		Clinical prediction studies in EMBASE	X	X	X	X				S	n/n
17	Hopewell 2002	X		MEDLINE search for RCTs[R]					X			s	n/n
18	Ingui 2001	X[R]	Other reviews	Searching for Clinical Prediction Rules in MEDLINE[R]	X	X	x					s	n/y
19	Jenuwine 2004	X[V]		PubMed searches[R]	X	X						s	n/n
20	Kassai 2006	X		Searching for DTA studies	X						Capture-recapture	s and f	n/y

		Reference Standard		Index	Metric or Method							Sum (s)/ form (f)	Threshold or cut-off defined/ CI reported
		Hand searching	Other		Sensitivity	Specificity	Precision	Accuracy	NNR	Yield	Other		
21	Layton 1988	X[R]		MEDLINE search strategy	X		X		X			s	y/n
22	Linder 2015		Database searching[R]	Citation chasing is more sensitive than keyword searches[R]	X		X					s	n/n
23	Marson 1996	X[R]		Searching MEDLINE for RCTs[R]	X		X					s	n/n
24	Mattoli 2012	X[R]		Identify studies in PubMed					X			s	n/n
25	McKibbon 2006	X[R]		PsycINFO search strategy[y]	X	X	X	X				s	y[2]/y
26	McKibbon 2009	X[R]		38 RCT search filters[R]	X	X	X	X				s	n/n
27	McKinlay 2006	X[v]		Cost and economic studies in EMBASE[R]	X	X	X	X				s	y[2]/n
28	Montori 2004	X		Identifying SRs in MEDLINE[VR]	X	X	X					s	n/n
29	Nasser 2006	X[R]		Identification of RCT	X							s	n/n
30	Rogerson 2015	X[R]		DTA search filters[R]	X	X	X		X			s	n/y
31	Spoor 1996	X		MEDLINE search for RCTs							Capture recapture	f	n/y
32	Taljaard 2010	X[RV]		Electronic search strategies to identify reports of cluster randomized trials in MEDLINE[R]	X	X	X					s	n/n[2]
33	Ugolini 2010	X[RV]		Search filters in MEDLINE[R]	X	X	X	X				s	n/n
34	van de Glind 2012	X[R]		MEDLINE search filters[R]	X	X	X	X	X			s	y[1]/n

48

		Reference Standard		Index	Metric or Method							Sum(s)/ form (f)	Threshold or cut-off defined/ CI reported
		Hand searching	Other		Sensitivity	Specificity	Precision	Accuracy	NNR	Yield	Other		
35	Walters 2005	X[RV]		Identification of qualitative studies in EMBASE	X	X	X	X				s	n[1]/y
36	Watson 1999	X[R]		Evaluation of MEDLINE and PsycINFO	X		X					s	n/n
37	Wilczynski 1993	X[RV]		Assessment of methodological search filters in MEDLINE[R]	X	X	X					s	n/n
38	Wilczynski 1994	X[RV]		Test of pre-explosions and subheadings in MEDLINE[R]	X	X	X					s	n/n
39	Wilczynski 2003	X[RV]		Detection of causation strategies in MEDLINE[R]	X	X	X	X				s	n/y
40	Wilczynski 2004	X[RV]		Detection of prognostic studies in MEDLINE[R]	X	X	X	X				s	y/y
41	Wilczynski-2004	X[RV]		Optimal search strategies for detecting health services research studies in MEDLINE	X	X	X					s	y/n
42	Wilczynski 2005	X[R]		Optimal Search Strategies for Detecting Clinically Sound Prognostic Studies in EMBASE: An Analytic Survey[R]	X	X	X	X				s	y/y
43	Wilczynski 2006	X[RV]		Search strategies to identify mental health content in MEDLINE[R]	X	X	X	X				s	y/y
44	Wilczynski 2007	X[RV]		Identification of systematic reviews in EMBASE[R]	X	X	X	X				s	y/y
45	Wilczynski 2010	X[RV]		Quality improvement search filters in MEDLINE[R]	X	X	X	X				s	y/y
46	Wilczynski 2011	X[RV]		Effect of NOT'ing content from search strategies	X	X	X		X			s	y/y

		Reference Standard		Index	Metric or Method							Sum (s)/ form (f)	Threshold or cut-off defined/ CI reported
		Hand searching	Other		Sensitivity	Specificity	Precision	Accuracy	NNR	Yield	Other		
47	Wong 2003	X[RV]		Identifying clinical prediction studies in MEDLINE	X	X	X	X				s	y/n
48	Wong 2004	X[RV]		Identifying clinical relevant qualitative studies in MEDLINE	X	X	X	X				s	n/y
49	Wong 2006	X[RV]		Identifying therapy studies in CINAHL[R]	X	X	X	X				s	y/y
50	Wong 2006	X[RV]		Identifying clinically sound studies in EMBASE	X	X	X	X				s	y/y
	TOTAL				45	34	40	22	8	4	3	s=49; f=2	17/26

Annex 4. Supplementary material to Publication Three, reported in Chapter 8

Hard-to-reach population search filter

1   poverty/

2   poverty area/

3   Vulnerable populations/

4   social support/

5   social conditions/

6   Social Stigma/

7   shame/

8   Social Isolation/

9   quality of life/

10   Prejudice/

11   Socioeconomic Factors/

12   ((hard$ adj2 reach) or (hard$ adj2 locate) or (hard$ adj2 find) or (hard$ adj2 treat) or (difficult$ adj2 locate) or (difficult$ adj2 engage) or social$ exclu$ or social inequalit$ or (difficult$ adj2 reach) or (difficult$ adj2 find) or (difficult$ adj2 treat)).ti,ab.

13   ((geograph$ or transport$ or physical) and barrier$).ti,ab.

14   ((low$ or poor or negative) and (quality adj2 life)).ti,ab.

15   ((vulnerable or disadvantaged or at risk or high risk or low socioeconomic status or neglect$ or affected or marginal$ or forgotten or non-associative or nonassociative or unengaged or hidden or excluded or transient or inaccessible or underserved or stigma$ or inequitable) and (people or population$ or communit$ or neighbourhood$1 or neighborhood$1 or group$ or area$1 or demograph$ or patient$ or social$)).ti,ab.

16   (poverty or deprivation or financial hardship$).ti,ab.

17   ((social or welfare or benefits) and claimant*).ti,ab.

18   ((low-income$ or low income$ or low pay or low$ paid or poor or deprived or debt$ or arrear$) and (people or person$1 or population$1 or communit$ or group$ or neighbourhood$1 or neighborhood$1 or famil$)).ti,ab.

19   (low$ and social class$).ti,ab.

20   (refuser$1 or nonuser$1 or non-user$1 or non user$1 or discriminat$ or shame or prejud$ or racism or racial discriminat$).ti,ab.

21   prisoners/

22   prisoner$1.ti,ab.

23   (recent$ adj2 release$ adj2 (inmate$ or prison$ or detainee$ or felon$ or offender$ or convict$ or custod$ or detention or incarcerat$ or correctional or jail$ or penitentiar$)).ti,ab.

24   ((prison$ or penal or penitentiar$ or correctional facilit$ or jail$ or detention centre$ or detention center$) and (guard$1 or population or inmate$ or system$ or remand or detainee$ or felon$ or offender$1 or convict$ or abscond$)).ti,ab.

25   (parole or probation).ti,ab.

26   ((custodial adj2 (care or sentence)) or (incarceration or incarcerated or imprisonment)).ti,ab.

27   Homebound Persons/

28   (immobile or (disabled and (house bound or home bound)) or ((house or home) adj3 bound)).ti,ab.

29   housing/st

30   (hous$ and (quality or damp$ or standard$ or afford$ or condition$ or dilapidat$)).ti,ab.

31   ((emergency or temporary or inadequate or poor$ or overcrowd$ or over-crowd$ or over-subscrib$ or oversubscrib$) and (hous$ or accommodation or shelter$ or hostel$ or dwelling$)).ti,ab.

32   exp homeless persons/

33   (rough sleep$ or runaway$1 or ((homeless$ or street or destitut$) and (population or person$1 or people or group$ or individual$1 or shelter$ or hostel$ or accommodation$1))).ti,ab.

34   Substance-Related Disorders/

35   Drug users/

36   Substance Abuse, Intravenous/

37   ((drug$ or substance) and (illegal or misus$ or abuse or intravenous or IV or problem use$ or illicit use$ or addict$ or dependen$ or dependant or delinquency)).ti,ab.

38   exp Alcohol-Related Disorders/

39   Alcoholics/

40   ((alcohol$ and (misus$ or abuse or problem$ use$ or problem drink$ or illicit use$ or addict$ or dependen$ or dependant or delinquency)) or (alcoholic$1 or alcoholism)).ti,ab.

41   Prostitution/

42   (prostitut$ or sex work$ or transactional sex$).ti,ab.

43   gypsies/

44   (traveller$1 or Gypsies or Gypsy or Gipsy or Gipsies or Romany or Romanies or Romani or Romanis or Rromani or Rromanis or Roma).ti,ab.

45   mental health/

46   Mentally Ill Persons/

47   (mental$ and (health or ill or illness)).ti,ab.

48   Communication Barriers/

49   Language/

50   (((language$ or communicat$) and (barrier$ or understand$ or strateg$ or proficien$)) or translat$ or interpret$ or (cultur$ and competen$)).ti,ab.

