

# Risk of bias in overviews of reviews: a scoping review of methodological guidance and four-item checklist

Madeleine Ballard\* and Paul Montgomery

**Objective:** To assess the conditions under which employing an overview of systematic reviews is likely to lead to a high risk of bias.

**Study Design:** To synthesise existing guidance concerning overview practice, a scoping review was conducted. Four electronic databases were searched with a pre-specified strategy (PROSPERO 2015: CRD42015027592) ending October 2015. Included studies needed to describe or develop overview methodology. Data were narratively synthesised to delineate areas highlighted as outstanding challenges or where methodological recommendations conflict.

**Results:** Twenty-four papers met the inclusion criteria. There is emerging debate regarding overlapping systematic reviews; systematic review scope; quality of included research; updating; and synthesizing and reporting results. While three functions for overviews have been proposed—identify gaps, explore heterogeneity, summarize evidence—overviews cannot perform the first; are unlikely to achieve the second and third simultaneously; and can only perform the third under specific circumstances. Namely, when identified systematic reviews meet the following four conditions: (1) include primary trials that do not substantially overlap, (2) match overview scope, (3) are of high methodological quality, and (4) are up-to-date.

**Conclusion:** Considering the intended function of proposed overviews with the corresponding methodological conditions may improve the quality of this burgeoning publication type. Copyright © 2017 John Wiley & Sons, Ltd.

**Keywords:** umbrella review; meta-review; overview of reviews; scoping review; synthesis methods

## 1. Introduction

Overviews of reviews compile information from multiple systematic reviews to provide a comprehensive synthesis of evidence (Higgins and Green, 2011). While systematic reviews bring together evidence from multiple studies in a rigorous fashion for a single intervention or group of interventions, a limitation of such systematic reviews as they have evolved is that they are often narrow in scope (Hartling *et al.*, 2014). They typically focus on direct pairwise comparisons and lack formal comparisons of competing interventions for a given condition or evidence on multiple applications of the same intervention (Hartling *et al.*, 2014). Overviews facilitate these broad comparisons and are similar to systematic reviews in that they use “systematic and explicit methods to identify, select and critically appraise relevant research, and to collect and analyse data” in order to answer clearly formulated questions that facilitate evidence-informed decision-making, but differ in their unit of synthesis: systematic reviews rather than primary studies (Hartling *et al.*, 2014; Higgins and Green, 2011).

Overviews examine: (1) different interventions for the same condition or population (e.g. Jones *et al.*, 2012), (2) the same intervention for different conditions or populations (e.g. Steultjens *et al.*, 2005), (3) multiple outcomes of the same intervention for the same condition or population (e.g. Flodgren *et al.*, 2011), or (4) adverse effects from

Centre for Evidence-Based Intervention, University of Oxford, Oxford, UK

\*Correspondence to: Madeleine Ballard, Centre for Evidence-Based Intervention, University of Oxford, 32 Wellington Square, Oxford OX1 2ER, UK.

E-mail: madeleine.ballard@gmail.com

the same intervention across multiple conditions (e.g. Singh Jasvinder *et al.*, 2011). Nomenclature for this review type varies (Hartling *et al.*, 2012); overviews are referred to as 'umbrella reviews', 'meta-reviews', 'overviews of systematic reviews', 'reviews of reviews', and 'systematic review of systematic reviews', among others; each of these terms has its adherents and has been used in recent publications (O'Donnell *et al.*, 2014; Theodoratou *et al.*, 2014; Martínez-González *et al.*, 2014; Hill and Ryan, 2014; Lamming *et al.*, 2014). This paper will use the term overview.

While overviews are not new (e.g. Meltzer, 1972), there remains a lack of clarity around overview methodology (Cochrane Comparing Multiple Interventions Methods Group, 2012; Pieper *et al.*, 2012). Given that the number of overviews published per year between 2004 and 2011 has risen fivefold, from six to thirty (Pieper *et al.*, 2012), and that overviews of reviews are the most downloaded content in certain evidence synthesis journals (Thomson, 2014), it is critical that this methodological haziness be addressed so as to avoid dissemination of potentially misleading overview results.

In order to (1) identify and synthesize existing guidance concerning overview practice and (2) assess when overviews are likely to produce valid results and when they are best avoided, a scoping review was conducted (PROSPERO 2015:CRD42015027592). This is the first attempt to systematically summarize such guidance.

## 2. Methods

### 2.1. Inclusion and exclusion criteria

Studies in any language were included if their primary focus was to describe or develop overview methodology. Studies were excluded if they were published before Jan 1, 2000 as (1) very few overviews were published before this and (2) any guidelines that exist are very likely to be out of date.

### 2.2. Systematic search strategy

Searches were run on Oct 27 2015 on MEDLINE and MEDLINE In-Process and Other Non-Indexed Citations, PsycINFO, and Embase, all on the Ovid platform. The following strategy was developed to achieve a high level of precision and specificity without sacrificing sensitivity.

1. (overviews adj3 review\*).m\_titl.
2. (umbrella adj5 reviews).m\_titl.
3. (meta-reviews or (meta adj reviews)).m\_titl.
4. (review adj reviews).m\_titl.
5. "systematic reviews of systematic reviews".m\_titl.
6. Or/1–5
7. (overview adj3 review\*).m\_titl.
8. (umbrella adj5 review).m\_titl.
9. (meta-review or (meta adj review)).m\_titl.
10. "systematic review of systematic reviews".m\_titl.
11. Or/7–10
12. (11 adj7 (challenges or method\* or how or approach or conduct\*)).m\_titl
13. 6 or 12
14. limit 13 yr = "2000–2015"

This is a specific search aiming to answer a well-focused question and that this is reflected in search terms. In addition to electronic searches, (1) contact was made with study authors to solicit relevant unpublished papers, on-going research, and suggestions for other contacts, (2) reference lists of included studies were examined for other studies that meet the inclusion criteria, and (3) webpages of leading evidence-synthesis experts were searched (listed in Appendix A of the Supporting Information).

### 2.3. Selection of studies

Titles and abstracts were independently screened by two authors in order to identify studies that met the eligibility criteria. Both authors then retrieved the full text of studies selected as meeting or possibly meeting the criteria and verified them again in order to produce a final list of included studies. Any discrepancies were discussed until agreement was reached.

### 2.4. Data extraction

For each paper, the following information was extracted by one reviewer into Microsoft Excel: author, year, title, overview methodological challenges delineated, methodological recommendations, and how recommendations

were derived. Results were then narratively synthesized by examining commonalities and differences across included studies. In order to improve data quality, a second reviewer audited extracted data.

## 2.5. Data synthesis

This review delineates (1) areas highlighted as outstanding challenges by overview methodology papers and (2) areas that have not been identified as challenges but where recommendations conflict.

Further to systematically summarizing overview methods guidance, the paper will attempt to operationalize relevant methodological debates into a decision making tool that will help improve the quality of overviews produced. This was an addition to the methods described in the review protocol; such a tool was believed to be the most practical way to summarise outstanding issues with included guidance.

## 3. Results

Excluding duplicates, a total of 48 records were screened for inclusion. Of these, 36 were recovered from electronic database searches, 2 from handsearching, and 9 from references. Thirty-one full text papers were considered for inclusion in the scoping review; 24 papers (Aromataris *et al.*, 2014; Cochrane Comparing Multiple Interventions Methods Group, 2012; Aromataris *et al.*, 2015; Baker *et al.*, 2014; Becker and Oxman, 2011; Caird *et al.*, 2015; Cochrane Comparing Multiple Interventions Methods Group, 2013; Conn and Coon Sells, 2014; Cooper and Koenka, 2012; Hartling *et al.*, 2012; Hartling *et al.*, 2014; Ioannidis, 2009; Li *et al.*, 2012; Pieper *et al.*, 2014a; Pieper *et al.*, 2014c; Pieper *et al.*, 2012; Piso *et al.*, 2015; Santaguida *et al.*, 2013; Smith *et al.*, 2011; Thomson *et al.*, 2013; Thomson *et al.*, 2010; Pieper *et al.*, 2014b; Ryan *et al.*, 2009; Yuan *et al.*, 2012) reporting 22 guidelines or descriptive studies met the inclusion criteria (see Figure 1).