51   (illiteracy or illiterate$).ti,ab.

52   "Emigration and Immigration"/

53   "Emigrants and Immigrants"/

54   "Transients and Migrants"/

55   refugees/

56   (immigrant$ or migrant$ or asylum or refugee$ or undocumented).ti,ab.

57   (displaced and (people or person$1)).ti,ab.

58   (born adj2 overseas).ti,ab.

59   ("foreign born" or "non uk born" or "non-uk born").ti,ab.

60   (complex adj2 (patient$ or need$)).ti,ab.

61   Community health aides/

62   Nurses/

63   Nurse's Practice Patterns/

64   Family practice/

65   Physicians, Family/

66   Voluntary Workers/

67   exp Voluntary health agencies/

68   exp Allied Health Personnel/

69   professional-family relations/

70   professional-patient relations/

71   nurse-patient relations/

72   physician-patient relations/

73   ((outreach or support or case or social or lay or allied or link or social care or socialcare
or healthcare or health care) adj2 (worker$ or professional$ or practitioner or advocate$1 or
personnel or staff or service provi$)).ti,ab.

74   (nurse or ((general or family) adj3 (practice$ or practitioner$ or physician$ or
doctor$))).ti,ab.

75   or/1-74

Hard-to-reach population search filter applied to search strategy to retrieve qualitative evidence on tuberculosis

Database(s): Ovid MEDLINE(R) 1948 to December Week 4 2010

1   exp Tuberculosis/
2   (tuberculosis or tb).ti,ab.
3   1 or 2
4   qualitative research/
5   interview/
6   Questionnaires/
7   Focus Groups/
8   Interviews as Topic/
9   Health Care Surveys/
10   Nursing Methodology Research/
11   (qualitative or focus group$ or case stud$ or field stud$ or interview$ or questionnaire$ or survey$ or ethnograph$ or grounded theory or action research or participant observation or narrative$).ti,ab.
12   (life and (history or stor$)).ti,ab.
13   (verbal interaction$ or discourse analys?s or narrative analys?s or social construct$ or purposive sampl$ or phenomenol$ or criterion sampl$).ti,ab.
14   exp Attitude/
15   Motivation/
16   (view$ or barrier$ or block$ or obstacle$ or hinder$ or constrain$ or facilitat$ or attitude$ or opinion$ or belief$ or perceiv$ or perception$ or aware$ or personal view$ or motivat$ or reason$ or incentiv$).ti,ab.
17   or/4-16
18   3 and 17
19   animals/ not humans/
20   18 not 19
21   limit 20 to yr="1990 -Current"
22   poverty/
23   poverty area/
24   Vulnerable populations/
25   social support/
26   social conditions/
27   Social Stigma/
28   shame/
29   Social Isolation/
30   quality of life/
31   Prejudice/
32   Socioeconomic Factors/

33   ((hard$ adj2 reach) or (hard$ adj2 locate) or (hard$ adj2 find) or (hard$ adj2 treat) or (difficult$ adj2 locate) or (difficult$ adj2 engage) or social$ exclu$ or social inequalit$ or (difficult$ adj2 reach) or (difficult$ adj2 find) or (difficult$ adj2 treat)).ti,ab.

34   ((geograph$ or transport$ or physical) and barrier$).ti,ab.

35   ((low$ or poor or negative) and (quality adj2 life)).ti,ab.

36   ((vulnerable or disadvantaged or at risk or high risk or low socioeconomic status or neglect$ or affected or marginal$ or forgotten or non-associative or nonassociative or unengaged or hidden or excluded or transient or inaccessible or underserved or stigma$ or inequitable) and (people or population$ or communit$ or neighbourhood$1 or neighborhood$1 or group$ or area$1 or demograph$ or patient$ or social$)).ti,ab.

37   (poverty or deprivation or financial hardship$).ti,ab.

38   ((social or welfare or benefits) and claimant*).ti,ab.

39   ((low-income$ or low income$ or low pay or low$ paid or poor or deprived or debt$ or arrear$) and (people or person$1 or population$1 or communit$ or group$ or neighbourhood$1 or neighborhood$1 or famil$)).ti,ab.

40   (low$ and social class$).ti,ab.

41   (refuser$1 or nonuser$1 or non-user$1 or non user$1 or discriminat$ or shame or prejud$ or racism or racial discriminat$).ti,ab.

42   prisoners/

43   prisoner$1.ti,ab.

44   (recent$ adj2 release$ adj2 (inmate$ or prison$ or detainee$ or felon$ or offender$ or convict$ or custod$ or detention or incarcerat$ or correctional or jail$ or penitentiar$)).ti,ab.

45   ((prison$ or penal or penitentiar$ or correctional facilit$ or jail$ or detention centre$ or detention center$) and (guard$1 or population or inmate$ or system$ or remand or detainee$ or felon$ or offender$1 or convict$ or abscond$)).ti,ab.

46   (parole or probation).ti,ab.

47   ((custodial adj2 (care or sentence)) or (incarceration or incarcerated or imprisonment)).ti,ab.

48   Homebound Persons/

49   (immobile or (disabled and (house bound or home bound)) or ((house or home) adj3 bound)).ti,ab.

50   housing/st

51   (hous$ and (quality or damp$ or standard$ or afford$ or condition$ or dilapidat$)).ti,ab.

52   ((emergency or temporary or inadequate or poor$ or overcrowd$ or over-crowd$ or over-subscrib$ or oversubscrib$) and (hous$ or accommodation or shelter$ or hostel$ or dwelling$)).ti,ab.

53   exp homeless persons/

54   (rough sleep$ or runaway$1 or ((homeless$ or street or destitut$) and (population or person$1 or people or group$ or individual$1 or shelter$ or hostel$ or accommodation$1))).ti,ab.

55   Substance-Related Disorders/

56   Drug users/

57   Substance Abuse, Intravenous/

58   ((drug$ or substance) and (illegal or misus$ or abuse or intravenous or IV or problem use$ or illicit use$ or addict$ or dependen$ or dependant or delinquency)).ti,ab.

59   exp Alcohol-Related Disorders/

60   Alcoholics/

61   ((alcohol$ and (misus$ or abuse or problem$ use$ or problem drink$ or illicit use$ or addict$ or dependen$ or dependant or delinquency)) or (alcoholic$1 or alcoholism)).ti,ab.

62   Prostitution/

63   (prostitut$ or sex work$ or transactional sex$).ti,ab.

64   gypsies/

65   (traveller$1 or Gypsies or Gypsy or Gipsy or Gipsies or Romany or Romanies or Romani or Romanis or Rromani or Rromanis or Roma).ti,ab.

66   mental health/

67   Mentally Ill Persons/

68   (mental$ and (health or ill or illness)).ti,ab.

69   Communication Barriers/

70   Language/

71   (((language$ or communicat$) and (barrier$ or understand$ or strateg$ or proficien$)) or translat$ or interpret$ or (cultur$ and competen$)).ti,ab.

72   (illiteracy or illiterate$).ti,ab.

73   "Emigration and Immigration"/

74   "Emigrants and Immigrants"/

75   "Transients and Migrants"/

76   refugees/

77   (immigrant$ or migrant$ or asylum or refugee$ or undocumented).ti,ab.

78   (displaced and (people or person$1)).ti,ab.

79   (born adj2 overseas).ti,ab.

80   ("foreign born" or "non uk born" or "non-uk born").ti,ab.

81   (complex adj2 (patient$ or need$)).ti,ab.

82   Community health aides/

83   Nurses/

84   Nurse's Practice Patterns/

85   Family practice/

86   Physicians, Family/

87   Voluntary Workers/

88   exp Voluntary health agencies/

89   exp Allied Health Personnel/

90   professional-family relations/

91   professional-patient relations/

92   nurse-patient relations/

93   physician-patient relations/

94   ((outreach or support or case or social or lay or allied or link or social care or socialcare or healthcare or health care) adj2 (worker$ or professional$ or practitioner or advocate$1 or personnel or staff or service provi$)).ti,ab.

95   (nurse or ((general or family) adj3 (practice$ or practitioner$ or physician$ or doctor$))).ti,ab.

96   or/22-95

97   21 and 96

Screening checklist for the retrospective analysis of the case study

1.	Was the study published in English?	YES/UNCLEAR – go to Q2	NO – exclude
2.	Was the study published in 1990 or later?	YES/UNCLEAR – go to Q3	NO – exclude
3.	Was the study conducted in an OECD country? (see below)	YES/UNCLEAR – go to Q4	NO – exclude
4.	Does the study present qualitative data? Does the study present qualitative data?	YES/UNCLEAR – go to Q5 YES/UNCLEAR – go to Q7 NO – exclude	NO – exclude
5.	Does the study have a focus on TB services of any kind?	YES/UNCLEAR – go to Q6	NO – exclude
6.	Does the study include data from any hard-to-reach group?	YES/UNCLEAR – go to Q7	NO – exclude
7.	Does the study include data from any hard-to-reach group that has not already been included in the search filter?	YES/UNCLEAR – include	NO – exclude
8.	Is there insufficient information in the title, abstract and keywords to make a judgement?	YES/UNCLEAR – temporary include and request full text then return to questions above or go to question 9	NO – exclude
9	No full text available	Discuss with second reviewer	

Note: at the time of reviewing the following were OECD member countries:

Australia, Austria, Belgium, Canada, Chile, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States

Annex 5.   Supplementary material to Publication Four, reported in Chapter 9

A5.1.   Searching

Databases searched:

1. ASSIA via Pro Quest;
2. BIOSIS via ISI;
3. British Education Index (BEI) via Pro Quest;
4. British Nursing Index (BNI) via Pro Quest;
5. Campbell Library via http://www.campbellcollaboration.org/Library/Library.php
6. Database of promoting health effectiveness reviews (DoPHER) via http://eppi.ioe.ac.uk/webdatabases/Intro.aspx?ID=2
7. EMBASE via OVID;
8. ERIC via Pro Quest;
9. Global Health via OVID;
10. GreenFILE via EBSCOHost;
11. HMIC (Health Management Information Consortium) via OVID;
12. IBSS via Pro Quest;
13. MEDLINE via OVID;
14. OpenGrey via http://www.opengrey.eu/
15. PsycINFO via OVID;
16. Social Policy and Practice (SPP) via OVID;
17. Social Services Abstracts via Pro Quest;
18. Sociological Abstracts via Pro Quest;
19. SPORTDiscuss via EBSCOHost;
20. The Cochrane Library (CDSR, DARE, HTA Library, CENTRAL and NHS EEDs) via Wiley interface;
21. The Trials Register of Promoting Health Interventions (TRoPHI) via http://eppi.ioe.ac.uk/webdatabases/Intro.aspx?ID=5
22. Web of Science (SCI-EXPANDED; SSCI; A&HCI; CPCI-S; CPCI-SSH)

MEDLINE search strategy:

Database(s): Ovid MEDLINE(R)
Host: OVID
Data Parameters: 1946 to September Week 3 2012
Date Searched: Wednesday October 3rd 2012
Searched By: CC
Strategy Checked by: KH, RL and RG
Search Strategy

# Searches	Results

1	(conservation$ and natural and environment$ and (renewal or volunteer$ or voluntary or participat$ or practical or regenerat$ or restor$ or maintain$ or care or enhance$ or preserve or creat$ or activ$ or action$ or involve$)).ti,ab.	377
2	(Conservation adj3 interventions).ti,ab.	29
3	((environmental$ adj3 (conservation$ or volunteer$ or steward$)) and (Regenerat$ or restore or restoration or redevelop or maintain or enhance or preserve or preserving or create or creation or establish or establishing or founding or build$ or cultivat$ or cultivation or participate or participation)).ti,ab.	73
4	(conservation$ adj3 (group$ or volunteer$ or voluntary or association$ or organisation$ or organization$ or participa$ or stakeholder$ or steward$ or trust or ranger$ or activit$)).ti,ab.	747
5	(conservation$ adj5 (nature or rural or countryside or outdoor$ or outside or backcountry or hinterland or outback or wood$ or park$1 or parkland or garden$ or meadow$ or farm$ or (farm adj1 land) or horticultural or floricultural or botanical or arboretum or allotment$ or forest$ or rainforest or moor$ or dale$1 or marsh$ or mountain$ or beach$ or wilderness or landscape$ or tree$ or copse$ or river$ or lake$ or canal$ or waterway or wetland$ or (open adj1 space$) or (protected adj1 area$) or green$ or planning$ or footpath$ or trail$ or coast$ or cliff$ or dune$ or (bio adj1 diversity) or (eco adj1 system) or (protected adj1 area$))).ti,ab.	1688
6	(geoconservation or (geo adj3 conservation)).ti,ab.	0
7	((activ$ or practical or participat$) adj3 conservation$).ti,ab.	481
8	exp "Conservation of Natural Resources"/ or *Environment/ or *Environment Design/	42248

60

9	(volunteer$ or voluntary).ti,ab. or *Voluntary Workers/ or *Consumer Participation/ or *Health Status/	199928
10	8 and 9	638
11	1 or 2 or 3 or 4 or 5 or 6 or 7 or 10	3520
12	((Volunteer$ or voluntary) adj5 (environment$ or nature or rural or countryside or outdoor$ or outside or backcountry or hinterland or outback or wood$ or park$1 or parkland or garden$ or meadow$ or horticultural or floricultural or botanical or arboretum or allotment$ or forest$ or rainforest or moor$ or dale$1 or marsh$ or mountain$ or beach$ or wilderness or landscape$ or tree$ or copse$ or river$ or lake$ or canal$ or waterway or wetland$ or (open adj1 space$) or (protected adj1 area$) or green$ or planning$ or footpath$ or trail$ or coast$ or cliff$ or dune$ or (bio adj1 diversity) or (eco adj1 system) or (protected adj1 area$))).ti,ab.	1142
13	(((voluntary or volunteer$) adj5 (group$ or association or stakeholder$ or steward$ or ranger$)) and (environment$ or nature or rural or countryside or outdoor$ or outside or backcountry or hinterland or outback or wood$ or park$1 or parkland or garden$ or meadow$ or horticultural or floricultural or botanical or arboretum or allotment$ or forest$ or rainforest or moor$ or dale$1 or marsh$ or mountain$ or beach$ or wilderness or landscape$ or tree$ or copse$ or river$ or lake$ or canal$ or waterway or wetland$ or (open adj1 space$) or (protected adj1 area$) or green$ or planning$ or footpath$ or trail$ or coast$ or cliff$ or dune$ or (bio adj1 diversity) or (eco adj1 system) or (protected adj1 area$))).ti,ab.	667
14	*Voluntary Workers/	3989
15	(environment$ or nature or rural or countryside or outdoor$ or outside or backcountry or hinterland or outback or wood$ or park$1 or parkland or garden$ or meadow$ or horticultural or	1253503

floricultural or botanical or arboretum or allotment$ or forest$ or rainforest or moor$ or dale$1 or marsh$ or mountain$ or beach$ or wilderness or landscape$ or tree$ or copse$ or river$ or lake$ or canal$ or waterway or wetland$ or (open adj1 space$) or (protected adj1 area$) or green$ or planning$ or footpath$ or trail$ or coast$ or cliff$ or dune$ or (bio adj1 diversity) or (eco adj1 system) or (protected adj1 area$)).ti,ab.

16	14 and 15	356
17	12 or 13 or 16	2010
18	(Green$ adj3 (space$ or gym or exercise or volunteer$ or voluntary or conservation or infrastructure or care or streets or communal or Guerrilla)).ti,ab.	402
19	greenspace.ti,ab.	25
20	18 or 19	425
21	(urban adj3 (green$ or park$1 or parkland or garden$ or horticultur$ or wood$ or forest$ or botanical or arboretum or allotment$ or (open adj1 space))).ti,ab.	579
22	((work$ or renewal or volunteer$ or voluntary or practical or regenerat$ or restor$ or maintain$ or care or enhance or preserve or creat$) and (urban or city or metropolis or town$) and (garden$ or park$1 or parkland or allotment$)).ti,ab.	370
23	*Cities/ and ((work$ or renewal or volunteer$ or voluntary or practical or regenerat$ or restor$ or maintain$ or care or enhance or preserve or creat$) and (garden$ or park$1 or parkland or allotment$)).ti,ab.	5
24	*Urban Health/ and (*Conservation of Natural Resources/ or *Voluntary Workers/)	19
25	21 or 22 or 23 or 24	914

26	((garden$ or horticulture or allotment$ or botanical or arboretum) adj5 (kitchen or school$ or college$ or university or campus or hospital$ or prison$ or penitentiary or institution or urban or green$ or communit$ or communal or group$ or guerrilla or (bio adj1 diver$) or eco or ((grow or pick) adj3 your own))).ti,ab.	595
27	((garden$ or horticulture or allotment$ or botanical or arboretum) adj5 (maintain$ or creat$ or culivat$ or enhance$ or preserve or voluntary or volunteer or conservation$ or participat$)).ti,ab.	120
28	Gardening/ and (*Conservation of Natural Resources/ or *Voluntary Workers/)	13
29	*Gardening/ and (kitchen or school$ or college$ or university or campus or hospital$ or prison$ or penitentiary or institution or urban or green$ or communit$ or communal or group$ or guerrilla or (bio adj1 diver$) or eco or maintain$ or creat$ or culivat$ or voluntary or volunteer or conservation$ or participat$).ti,ab.	118
30	26 or 27 or 28 or 29	754
31	((communit$ adj5 (group$ or team$ or association$ or organisation or organization or participa$ or stakeholder$ or steward$ or trust$ or ranger$ or activit$)) and (garden$ or allotment$ or forest or (natural and environment) or conservation$)).ti,ab.	363
32	(communit$ and (work$ or renewal or volunteer$ or voluntary or practical or regenerat$ or restor$ or maintain$ or care or enhance$ or preserve or creat$ or activ$ or action$ or involve$) and ((natur$ adj3 environment$) or (environmental$ and conservation$))).ti,ab.	479
33	(((communit$ or local) adj5 (garden$ or park$ or green$ or greenspace or outdoor$ or outside$ or pavement$ or sidewalk$ or wood$ or allotment$ or lake$ or canal$ or river$)) and (work$ or renewal or volunteer$ or voluntary or practical or participat$ or	813

regenerat$ or restor$ or maintain$ or enhance or preserve or
creat$)).ti,ab.