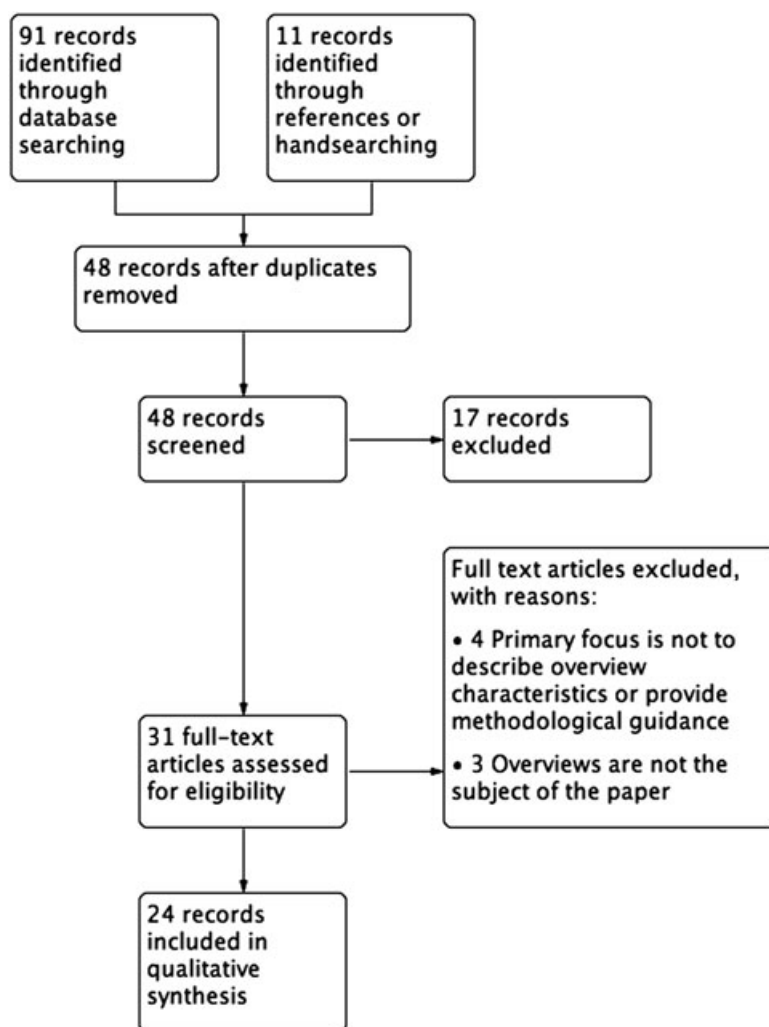


Figure 1. PRISMA study flow diagram

### 3.1. Characteristics of the sample

Of the papers included in the scoping review, six described four sets of methodological guidelines assembled by a working group or editorial team (Aromataris *et al.*, 2014; Aromataris *et al.*, 2015; Becker and Oxman, 2011; Cochrane Comparing Multiple Interventions Methods Group, 2013; Conn and Coon Sells, 2014; Cochrane Comparing Multiple Interventions Methods Group, 2012); 10 were narrative reviews or descriptions of authors' anecdotal experiences conducting overviews (Baker *et al.*, 2014; Caird *et al.*, 2015; Cooper and Koenka, 2012; Hartling *et al.*, 2014; Ioannidis, 2009; Santaguida *et al.*, 2013; Smith *et al.*, 2011; Thomson *et al.*, 2010; Piso *et al.*, 2015; Ryan *et al.*, 2009); seven were systematic reviews in which authors used a pre-specified, comprehensive search strategy to assemble all overviews published during a given date range in order to describe their methodological characteristics and reporting (Hartling *et al.*, 2012; Li *et al.*, 2012; Pieper *et al.*, 2014a; Pieper *et al.*, 2014c; Pieper *et al.*, 2012; Thomson *et al.*, 2013; Yuan *et al.*, 2012); and one was a systematic review of HTA guidelines (Pieper *et al.*, 2014b). Papers are listed by type in Table 1 and the objective and methodological issues addressed by each summarized. Excluded papers, along with reasons for exclusion, are listed in Appendix B of the Supporting Information.

Included records define overviews as “systematic reviews of systematic reviews” (Pieper *et al.*, 2012; Baker *et al.*, 2014; Caird *et al.*, 2015; Ryan *et al.*, 2009; Smith *et al.*, 2011) and indicate that they serve the following three functions:

1. To identify gaps in the literature where multiple comparable studies may exist but a research synthesis has not been performed (Caird *et al.*, 2015; Piso *et al.*, 2015; Ryan *et al.*, 2009; Cooper and Koenka, 2012; Santaguida *et al.*, 2013)
2. To compare and contrast existing systematic reviews (Aromataris *et al.*, 2014; Aromataris *et al.*, 2015; Baker *et al.*, 2014; Conn and Coon Sells, 2014; Cooper and Koenka, 2012; Pieper *et al.*, 2014a; Pieper *et al.*, 2012; Santaguida *et al.*, 2013; Smith *et al.*, 2011).
3. To provide a summary of evidence from existing systematic reviews, with or without synthesis (Becker and Oxman, 2011; Caird *et al.*, 2015; Cochrane Comparing Multiple Interventions Methods Group, 2012; Cochrane Comparing Multiple Interventions Methods Group, 2013; Cooper and Koenka, 2012; Hartling *et al.*, 2012; Hartling *et al.*, 2014; Li *et al.*, 2012; Pieper *et al.*, 2014a; Pieper *et al.*, 2012; Piso *et al.*, 2015; Ryan *et al.*, 2009; Smith *et al.*, 2011; Thomson *et al.*, 2013; Thomson *et al.*, 2010).

### 3.2. Methodological guidance and debates

The majority of included papers noted, either explicitly or implicitly, that many of the methodological standards for systematic reviews could also be applied to overviews (Pieper *et al.*, 2012; Becker and Oxman, 2011; Pieper *et al.*, 2014a; Santaguida *et al.*, 2013). For example, there was a consensus that, as with systematic reviews, overviews ought to set a-priori inclusion and exclusion criteria that specify the purpose of the overview (Cochrane Comparing Multiple Interventions Methods Group, 2013; Smith *et al.*, 2011) and describe features of systematic reviews to be included in enough detail so as to enable a reader to repeat the selection (Ioannidis, 2009; Li *et al.*, 2012; Smith *et al.*, 2011; Thomson *et al.*, 2010; Santaguida *et al.*, 2013).

Similarly, overview guidelines indicated that search strategies ought to be pre-specified (Aromataris *et al.*, 2014; Becker and Oxman, 2011; Conn and Coon Sells, 2014), incorporate manual searches of key journals and of reference lists of included systematic reviews (Conn and Coon Sells, 2014), and—with the exception of the Cochrane Collaboration (Becker and Oxman, 2011) which notes that overview authors should “typically” search only for relevant Cochrane reviews—include multiple online databases (Conn and Coon Sells, 2014; Aromataris *et al.*, 2014). These recommendations differ from systematic review methodology only in that, rather than cover many decades as is typical in a systematic review search, overview searches rarely need to extend prior to 1990 as all but a small minority of systematic reviews were conducted after that time (Baker *et al.*, 2014; Smith *et al.*, 2011; Aromataris *et al.*, 2014).

There is, however, emerging debate related to (i) overlapping systematic reviews, (ii) the scope of systematic reviews, (iii) evaluating the quality and reporting of included research, (iv) updating included systematic reviews, and (v) synthesizing and reporting the results of included systematic reviews.

The questions of whether overviews will save time, when they are likely to produce valid results and when they are best avoided, will be explored through a discussion of these five debates.

#### (i) Overlapping systematic reviews

A systematic review of overviews, found that nearly half (46.7%) of overviews published between 2009 and 2011 did not mention overlaps (Pieper *et al.*, 2014a). When included systematic reviews are meta-analysed, using data from individual studies more than once without accounting for overlap will overestimate statistical power and thus risk producing a misleading, overly precise estimate (Pieper *et al.*, 2014c; Smith *et al.*, 2011). When systematic reviews are narratively summarized, a similar problem occurs if steps are not taken to ensure certain primary studies are not overrepresented (Thomson *et al.*, 2010; Pieper *et al.*, 2014c; Caird *et al.*, 2015). Several possible approaches to this issue have been proposed, though no consensus has been reached.

**Table 1.** Characteristics of included studies.

Reference	Objective	Inclusion criteria	Search strategy	Data extraction and analysis	Quality: individual studies	Quality: systematic reviews	Quality of evidence	Updating	Overlap	Scope
<b>Methodological guidelines</b>										
Aromataris <i>et al.</i> , 2014,	Provide methodological guidance for the conduct of JBI overviews	✓	✓	✓	✓	✓	✓	✓	✓	✓
Aromataris <i>et al.</i> , 2015	Provide guidance to authors for the preparation of Cochrane Overviews	✓	✓	✓	✓	✓	✓	✓	✓	✓
Becker and Oxman, 2011	Present a set of recommendations for comparing multiple interventions	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cochrane Comparing Multiple Interventions Methods Group, 2013;										
Cochrane Comparing Multiple Interventions Methods Group, 2012										
Conn and Coon Sells, 2014	Provide guidance for authors contemplating preparation of overviews for WJNR	✓	✓	✓	✓	✓	✓	✓	✓	✓
<b>Narrative reviews and descriptions of authors' anecdotal experiences</b>										
Baker <i>et al.</i> , 2014	Describe the first overview of the Cochrane Public Health Group (CPHG) and propose a way forward for the methodologies under consideration	✓	✓	✓	✓	✓	✓	✓	✓	✓
Caird <i>et al.</i> , 2015	Highlight the challenges inherent in the overview method and present, where possible, potential solutions to these challenges	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cooper and Koenka, 2012	Describe what methods overviews have developed or have adopted from other forms of scholarship.	✓	✓	✓	✓	✓	✓	✓	✓	✓
Hartling <i>et al.</i> , 2014	Describe and compare the contributions made by, and the challenges of producing three types of evidence synthesis	✓	✓	✓	✓	✓	✓	✓	✓	✓

(Continues)

**Table 1.** (Continued)

Reference	Objective	Inclusion criteria	Search strategy	Data extraction and analysis	Quality: individual studies	Quality: systematic reviews	Quality of evidence	Updating	Overlap	Scope
Ioannidis, 2009	Review the main features, strengths, and limitations of methods that integrate evidence across multiple meta-analyses.	✓		✓	✓			✓		
Piso <i>et al.</i> , 2015	Assess methods and strengths/limitations of overviews and discuss their applications using examples from two Austrian institutions.	✓	✓	✓	✓	✓			✓	
Ryan <i>et al.</i> , 2009	Outline a method for preparing data integration tables to enable review-level evidence synthesis			✓	✓	✓				
Santaguida <i>et al.</i> , 2013	Describe the methodology used to conduct an overview	✓	✓	✓	✓	✓		✓	✓	
Smith <i>et al.</i> , 2011	Provide a guide to clinicians and researchers who wish to conduct an overview.	✓	✓	✓	✓	✓	✓			
Thomson <i>et al.</i> , 2010	Discuss outstanding challenges for overview authors and the limitations of overviews.	✓	✓	✓	✓	✓	✓	✓	✓	✓
<b>Systematic reviews of overview methodology and reporting or HTA guidelines</b>										
Hartling <i>et al.</i> , 2012	Inform the methodology for overviews by (1) identifying all overviews of health care interventions published 2000–2011, and (2) describing the overviews with respect to their methodological approaches.	✓	✓	✓	✓	✓	✓			
Li <i>et al.</i> , 2012	To evaluate the reporting and methodological quality of overviews of systematic reviews.	✓	✓			✓	✓			
Pieper <i>et al.</i> , 2012		✓	✓			✓	✓	✓	✓	

(Continues)

**Table 1.** (Continued)

Reference	Objective	Inclusion criteria	Search strategy	Data extraction and analysis	Quality: individual studies	Quality: systematic reviews	Quality of evidence	Updating	Overlap	Scope
	Examine published overviews of systematic reviews in terms of descriptive and methodological characteristics.									
Pieper 2014A	Summarize Health Technology Assessment (HTA) agency guidance for and applied methodology in overviews		✓	✓		✓	✓		✓	
Pieper 2014B	Examine published overviews of systematic reviews in terms of updating techniques, publication lag in overviews and to develop recommendations for authors of overviews.	✓	✓			✓	✓	✓	✓	
Pieper 2014C	Determine if the authors mention overlapping reviews in overviews and calculate the actual overlap in published overviews using newly introduced, validated measures.								✓	
Thomson <i>et al.</i> , 2013	Describe some aspects of the approaches taken in EBCH overviews	✓	✓					✓		✓
Yuan <i>et al.</i> , 2012	To survey and analyze the quality assessment of the included studies in overviews				✓	✓	✓			



A first approach, put forth by Pieper *et al.* (2014c), is the corrected covered area (CCA)—a validated method to calculate the degree of overlap in an overview. The CCA calculates overlap by dividing the frequency of repeated occurrences of the index publication (the first occurrence of a primary publication) in other systematic reviews by the product of index publications and systematic reviews less the number of index publications. The CCA prevents potential pitfalls of simply reporting primary studies in a grid (Cooper and Koenka, 2012); because the numerator only counts primary publications included in more than one systematic review it diminishes the impact of large systematic reviews. While two guidelines (Aromataris *et al.*, 2014; Conn and Coon Sells, 2014) both indicate that a “clear indication” of overlap should be presented, reporting overlap does not solve the problem of producing misleadingly precise estimates as it does not eliminate the overlap that can influence the precision of the results.

Rather than simply including all syntheses regardless of overlap, a second approach is to develop a priori criteria for choosing a single systematic review for inclusion when multiple potential candidates are available (Cochrane Comparing Multiple Interventions Methods Group, 2012). Cooper and Koenka (2012) suggest selecting the synthesis that (1) provides the most complete description, (2) is most recent, (3) contains the most evidence, (4) is methodologically most rigorous, or (5) is published (vs. in the grey literature). It is possible, however, that this approach would still result in excluding a systematic review that had relevant studies.