34	31 or 32 or 33	1556

35	11 or 17 or 20 or 25 or 30 or 34	8554

36	exp animals/ not humans.sh.	3785951

37	35 not 36	6941

(clinical or surgery or surgical or cell or cells or laboratory or

38	placebo or bladder or uterus or breast or gene or genes or genetic	9317419

or bowel or liver or enzymes or viral or lymph or molecular).mp.

39	37 not 38	4815

40	limit 39 to english language	4349

41	limit 40 to yr="1990 -Current"	3896

A5.2. Qualitative analysis

Personal achievement

All of the included qualitative studies reported this complex category, and
therefore all three good quality studies. There were two main focuses to the
discussions reported: first, there were those studies (usually those where
mental health issues were a factor) in which respondents were engaging as a
means of recuperation, and second, those in which participation in
environmental improvement was the motivating factor. The types of
achievement valued and experienced by the participants varied between these
two groups.

In the first group, the richest and most pronounced reports of achievement
came from projects dealing with individuals experiencing mental ill health
(BTCV 2010a (n = 19), Wilson 2009 (n = 29) particularly). Achievement came
about through the provision of, and then adherence to the daily structure of the
programme, thereby increasing motivation and ability to engage in activities
and, finally, the impact that completing the activities had:

*"The very fact of participation was an achievement in itself for some volunteers.
Depression and linked illnesses can limit daily activity and so for some to get
dressed and attend was significant."* (Author, BTCV 2010a (n =19)).

Commented [CC1]: s1 achievement in ill health: adherence to daily structure and increase in motivation and engagement

64

"Aye it's been great I've thoroughly enjoyed it, Aye. I wouldn't say I've been great at it. I've tried it anyway; I've came along and tried it. I wasn't too good at it (willow weaving) but at the end I done it. At least I tried … I feel in myself I've achieved something … Like see when I gae home after leaving here I'm puffed oot and I feel as if I've achieved something. I'm knackered and I'm quite proud of myself cause I've done it." (Participant, Wilson 2009 (n = 29) review team ellipsis)

For these people the sense of achievement focused, at least initially, less on the nature of the activity undertaken, rather through attempting and adherence to the programme. It was considered to be a progressive and reinforcing process, with some participants developing the self-confidence and the skills to re-enrol as team leaders after successful completion. This was seen as significant progress and achievement and, perhaps, shows developing commitment to the actual activities involved.

The second group of people for whom achievement was important were those seeking to improve the environment, particularly those engaged with the Australian Landcare movement and other associated programmes (Christie 2004 (n = 18), Gooch 2005 (n = 85), Townsend 2004 (n=18), Townsend 2006 (n=80)). For these participants it was primarily the environmental impact which was important, however some individuals reported that this led to a negative feeling of futility when activities resulted in little or no impact (Christie 2004 (n = 18)).

Related to this, were those who found benefits accrued through taking part in socially or environmentally valuable activities. For some, environmental enhancement and conservation activities provided an opportunity to 'give something back' Christie 2004 (n = 18). This was of particular importance for those who felt they had drawn on societal resources, or who had a strong environmental ethic:
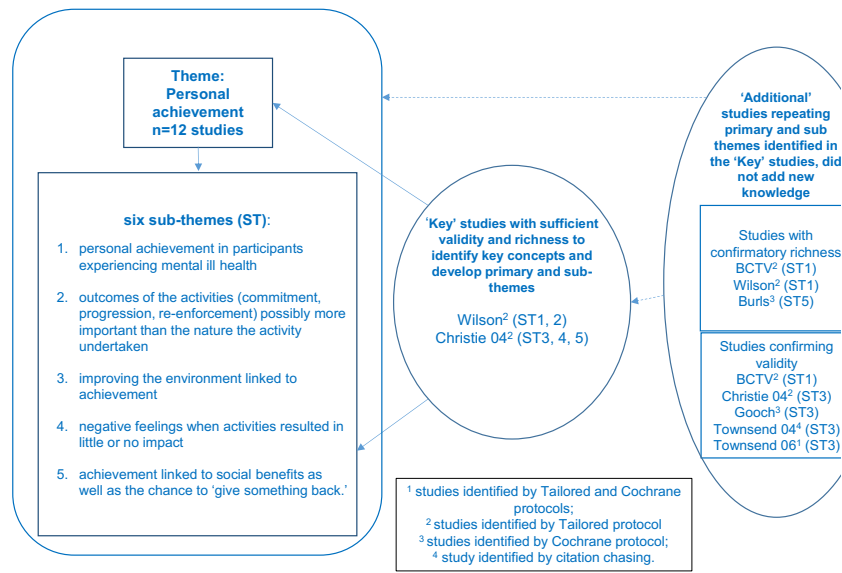
"Our work is beneficial to nature; for the benefit of the birds; we create an environment for wildlife; we've got trees established now, probably some of them are 25 feet tall; it's not just this plot of land, it's not just for these birds and this wildlife but it's for the people as well; for other people to look at in years to come; greater understanding of plants, nature and ecology; regeneration; the birds have somewhere to nest, the frogs have somewhere to spawn, it makes the world go round." (Participant, Burls 2007 (n = 11))

This category seemingly exists on a continuum of personal achievement: with completing structured daily activities (which, for some, amounted to getting out of bed) at one end, and impacting on global environmental troubles at the other, and as the last quote illustrates, these impacts were interconnected with many of the other themes discussed here.

Commented [CC2]: s2 outcome of process possibly more important than work done

Commented [CC3]: s2 feelings of progression, reinforcement, commitment to activities.

Commented [CC4]: s3→ – achievement linked to improvement in the environment

Commented [CC5]: s4 – negative feelings of achievement where activates resulted in to little or no impact

Commented [CC6]: s5 – social benefits of conversation

Commented [CC7]: s7 'payback'

Figure 10 personal and achievement

Theme:
Personal
achievement
n=12 studies

six sub-themes (ST):

1. personal achievement in participants
   experiencing mental ill health

2. outcomes of the activities (commitment,
   progression, re-enforcement) possibly more
   important than the nature the activity
   undertaken

3. improving the environment linked to
   achievement

4. negative feelings when activities resulted in
   little or no impact

5. achievement linked to social benefits as
   well as the chance to 'give something back.'

'Key' studies with sufficient
validity and richness to
identify key concepts and
develop primary and sub-
themes

Wilson[2] (ST1, 2)
Christie 04[2] (ST3, 4, 5)

'Additional'
studies repeating
primary and sub
themes identified in
the 'Key' studies, did
not add new
knowledge

Studies with
confirmatory richness
BCTV[2] (ST1)
Wilson[2] (ST1)
Burls[3] (ST5)

Studies confirming
validity
BCTV[2] (ST1)
Christie 04[2] (ST3)
Gooch[3] (ST3)
Townsend 04[4] (ST3)
Townsend 06[1] (ST3)

[1] studies identified by Tailored and Cochrane
protocols;
[2] studies identified by Tailored protocol
[3] studies identified by Cochrane protocol;
[4] study identified by citation chasing.

Personal/social identity

Six of the included studies (including one good quality study) discussed the impact that participation in EECA had on individuals' sense of personal and social identity, and related to the sense of self-worth, of community, belonging, environmentalism, and a reinforcement of a sense of self as connected to nature (Burls 2007 (n = 11); Carter 2008 (n = unknown); Christie 2004 (n = 18); Gooch 2005 (n = 85); O'Brien 2008a (n = 88, good quality); Wilson 2009 (n = 29).

Carter 2008 (n = unknown), in a study of offenders, examined the impact that environmental work had on participants' integration into society. Participants discussed the notion of re-building a sense of self-worth and identity through engaging in EECA, during which they came into contact with non-offenders, and through which they felt they were making a direct contribution to society. Of particular importance was the sense of being trusted, to be out and talking to the public, which although difficult for some of the individuals, was felt to contribute to the process of de-stigmatisation and development of self-esteem:

"It's nice feeling part of, ehm, part of society again." (Participant, Carter 2008 (n = unknown)).

Similarly, volunteers in an Australian stewardship programme, some of whom had mental health issues, noted the importance of rebuilding a positive social identity through the group based on meaningful and collaborative activities (Burls 2007 (n = 11). Contact with the public was also noted to be important for these individuals.

The results reported by Christie 2004 (n = 18) differed from those of Carter and Burls in that respondents, who were conservation volunteers, focused more on environmental outcomes and their contribution to them. Participants reported that their sense of identity was linked to the impact they were having on environmental issues. Similarly Gooch 2005 (n = 85), who reported on the impacts of a catchment volunteering programme, found that developing and maintaining an environmentalist identity and having an impact on nature was valued:

"The study suggests that the social identity formed by members of a particular group contributes to a sense of belonging, responsibility, values and emotions." (Author, Gooch 2005 (n = 85)).

Comparison of the results from Burls 2007 (n = 11) and Carter 2008 (n = unknown) demonstrates an apparent difference in impact between different user categories: for the marginalised groups, the meaningful activities facilitated the rebuilding or maintaining of a "normal" identity, this was articulated by those who may have felt they had been defined by illness or status (for instance as a 'prisoner' or 'depressed'), while for others the activities allowed the participants to demonstrate and validate their "environmentalist" identities.