The Cochrane Collaboration’s stipulation that their overviews should typically only include Cochrane systematic reviews avoids this issue and so can be counted as a third approach (Becker and Oxman, 2011). Pieper *et al.* (2012), however, point out that while surveys of these systematic reviews have demonstrated that Cochrane reviews tend to be of higher quality than non-Cochrane reviews (Shea *et al.*, 2002; Moseley *et al.*, 2009; Jørgensen *et al.*, 2006; Olsen *et al.*, 2001), it is an ecological fallacy to suggest that for this reason they merit being preferentially included. Moreover, as both narrative (Caird *et al.*, 2015) and systematic papers (Pieper *et al.*, 2012) have argued, other systematic reviews may be more up-to-date, have a different focus, include additional data (e.g. on adverse effects), or incorporate new analyses (e.g. on various subgroups) that could increase the value of an overview. Excluding these systematic reviews from an overview would omit potentially important information. Thus, the problem of overlap in overviews currently remains unresolved (Cochrane Comparing Multiple Interventions Methods Group, 2013).

### (ii) Scope of systematic reviews

Where the scope of the individual systematic reviews is more heterogeneous and does not align as closely with the overview question, several papers (Caird *et al.*, 2015; Cooper and Koenka, 2012; Conn and Coon Sells, 2014; Cochrane Comparing Multiple Interventions Methods Group, 2012) note that authors of overviews may encounter the problem of scope mismatch, in which the scope of the overview does not match that of individual systematic reviews (see Figure 2). Thomson *et al.* (2010) provide the example of overviews about child health; if a systematic review includes both adults and children, it is unclear whether it ought to be included. Where the scope of some systematic reviews lies outside the scope of the overview, authors of such overviews face a dilemma: either they (1) include the systematic review, undertaking the time-consuming process of extracting relevant data from primary studies (Thomson *et al.*, 2010; Caird *et al.*, 2015) or (2) in cases where guidelines do not recommended that authors retrieve primary studies (Aromataris *et al.*, 2014), exclude the systematic review, risking omission of relevant results and thereby potentially introducing bias into the outcome (Caird *et al.*, 2015).

Exacerbating this difficulty, as Caird *et al.* (2015) note, systematic reviews do not always report all data from included primary research. This is, of course, impossible to tell in the absence of examining primary trials, yet retrieving and examining primary trials negates the anticipated efficiency gain of conducting an overview (Caird *et al.*, 2015).

### (iii) Evaluating the quality and reporting of included research

Included studies recommend overview authors address three quality assessments: methodological quality of trials included in systematic reviews, of included systematic reviews themselves, and of evidence across included systematic reviews.

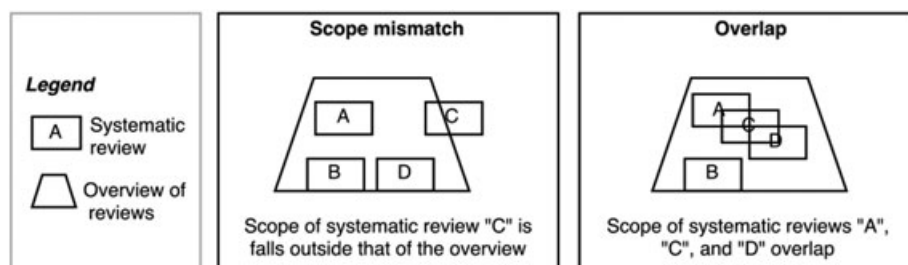


Figure 2. Scope mismatch versus overlap



3.3.1. *Primary trials.* Several papers (Cooper and Koenka, 2012; Conn and Coon Sells, 2014; Caird *et al.*, 2015; Aromataris *et al.*, 2015; Yuan *et al.*, 2012) note that considering the quality-assessment system used in included syntheses and evaluating whether it is appropriate is an important step in conducting an overview. The Cochrane Handbook recommends the GRADE approach for systematic reviews and that overview authors should report the quality of evidence as assessed by the systematic review authors (Becker and Oxman, 2011).

In the event that an appropriate quality assessment was not used in the original systematic review, overview authors will need to either (1) assess the risk of bias themselves (Becker and Oxman, 2011)—a time-consuming process in which data is extracted at trial level—or (2) include only systematic reviews that have detailed reporting of the quality of primary studies—an approach that increases efficiency by possibly excluding relevant trials from the resulting evidence base (Caird *et al.*, 2015). This may explain why information on quality of individual studies was extracted from fewer than 40% of overviews over the last decade (Hartling *et al.*, 2012).

3.3.2. *Systematic reviews.* While included guidelines agree that included systematic reviews must be assessed for methodological quality (Aromataris *et al.*, 2014; Aromataris *et al.*, 2015; Becker and Oxman, 2011; Cochrane Comparing Multiple Interventions Methods Group, 2013; Conn and Coon Sells, 2014; Cochrane Comparing Multiple Interventions Methods Group, 2012), there is no consensus on what instrument should be used (Yuan *et al.*, 2012). A descriptive analysis of overviews (Hartling *et al.*, 2012) found that among overviews published between 2000 and July 2011, only 37% assessed the quality of included systematic reviews at all, where this was done a diversity of instruments (Oxman and Guyatt, 1991; Shea *et al.*, 2007; Moher *et al.*, 1999; Jadad *et al.*, 1997) were used—a result echoed in the other descriptive analyses (Li *et al.*, 2012; Hartling *et al.*, 2012; Hartling *et al.*, 2014).

As Pieper *et al.* (2012) point out, it is important to differentiate between methodological quality, which considers how well the systematic review was conducted, and reporting quality, which considers how well systematic reviewers have reported their methodology and findings. For this reason, it has been recommended that PRISMA be used in conjunction with a comprehensive, validated critical appraisal tool (Pieper *et al.*, 2012; Shea *et al.*, 2007; Oxman and Guyatt, 1991).

The question of what to do if systematic reviews are of low quality remains similarly unresolved (Pieper *et al.*, 2014b). Excluding low-quality systematic reviews will likely give an incomplete picture of the data (Caird *et al.*, 2015), and attempting to compensate by performing new searches (Pieper *et al.*, 2014c) has its own difficulties that will be discussed in the next section.

3.3.3. *Quality of the body of evidence across included systematic reviews.* No method exists to evaluate the strength of evidence for findings across different systematic reviews (Santaguida *et al.*, 2013). As Santaguida *et al.* (2013) explain, some GRADE criteria are only applicable to primary studies. For example, GRADE stipulates that quality of evidence can be upgraded in the case of large effect sizes—a criterion not applicable to narratively synthesized systematic reviews. Although they describe an approach to apply GRADE to systematic reviews (Santaguida *et al.*, 2013), as Hartling *et al.* (2014) note, guidance is needed to ensure appropriate use and interpretation of GRADE when not applied to the primary studies for which the tool was designed. This has yet to be done.

#### (iv) Updating included systematic reviews

For an overview to be useful for decision-making, underlying systematic reviews must be up-to-date (Pieper *et al.*, 2014c; Cooper and Koenka, 2012; Hartling *et al.*, 2014; Thomson *et al.*, 2010). In a study of how quickly systematic reviews go out of date cited by several included papers (Pieper *et al.*, 2014c; Pieper *et al.*, 2012; Caird *et al.*, 2015), the median duration of survival free of a signal for updating<sup>1</sup> is 5.5 years (CI, 4.6 to 7.6 years), with a signal having occurred for 23% of systematic reviews after 2 years and for 15% of systematic reviews within 1 year (Shojania *et al.*, 2007).

These figures are striking because in their systematic review of overview characteristics, Pieper *et al.* (2014c) found that the mean publication lag in overviews (defined as publication date of the overview minus publication date of the systematic review) was more than 5 years; only one in four overviews considered up-to-dateness and methods for updating systematic reviews had high heterogeneity.

While some guidelines ignore this issue, recommending that authors not undertake an update of included systematic reviews or a new systematic review within the overview (Aromataris *et al.*, 2015; Becker and Oxman, 2011), others propose asking subject specialists to provide relevant primary research (Caird *et al.*, 2015) or considering whether a new search for recent primary studies may be necessary (Pieper *et al.*, 2012; Pieper *et al.*, 2014b). Pieper *et al.* (2014c) distinguished two ways of searching for primary studies in overviews: (1) updating

<sup>1</sup>Quantitative signals for updating were defined as changes in statistical significance or relative changes in effect magnitude of at least 50% involving a primary outcome or any mortality outcome; qualitative signals were defined as substantial differences in characterizations of effectiveness, new information about harm, and caveats about the previously reported findings that would affect clinical decision making (Shojania *et al.*, 2007).

included systematic reviews and (2) searching for secondary and primary literature simultaneously. Both approaches are also relevant in instances where there is not a systematic review for all interventions relevant to a particular overview (Ioannidis, 2009; Thomson *et al.*, 2010).

Updating included systematic reviews in the context of an overview, however, creates the added complexity of deciding how to apply inclusion/exclusion criteria of the systematic reviews to new trials: it is unclear how to proceed if a trial examines a comparison clinically meaningful to the overview but does not clearly match the systematic review criteria; for example, the population specified (Thomson *et al.*, 2013). Moreover, updating low-quality systematic reviews by searching only the time span not covered by the systematic review might be pointless if it did not initially include a comprehensive search strategy (Pieper *et al.*, 2014c).

Although option (2), performing a search for secondary and primary literature simultaneously solves both problems, it will undermine the purported time-savings of conducting an overview and raises the question of why one would undertake an overview, rather than a broad systematic review, in the first place (Pieper *et al.*, 2014c). This is particularly true given that trials introduce complexity into the overview synthesis process (discussed in the next section) that they do not in the context of systematic reviews (Baker *et al.*, 2014). Indeed, it was as a result of confusion generated by the Cochrane Comparing Multiple Interventions Methods Group (2012) (CMIMG) attempt to broaden the Cochrane definition of an overview such that individual studies could be included that the CMIMG reversed their decision a year later (Cochrane Comparing Multiple Interventions Methods Group, 2013).

Either way, while Pieper *et al.* (2012) suggest that the decision to update/search for primary trials could be based on the strength of evidence in included systematic reviews, or the research activity in the respective field, it remains the case that there is no way to systematically investigate whether an update in the context of overviews is necessary (Pieper *et al.*, 2014c; Piso *et al.*, 2015).

(v) Synthesizing and reporting the results of included systematic reviews

While debates i–iv deal with the trial selection stage, methods for synthesizing the results of multiple systematic reviews are also evolving.

Some guidelines (Aromataris *et al.*, 2014; Aromataris *et al.*, 2015; Becker and Oxman, 2011; Ryan *et al.*, 2009) indicate that overview findings should not be re-synthesized, and should be presented in tables or figures. While Caird *et al.* (2015) argue that such overviews are a useful tool in comparing the consistency of scope and findings across systematic reviews, others have called the qualitative juxtaposition of data from separate meta-analyses “subjective and suboptimal” (Ioannidis, 2009) as it facilitates “informal” indirect comparison that may be misleading (Cochrane Comparing Multiple Interventions Methods Group, 2013).

The CMIMG (Cochrane Comparing Multiple Interventions Methods Group, 2013) suggests this simple reformatting of summary estimates into a table or figure should only be used if included trials in these systematic reviews meet the transitivity assumption: that two or more interventions can validly be compared using trials that did not compare them directly but which compared one or other of them against a common comparator (Cochrane Comparing Multiple Interventions Methods Group, 2013). If the trials do not meet this assumption, the CMIMG (Cochrane Comparing Multiple Interventions Methods Group, 2013) suggests overview authors present results of included systematic reviews on a review-by-review basis— despite their being no reporting guidelines for systematically presenting data from narrative reviews (Pieper *et al.*, 2014a) and no user testing on the method of presenting data in overviews (Smith *et al.*, 2011).