There were several other ways in which participation had an impact on identity. A number of the respondents interviewed by O'Brien 2008a (n = 88, good

Commented [CC8]: s1 ire-building identity (offenders): self worth, trust, self estee.

Commented [CC9]: s2 identity linked to impact in the environment (as opposed to re-building identify).

Commented [CC10]: s1 again

Commented [CC11]: this is s2 as above?

quality) highlighted the role of participation in maintaining a positive self-identity post retirement. While for others in the study, particularly those who had struggled to find paid work, volunteering contributed to their sense of self-worth and status.

The role of the activities in enabling a continuation of a sense of self as connected to nature, a notion which had developed in childhood, was also identified by both O'Brien 2008a (n = 88, good quality) and by Wilson 2009 (n = 29):

"Ah, well I've always enjoyed the outdoors. But since I've became not well, it's just as if I've been housed. Just locked up in the house which is not me. So this was a chance to get out, get fresh air, some exercise and do something for the community and that." (Participant, Wilson 2009 (n = 29)).

Through engaging with meaningful activities that were seen to be valuable socially and environmentally, individuals had access to resources (personal, social and cultural) which allowed them to develop more positive self-identities.

Figure 11 Personal social/identity
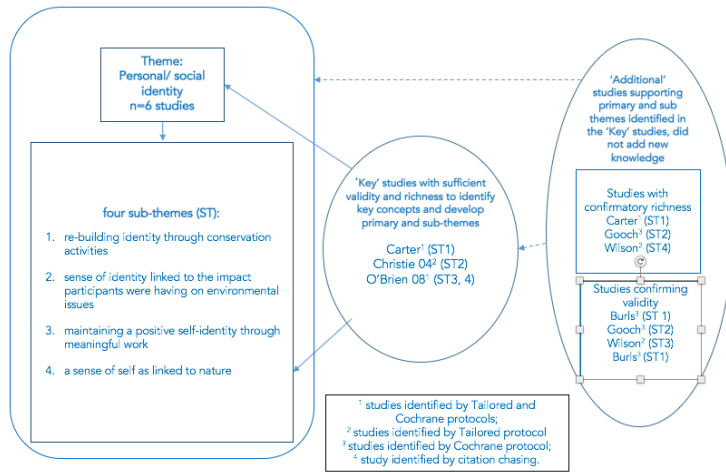
Developing knowledge

Nine of the 12 included qualitative studies (including two good quality studies) reported results relating to participants' perceived increases in knowledge of the environment, not only of more specific associated conservation skills, but also in relation to social and personal abilities.

The immediate impact of participation in EECA on knowledge gain could be found in the development of the skills necessary to carry out the EECA effectively, through knowledge of what to do, how and when to do it. For instance:

"Improved confidence was felt to be linked to enhanced knowledge about how to use tools properly." (Author, BTCV 2010a (n = 19)).

This immediate acquisition of relevant skills improved self-confidence and appeared to contribute to the positive impacts of participation. Participants in studies by Burls 2007 (n = 11), Carter 2008 (n = unknown), Christie 2004 (n = 18), Gooch 2005 (n = 85), O'Brien 2008a (n = 88, good quality), Townsend 2004 (n = 18), and Wilson 2009 (n = 29), reported increases in their knowledge of nature and the environment.

"I get a better understanding of the river system in doing it. I get a better understanding of the whole environment…and it stimulates me." (Participant Gooch 2005 (n = 85) review team ellipsis).

This acquisition of knowledge directly contributed to one participant's enjoyment of the activities:

"I've loved the activities, you know, finding out about the trees and, and you know, the plants and things. I love all that." (Participant, Wilson 2009 (n = 29)).

In some cases the knowledge gained was more widely applicable beyond the EECA programmes. For instance, one of the major themes emerging from the analysis by O'Brien 2010a (n = 10, good quality) was the development of transferable employability skills alongside the more sociable and physical benefits. Burls 2007 (n = 11) also noted that participants, some of whom received unemployment benefits, felt more positive about their employment prospects as a result of taking part in the programmes. Benefits to wider skills such as increased vocabulary and team working were highlighted. Similarly, the participants in the study by BTCV 2010a (n = 19) who also had mental health issues, received practical training in environmental conservation. The participants highlighted the specific nature of the knowledge gained, for example using soil rather than concrete to erect fencing, and how this had led them to question how they undertook other tasks. Participants also received a certificate of proficiency, which was a major achievement, and increased confidence as well as demonstrating their knowledge of conservation techniques:

"Developing new perspectives was also central, and this in turn led to some volunteers studying for qualifications in conservation. A proportion of the volunteers had been unemployed for some time and so the structure and

Commented [CC14]: s1 knowledge gain and developing skills

Commented [CC15]: s1 developing knowledge and improved self confidence

Commented [CC16]: ??developing knowledge linked to positive impact of participation

Commented [CC17]: ?? quotes to support s2

Commented [CC18]: s3 development of skills beyond conservation

69

activity of the sessions was beneficial. Skills learned also contributed to feelings of enhanced employability." (Author, BTCV 2010a (n = 19)).

Learning ranged from specific tool use to broader knowledge of nature and the environment, as well as how to function as part of a team to achieve a goal.

Figure 12 Developing knowledge
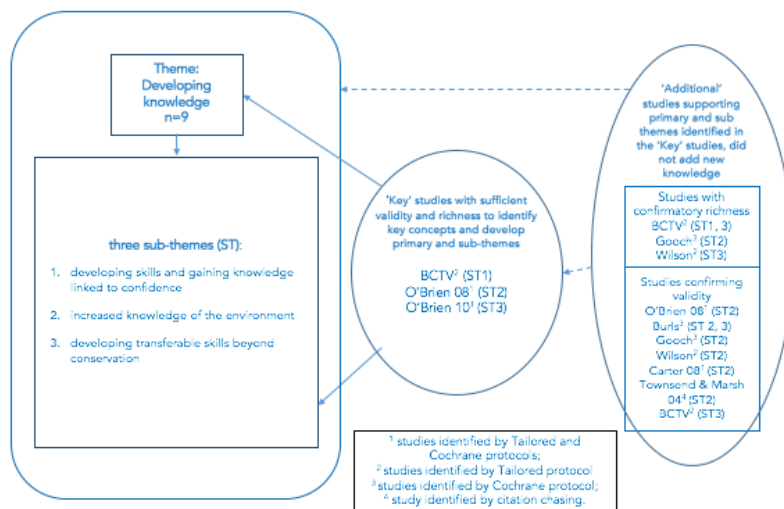
Benefits of place

The benefits of the space and places in which the activities took place was a key theme and was present in all included qualitative studies, including all three good quality studies. There were three specific aspects: 1) being in nature, 2) the development of a sense and pride of place, and 3) being away from usual environments. The impacts of contact with natural environments were particularly important and simply being outdoors was a positive element (O'Brien 2008a (n = 88) good quality). Several of the participants in the study by Burls 2007 (n = 11)) described multiple values of being in nature, including:

"the beauty has a calming effect." (Participant, Burls 2007 (n = 11)).

For others the benefits of being in a natural space related to a perception of a cleaner environment (Townsend 2004 (n = 18)), and for others the variety of natural life was important:

"I don't think there is anything more enjoyable than being out in the fresh air with nature, you never know what you're going to see, what you're going to bump into" (Participant, Burls 2007 (n = 11)).

These opportunities to be in nature were motivational and helped maintain adherence to the projects (Caissie 2003 (n = 10) good quality). There was also the suggestion that contact with the natural world helped give participants a broader perspective of the world and their place within it. Burls 2007 (n = 11) argued that the new and intimate connection with nature allowed the participants to develop the feeling they were part of something fundamental; a cycle of growth, of nature and life:

"Taking care of our environment and feeling that we are part of it; some level of power and energy " (Participant, Burls 2007 (n = 11)).

Some respondents stated that their sincere relationship with nature and the local place in which the activity was undertaken was both a motivator and outcome of participation (Burls 2007 (n = 11)). Regular work in, and care for a particular environment resulted in a strong sense of place and attachment:

"When we pass round that roundabout and see those trees growing it's very rewarding. I can see that I've done my little bit for the environment. I live around here – I intend to come back" (Christie 2004 (n = 18)).

Many had a broad vision for the conservation of the environment and participation in EECA provided a route through which to contribute and something to be proud of (Christie 2004 (n = 18)). A participant in the study undertaken by Gooch 2005 (n = 85) argued that there was a clear need for someone to take a stand and protect the environment of her local community:

"The biggest thing for me when I came here was meeting like-minded people. It feels good to give something that nobody else is prepared to give" (Participant, Gooch 2005 (n = 85)).

Commented [CC20]: st1 – contact with natural environment

Commented [CC21]: s2 benefits of **natural** environment

Commented [CC22]: s3 opportunities of place; motivator, offering perspective

Commented [CC23]: s3 again

Commented [CC24]: s4 sense of place and attachment

Commented [CC25]: s4 ecca as a way to contribute to a place and be proud of contribution

For many, especially those who had experienced various forms of marginalisation, deprivation and, perhaps, institutionalisation, the benefits of place were associated with 'being away':

"[It is] a chance to get people out into a green space, it's very different to all of the environments in mental health services elsewhere, day centres are just not going to have this kind of atmosphere." (Group leader/participant, Burls 2007 n = 11)).

This sentiment was echoed by Wilson 2009 (n = 29) study of the impacts of the Branching Out programme:

"It's been very therapeutic I think - all the different sights and sounds and smells is very different from the hospital environment that I'm used to, you know and the city environment of course, and I've really enjoyed being out in the countryside." (Participant, Wilson 2009 (n = 29))

However, more broadly, and for participants who had not been referred to the programme for health reasons, being away from normality, from urban living or from the everyday day stresses and strains of working life was important (Burls 2007 (n = 11), O'Brien 2008a (n = 88), Townsend 2004) (n = 18)).

Figure 13 Benefits of place
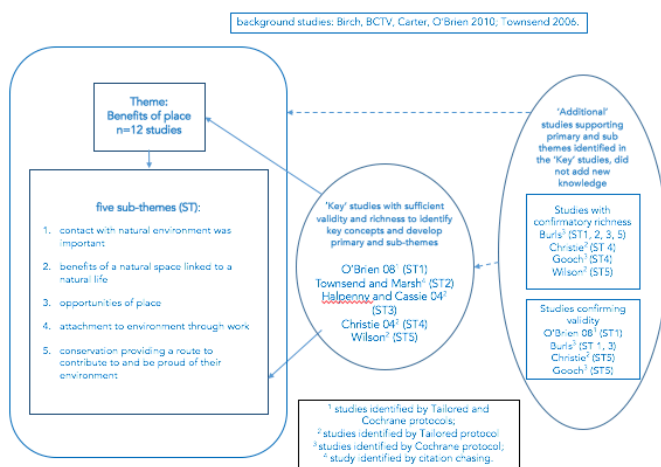
Social contact

The theme for which there was most frequent and rich description in the included qualitative studies was social contact. All included qualitative studies (and so all three good quality studies) reported themes relating to this, and there was little variation in content across different participant groups. From the studies, it appears that the activities were not completed in isolation but as part of a small team, which may have been part of a wider group of projects or programmes. Where descriptions of the projects were available, they showed that many aimed to harness the benefits of social contact.

There were clearly groups for which opportunities for social contact had greater impact; those experiencing previous isolation through mental ill health (BTCV 2010a (n =19), Wilson 2009 (n = 29)), and retirees (O'Brien 2008a (n = 88, good quality)) all reported benefits in terms of improved social networks. For these groups contact with other people had a positive effect and was seen as part of the recovery process:

"It helped me get out the house and meet people and join in the activities a bit more. I don't know if you're aware, I had depression, I wouldn't go out at all, I mean it's about a year ago, I wouldn't go out at all…" (Participant, O'Brien 2010a (n = 10, good quality)).

An important aspect of this was the unforced, relaxed nature of the social contact. Additional factors included undertaking shared activities, collaborative learning and companionable interactions.

"Everybody seems to get on and muck in together and if somebody was struggling you'd try to help them along…" (Participant, Wilson 2009 (n = 29)).

The neutrality of the setting and social contact was important for some:

"We all get on very well it's quite a close band of people. There's no hidden agenda; you don't need to know who the people are or what they do. You just come [and] enjoy the day that's the beauty of it." (Participant, O'Brien 2010a) (n = 10, good quality)).

Findings reported from the study of offenders in nature (Carter 2008 (n = unknown)) indicated that, for this group, it was seen as an achievement to be part of the general public without being verbally abused or derided, and engagement with visible improvements to the local natural environment enabled them to accomplish that. Participation had additional outcomes in the potential to facilitate positive re-engagement with family members:

"One offender, after a few weeks on the scheme, took his father out to show him the work he had completed. "It's nice feeling part of, ehm, part of society again" " (Participant, Carter 2008 (n = unknown))

The social contact through taking part in EECA also allowed individuals to develop wider support networks and to meet new people. In some cases the friendships were strong enough to encourage people to meet outside of the formal activity programme. For others, social contact was more focused on a

73

Commented [CC26]: s1 improved social networks (partic. for those with previous isolation)

Commented [CC27]: s2 neutrality of setting was important

Commented [CC28]: s3 social contact with public was achievement

Commented [CC29]: s4 meeting new or like-minded people

coming together of like-minded people with the purpose of improving the environment (Christie 2004 (n = 18), Caissie 2003 (n = 10, good quality), Gooch 2005 (n=85)).

Whilst subtly different to those at risk of isolation, the effect of social contact was no less frequently reported.

Figure 14 Social contact
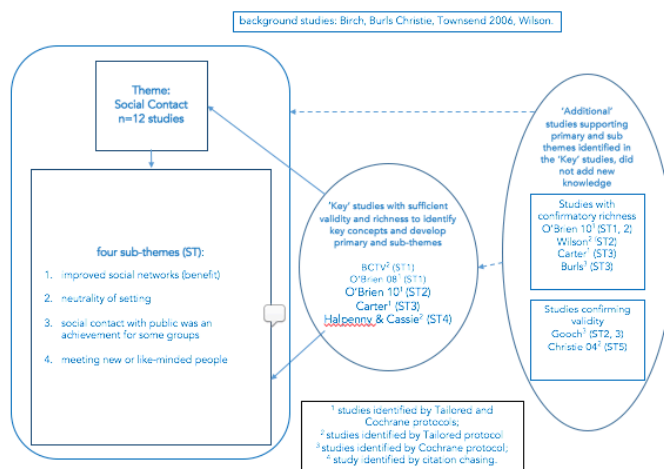
Physical activity

Seven studies, including one good quality study, specifically discussed the perceived benefits that participation in EECA could have in terms of opportunities for physical activity across the studies and activity types: Birch 2005; Burls 2007; Carter 2008; O'Brien 2008a (good quality); Townsend 2004; Townsend 2006; Wilson 2009. In the interviews and focus groups held by Townsend 2006 (n = 80) and Townsend 2004 (n = 18) participants associated membership of conservation volunteering groups with increased levels of physical activity. Indeed for some environmental volunteers the opportunity for activity was a primary motivation:

"I was advised to get exercise, so he re I am." (Participant, O'Brien 2008a (n = 88))

While these motivations might have been expected in those referred to the activities through health services, it was also found in other studies where participants could be considered the more 'traditional' type of volunteer, for example Townsend 2006 (n = 80). The conservation activities were also felt to be more engaging and interesting than other forms of exercise, perhaps aiding adherence to an exercise referral:

"The value of undertaking practical, outdoor, work was highlighted. This was felt to be rewarding compared to activity in a gym, for example." (Author, BTCV 2010a (n = 19))

For the offenders in the study by Carter 2008 (n = unknown) taking part in EECA provided an invaluable opportunity to be physical active:

*"Access to a gym is rare for prisoners; access to nature is even rarer. Those taking part in the schemes often comment how good it feels to be outside in the fresh air, and to be physically active throughout the day." (*Author, Carter 2008 (n = unknown)

Also of importance was the notion that participation in EECA was a route to better health through these increased levels of physical activity. All three respondents in the study by Birch 2005 (n = 3) reported that they felt that taking

75

Commented [GR30]:

Commented [GR31]: Subtheme 1 – PA associated with increased activity

Commented [GR32]: Subtheme 2 – people were motivated to take part in conservation BECAUSE they wanted to do more exercise.

Commented [GR33]: Subtheme 3 – conservation activities more interesting/rewarding than traditional types of exercise

Commented [GR34]: Subgroup 4: Exercise of particular relevance to offenders

part in Green Gym provided them with the opportunity to improve their health through this increased activity. The participants reported that the activity was linked to increased stamina, fewer injuries and reduced stress.

"I feel exhausted...but it's a de-stress ."(Participant, Birch 2005 (n = 3) review team ellipsis)

The participants in the studies undertaken by Burls 2007 (n = 11) and O'Brien 2008a (n = 88) also reported similar notions of increased physical health through higher levels of physical activity, levels of activity that were greater and potentially more varied than would have been undertaken without EECA:

"This is a superb way of keeping relatively fit. The physical is important, it's the buzz, tree felling it's a bit of a man's thing. Generally we want to get on and we are out there for the physical. It's good for muscle tone and keeping the beer off the belly." (Participant, O'Brien 2008a(n = 88)).
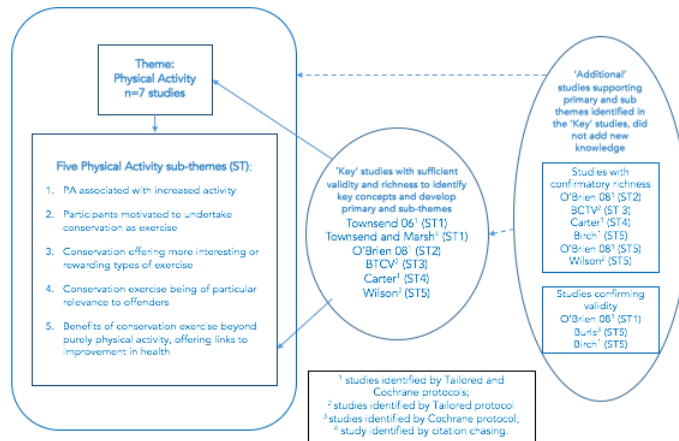
Weight loss, amongst other benefits, was also of importance to a participant, who had been referred to the programme by social and mental health services, in the Scottish 'Branching Out' programme:

"I feel it's actually benefited my health, because I do suffer from asthma. It seems as if I'm getting more fresh air and I feel a wee bit healthier and plus some of the work that they dae. I feel that, in a way it is making me lose a wee bit of weight. I used to be twenty stone now I'm only eighteen." (Participant, Wilson 2009 (n = 29)).