A second strategy for integrating the results of research syntheses is to employ a formal statistical analysis to make comparisons across interventions (Hartling *et al.*, 2014). These analyses can incorporate direct (head-to-head) comparisons, indirect comparisons, or both (i.e. network meta-analysis- NMA) (Thomson *et al.*, 2010). Although Ioannidis (2009) and Hartling *et al.* (2014) caution that merging results from pre-existing systematic reviews is limited by differences in eligibility criteria; outcome selection, reporting, and definition; methods of combining estimates (random vs. fixed effects models); approaches to missing data, updating, and overlap.

Nonetheless, a re-analysis of a published overview of treatments for childhood nocturnal enuresis cited by Thomson *et al.* (2010) used NMA to establish internally consistent estimates for all possible pairwise treatment comparisons, thereby correcting the results of the original overview which had been based on a summary table (Caldwell *et al.*, 2010).

While this proves that an NMA can be performed based on pooled summaries from systematic reviews, this shortcut is rarely available (Cochrane Comparing Multiple Interventions Methods Group, 2013). First, while the enuresis systematic reviews used the same two key outcomes, systematic reviews included in other overviews have not demonstrated the same consistency in their outcome selection (Wang *et al.*, 2012). Second, a strong form of the transitivity assumption is employed in indirect comparison (i.e.  $(A - B) = (A - C) + (C - B)$  and vice-versa) (Cochrane Comparing Multiple Interventions Methods Group, 2013). Even if included systematic reviews use the same outcome measures, appropriate use of NMA requires that authors ensure that the method's underlying assumptions have been met, which is best achieved by examining the primary trials in detail. As Cooper and Koenka (2012) point out, such an approach—or attempting to analyse systematic reviews and primary trials simultaneously—may not even be called an overview, as at that point overview authors would simply be using the research syntheses as a source of references for their new effort.

Another related issue concerns the methods used to explain heterogeneity (Conn and Coon Sells, 2014). While overviews are uniquely placed to address the issue of discordance, defined as conflicting results of systematic reviews on the same question, Pieper *et al.* (2014a) found that strategies to deal with discordant systematic reviews were reported in only 5% of overviews. This despite Jadad *et al.* (1997) having developed a relevant algorithm (summarized in Figure 3) nearly two decades ago (Conn and Coon Sells, 2014).

While a tool exists to improve exploration of heterogeneity in overviews, how to appropriately synthesize and report their findings remains debated.

## 4. Discussion

A scoping review was conducted in order to synthesise existing guidance concerning overview practice and delineate areas highlighted as outstanding challenges or where methodological recommendations conflict. Systematic searches identified 48 unique records, of which 24 met the inclusion criteria.

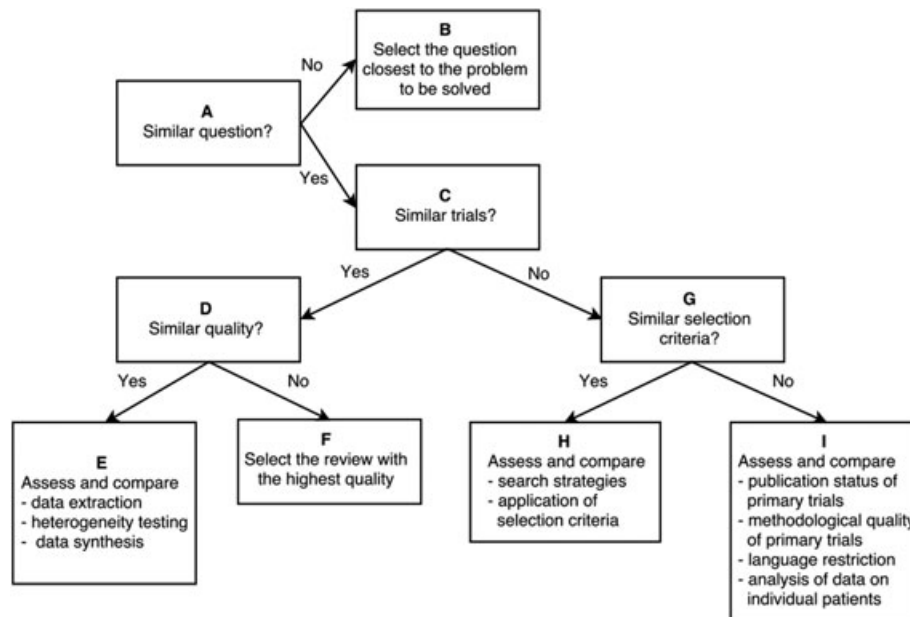
The scoping review demonstrates that, while many of the methodological standards for systematic reviews could also be applied to overviews, the problems of overlap, scope, evaluating the quality and reporting of included research, updating included systematic reviews, and how to synthesize and report findings remain unresolved.

These debates have implications for the potential risk of bias of prospective overviews. Given these outstanding issues, the question remains: when should overviews be undertaken? Caird *et al.* (2015) answer by arguing that overviews are a “solution” when (1) the systematic review addresses a broad research question, (2) results are required within a short timescale, and (3) there are not sufficient resources to conduct a full systematic review in the time available. While Caird *et al.* (2015) acknowledge the trade-off between comprehensiveness and lessened uncertainty and note that it may be “uncomfortable for reviewers,” they nonetheless conclude that overviews are a useful means of synthesizing evidence across broad topic areas (p. 95).

The question of when and how to make such trade-offs, however, requires additional specificity. Current guidance is heavily informed by consensus and personal experience rather than empirical evidence (Smith *et al.*, 2011). Enough is known, however, to differentiate between circumstances in which the efficiency of an overview does not compromise the validity of its outcome and those in which such expediency is likely to introduce serious bias in the form of overly precise estimates and the improper exclusion of relevant data. These conditions can be elucidated by clarifying a prospective overview’s intended purpose.