Physical activity was one of the key mechanisms though which the participants felt they benefited from engagement with EECA leading to increased fitness, weight loss, lowered stress and increased muscle strength. Positive attitudes were found across all user groups and activity types, but were, predictably, a particularly important focus of those taking part in the 'Green Gym' type activities.

Figure 15 Physical activity

Theme:
Physical Activity
n=7 studies

Five Physical Activity sub-themes (ST):

1. PA associated with increased activity

2. Participants motivated to undertake conservation as exercise

3. Conservation offering more interesting or rewarding types of exercise

4. Conservation exercise being of particular relevance to offenders

5. Benefits of conservation exercise beyond purely physical activity, offering links to improvement in health

'Key' studies with sufficient validity and richness to identify key concepts and develop primary and sub-themes
Townsend 06[1] (ST1)
Townsend and Marsh[4] (ST1)
O'Brien 08[1] (ST2)
BTCV[2] (ST3)
Carter[1] (ST4)
Wilson[2] (ST5)

'Additional' studies supporting primary and sub themes identified in the 'Key' studies, did not add new knowledge

Studies with confirmatory richness
O'Brien 08[1] (ST2)
BCTV[2] (ST 3)
Carter[1] (ST4)
Birch[1] (ST5)
O'Brien 08[1] (ST5)
Wilson[2] (ST5)

Studies confirming validity
O'Brien 08[1] (ST1)
Burls[3] (ST5)
Birch[1] (ST5)

[1] studies identified by Tailored and Cochrane protocols;
[2] studies identified by Tailored protocol
[3] studies identified by Cochrane protocol;
[4] study identified by citation chasing.

Spirituality

Notions of spirituality were reported in studies by BTCV 2010a (n = 19), Burls 2007 (n = 11), O'Brien 2008a (n = 88, good quality) and O'Brien 2010a (n = 10, good quality). This related to the previous theme (the importance of place) and was primarily understood as a connectedness to nature:

"On a personal level participants found their relationship with nature facilitating spiritual growth. Finding solace in nature." (Author, O'Brien 2010a (n = 10)).

The notions of peace and solitude in relation to being in the natural environment were common to each of the reports that considered spirituality. Christie 2004 (n = 18) reported participants feeling part of the land in which they were engaged and that was the single greatest motivator for being involved and outcome of engagement.

> **Commented [CC39]:** s1 spirituality as connected to nature
>
> **Commented [CC40]:** s2 peace and solitude – but perhaps not a theme rather defining what spirituality is.
>
> **Commented [CC41]:** s2 feeling part of land

Figure 16 Spirituality
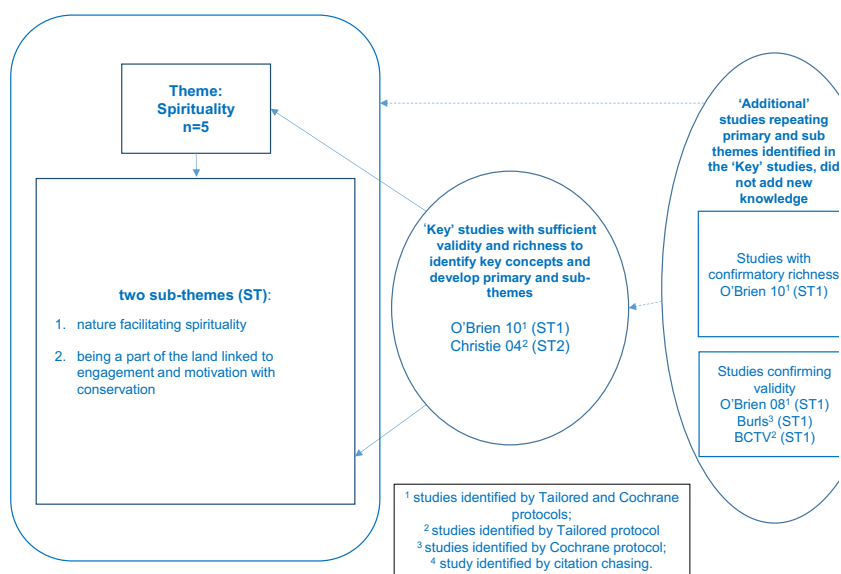
Psychological benefits

The myriad of perceived psychological benefits of EECA, aspects of which were also reported in all of the included qualitative studies (and therefore all three good quality studies), was strongly associated with each of the other themes. The impacts of achievement, for instance, were strongly linked to the positive emotions of accomplishing something, whether it was getting out of bed for someone suffering from depression, or, for a committed environmentalist, in making a real difference to an environment. This category encompasses discussions by participants on a range of mental benefits of participation in EECA including emotional response, quality of life or recovery from depression. Impacts could be multiple. For example, the respondents in the study by Wilson 2009 (n = 29) spoke about feeling more confident, having improved self-esteem, and better overall mental well-being.

The structure provided by repeated involvement with programmes of activities was again raised as having a central psychological effect, particularly for those experiencing some level of mental ill health or those at risk of social isolation (Birch 2005 (n = 3), BTCV 2010a (n = 19), and Wilson 2009 (n = 29) most markedly).

"it's getting me out the house and to me that in itself is a task, but it's a task worth doing, you know. I like to see the fruits of my labour." (Participant, Wilson 2009(n=29)).

The type of work which was undertaken in this structure was also important. Whilst it was physically (and occasionally emotionally) demanding work, it was also un-pressurised and flexible, which was important to respondents. Furthermore, being able to see the tangible impact of what was achieved appeared to motivate participation.

Related, though markedly different, were the impacts felt by the groups of participants who considered involvement in EECA to be altruistic. For this group, psychological benefits were accrued through the leaving of a legacy for future generations (Christie 2004 (n = 18) and Gooch 2005 (n = 85)). Indeed, one of the respondents in the study by Gooch 2005 (n = 85) referred to EECA participation as empowering.

*"Basically giving something back to nature because I've taken a lot from it."*
(Participant, Caissie 2003 (n = 10, good quality)).

As the individuals interviewed by Caissie 2003 (n = 10, good quality) had taken trips solely for the purpose of environmental enhancement it is not surprising that altruism was a major theme; respondents wanted to give something back to the environment.

Figure 17 Psychological benefits
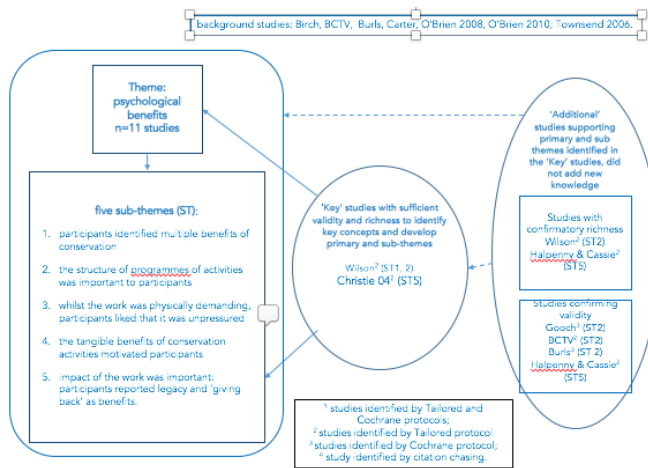
Risks and negative impacts

Very few included qualitative studies (and no good quality studies) reported any perceived risks or negative impacts associated with participation in EECA. Some even argued the potential risks were minimal:

"No more than normal life risks; only risks you put yourself in, but not other than that; it could happen in life anyway; it's safer than me riding my bike on the road." (Author, Burls 2007 (n = 11)).

Christie 2004 (n = 18) examined the experiences of Australians enrolled on a peri-urban environmental regeneration scheme and reported limited feelings of 'well-informed futility', amongst some participants. This sense of pointlessness came about when they realised the extent of the perceived problem and their in/ability to make a meaningful impact through activities. Similarly, Gooch 2005 (n = 85) reported some aspects of negative feeling associated with water-catchment restoration in Australia, where participants felt that their input was not sufficient and that more needed to be done. In this case a more positive connotation was reported, with individuals citing motivating future generations and sustainability of action as motivators to continued participation:

"There's a need here, I don't enjoy this [volunteering] at the moment, I must admit it. It's...it's killing me, but I've got to keep going, there's just too much at stake." (Participant, Gooch 2005 (n =85) review team ellipsis).

In both cases these participants were volunteers motivated to take part through their deeply held environmental concerns. Their focus on the significant challenges to the environment may have been greater than for those whose motivations were more modest.

Commented [CC49]: s1 limited data to support risks or negative impacts

Commented [CC50]: s2 risk minimal

Commented [CC51]: s3 futility, size of the task relative to the perceived problem contributed to negative feelings

Figure 18 Risks/ Negatives
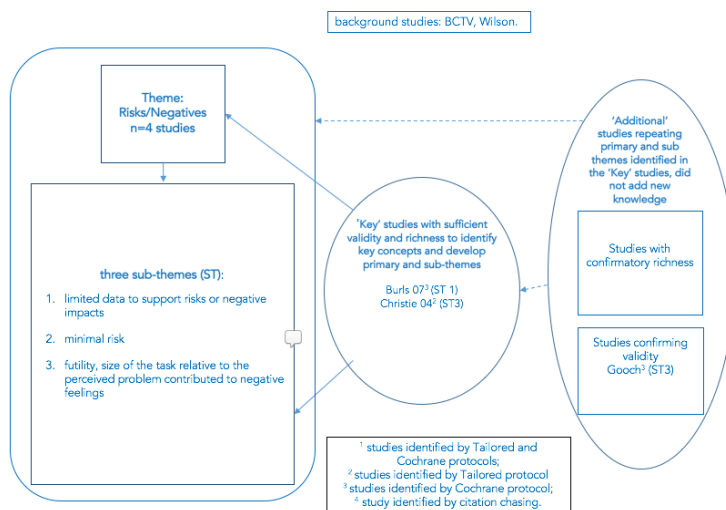
List of organisations contacted

The Conservation Volunteers (BTCV)	Hush Farms
2020 Vision	Isle of Anglesey County Council
Aaron Pyecroft	Isle of Wight AONB
Active Wales	Isles of Scilly AONB
Age UK	Keep Britain Tidy, Beach Care
Ambios	Keep Wales Tidy
Arnside and Silverdale AONB	Kent Downs AONB
Avon Wildlife Trust	Kent High Weald Partners
Bailies of Bennachie	Kent nat tr vol
Basingstoke con vol	LANTRA
BeachCare (Keep Britain Tidy)	Lea Bridge con vol
Berkshire con vol	LEAF/Let nature feed your senses
B'ham Guild (Broader)	Leicester con vol
Biodiversity SW	Lincolnshire Wolds AONB
Biosphere CLS	Liverpool PCT
Birmingham Guild for Student Colunteers	Llyn Peninsula AONB

Blackdown Hills AONB	London and w/msex vol
Blackdown Hills Hedge Association	Lothian con vol
Blackwater Valley countryside volunteers	Love where you live
Bolton conservation vol	Malvern Hills AONB
Bolton Wildlife Programme	Manchester nat tr vol
Bournemouth nat tr vol	Marine Conservation Society
Bracknell con vol	Medway Valley Countryside Partnership
British Waterways	Mendip Hills AONB
BVSC (Birmingham)	MIND (Eco Minds)
Cambridge con vol	MoD
Camp Kernow	Moor Trees
Cannock Chase AONB	NAAONB
Canterbury Environmental Education Centre	Nat Eng Big Lottery projs
Cardiff con vol	National Parks
Carymoor Env trust	National Trust
Causeway Coast and Glens Heritage Trust	National Trust for Scotland
CCD	Natural England
Change Agents UK	Natural England
Chichester Harbour AONB	Naturally Active project - Kent
Chichester Harbour AONB Officer	Neroche
Chilterns Conservation Board	New Forest Volunteers
City Farms and Community Gardens	Newlands Project
Clwydian Range AONB	Newquay Zoo
CN4C	NHS Forest
CoAST	Nidderdale AONB
Coastnet	Norfolk Coast AONB Partnership
Community Environmental Trust	Norfolk nat tr vol
Community Payback	North Devon AONB
Confor SW	North Devon Council
Conservation Foundation	North East Wales Wildlife

Conservation Volunteers Australia	North Pennines AONB Partnership
Glasgow Con Vol	North Wessex Downs AONB
Cornwall AONB	Northumberland Coast AONB
Cornwall Council	OPAL
Cotswolds Conservation Board	Outdoor and Experiential Learning Group
Countryside Recreation Network	Outdoor health forum
Countryside Trust	Oxford cons vol
Cove Brook Greenway group	Oxford Urban Wildlife Group
Coventry nat tr vol	Pembroke 21C
CPRE	People and Planet
Cranborne Chase and West Wiltshire Downs AONB	Plantlife International
CRCC	Plymouth Environmental Action
CRESH	Plymouth Student Scientist
CSV	PROSPECTS
Cusgarne Organic Farm	Quantock Hills AONB
CVS	Reforesting Scotland
Dartmoor Preservation Association	Rowhill con vol
Dean Green Team Volunteers	Royal Horticultural Society
Dedham Vale AONB and Stour Valley Project	RSPB
Derbyshire con vol	Scarborough con vol
DofE	Scottish Wildlife Trusts
Dorset AONB Partnership	Scouts
Durham Uni con vol	SeaSchool
Durlston Volunteers	Sheffield W'experience programmes
Earth Trust	Shropshire Hills AONB
East Devon AONB Partnership	Silvanus Trust
Egham/Staines con vol	Small Woods Association
Environment Kernow	SNCV (Sutton)
Epping forest con vol	SNH

Europarc	Snowdonia Society
FEVA	Solway Coast AONB
Fleet Pond Soc	Somerset Community Food
Forest of Bowland AONB	South Devon AONB Partnership
Forest Research	South Down National Park
Forest School	South West Environmental Action Trust
Forestry Commission	South West Lakes Trust
Forestry Commission Scotland	South West London Environment Network
Forum for Environmental Volunteering Activity	Steeple Woodland Reserve
Friends of Par Beach	Suffolk Coast & Heaths AONB Partnership
Friends of the earth	Surfers Against Sewage
Frimley Fuel Allot con team	Surrey Hills AONB
Froglife	Sustrans
Gibbonsdown and Court Partnership	Tamar Valley AONB Partnership
Glentress Trail Fairies	Teignbridge vols
Global Boarders	TFL volunteers
Gloucester vale con vol	Thames 21
Gower AONB	The Mendip Society
Green Space Community Network	THRIVE
Green Team	Venture Scotland
Greener Ilfracombe	vInspired
Greenham and Crookham con vol	Volunteer Bristol
Greenpeace Cornwall	Volunteer Cornwall
Groundwork	Volunteer development Scotland
Guernsey con vol	Wandle Trust
Haldon Forest Volunteers	Wednesday con vol
Haldon4Horses	West Country Rivers Trust
Hampshire con vol	Wicken Fen con vol
Harlow con vol	Wildlife Trust

Haven Holidays	Wirral county vols
High Weald AONB JAC	Woodland Trust
Highland Environmental Network	Wychwood Project
Hill Holt Wood	Wycombe District Council
Howardian Hills AONB	Wye Valley AONB

Annex 6.   Website hand searches

1. Environment;
2. Conservation;
3. (1) and (2);
4. Environmental enhancement;
5. Volunteering; and
6. Health/well-being.

Website URL	Website Name
http://www.dtic.mil/whs/directives/index.html	US Military DoD
http://www.ccw.gov.uk/default.aspx	Countryside Council for Wales
http://www.tsrc.ac.uk/	Third Sector Research
http://www.vssn.org.uk/	Voluntary Sector Studies Network - Journal
http://www.ivr.org.uk/ivr-evidence-bank?q=&t%5B%5D=362	Institute for Volunteer Research
http://www.naturaleconomynorthwest.co.uk/	Natural Economy North West
http://www.oecd-ilibrary.org/;jsessionid=136d54v2tehqa.delta	OECD iLibrary
http://www.oecd.org/department/0,3355,en_2649_33713_1_1_1_1_1,00.html	OECD Environmental Directorate
http://www.epa.gov/	US Environmental Protection Agency
http://hero.epa.gov/index.cfm	Health and Environmental Research Online - US
http://www.eea.europa.eu/	European Environment Agency
http://www.npca.org/	US National Parks Conservation
http://www.environment.gov.au/	Australian Environment Agency
http://www.ec.gc.ca/default.asp?lang=en	Environment Canada
http://www.npws.ie/	Ireland Parks
http://www.epa.ie/	Environmental Protection Ireland
http://www.doeni.gov.uk/niea/	Dept of Env Ireland
http://www.epa.govt.nz/Pages/default.aspx	NZ EPA
http://www.doc.govt.nz/publications/	NZ Conservation Authority
http://www.forestry.gov.uk/publications	Forestry Commission
http://www.forestry.gov.uk/forestresearch	Forest Research

http://www.snh.gov.uk/publications-data-and-research/	Scottish Natural Hertitage
http://www.feva-scotland.org/display/library	FEVA
http://publications.naturalengland.org.uk/	Natural England
http://www.sehn.org/	Science and Environmental Health Network
http://www.sustainweb.org/publications/	Sustain Web
http://www.fph.org.uk/policy%2c_publications_and_events	Faculty of Public Health College - London
http://www.carefarminguk.org/case-studies.aspx	Care Farming UK
http://www.mod.uk/DefenceInternet/MicroSite/DIO/OurPublications/EstateAndSustainableDevelopment/Sanctuary.htm	Sanctuary Magazine, MoD UK
http://www.hphpcentral.com	International 'healthy parks healthy people' network
http://www.ecohealth.net	International association for ecology and health
http://www.ramsar.org/cda/en/ramsar-pubs/main/ramsar/1-30_4000_0__	Healthy wetlands and healthy people initiative of Ramsar Convention on Wetlands
http://www.cbd.int/	Healthy planet healthy people initiative of the convention on biodiversity
www.saveourseine.com/	Save our Seine
http://www.landcareonline.com/; http://www.landcareonline.com.au/?page_id=9608	Landcare online