### 4.1. Overview functions

As explained above, three distinct, not entirely consistent functions have been proposed for overviews: (1) identify gaps in the literature where multiple comparable studies may exist but a research synthesis has not been performed (Caird *et al.*, 2015; Piso *et al.*, 2015; Ryan *et al.*, 2009; Cooper and Koenka, 2012; Santaguida *et al.*, 2013), (2) compare and contrast existing systematic reviews (Aromataris *et al.*, 2014; Aromataris *et al.*, 2015; Baker



**Figure 3.** Algorithm to identify and resolve causes of discordance between systematic reviews (reproduced from Jadad *et al.*, 1997)

*et al.*, 2014; Conn and Coon Sells, 2014; Cooper and Koenka, 2012; Pieper *et al.*, 2014a; Pieper *et al.*, 2012; Santaguida *et al.*, 2013; Smith *et al.*, 2011), and (3) provide a summary of evidence from existing systematic reviews, with or without synthesis (Becker and Oxman, 2011; Caird *et al.*, 2015; Cochrane Comparing Multiple Interventions Methods Group, 2012; Cochrane Comparing Multiple Interventions Methods Group, 2013; Cooper and Koenka, 2012; Hartling *et al.*, 2012; Hartling *et al.*, 2014; Li *et al.*, 2012; Pieper *et al.*, 2014a; Pieper *et al.*, 2012; Piso *et al.*, 2015; Ryan *et al.*, 2009; Smith *et al.*, 2011; Thomson *et al.*, 2013; Thomson *et al.*, 2010).

That these purposes are typically listed consecutively (Cooper and Koenka, 2012; Pieper *et al.*, 2012) belies the fact that a different methodological approach is required to achieve each one. This conflation of functions and equivocation of the term overview is the source of considerable methodological confusion.

The methodological conditions of each proposed aim will be considered in turn. A summary of this discussion is presented in Figure 4.

(i) Identify gaps in the literature

Contrary to the claim made by some included papers (Caird *et al.*, 2015; Piso *et al.*, 2015; Ryan *et al.*, 2009; Cooper and Koenka, 2012; Santaguida *et al.*, 2013), overviews that fail to find a systematic review for every relevant comparison will not, by default, detect evidence gaps. In contrast to empty systematic reviews, such overviews will not serve an agenda-setting function by directing future research and are likely to be misleading (Cochrane Empty Reviews Project, 2013). For example, consider a situation where there are three possible interventions (A, B, and C) for a particular condition. Although there exist primary trials evaluating each intervention, only intervention A and intervention B have been systematically reviewed (Figure 5).

Because conducting an overview does not involve searching for primary trials, an overview of this particular condition would include the systematic review of intervention A and the systematic review of intervention B, but no information on intervention C. While it is possible that authors well-versed in possible intervention options would note the absence of a systematic review for option C, such an observation would be ad hoc, rather than systematic. More critically, the resulting overview would still be incomplete and potentially misleading in that it does not reflect current evidence: although evidence about option C exists, the overview will not capture it.

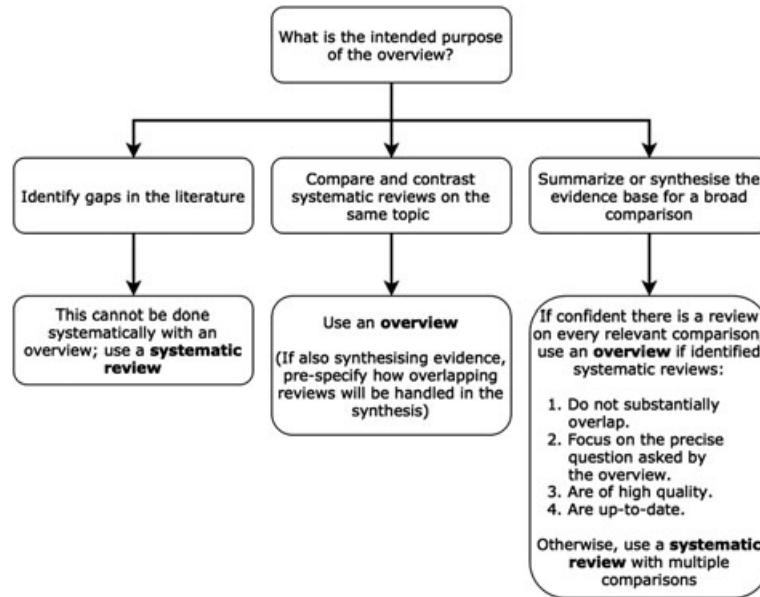


Figure 4. When to conduct an overview—a pilot decision tool

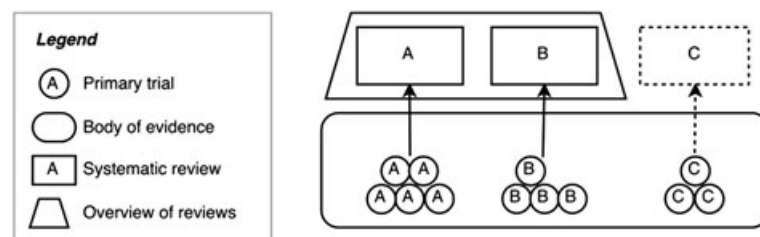


Figure 5. Overview in the absence of a systematic review on every relevant comparison

Identifying gaps in the literature is best done using a systematic review and should not be considered a function of overviews.

(ii) Compare and contrast existing systematic reviews

The second proposed function for overviews is to assess whether systematic reviews addressing similar questions independently report similar results and arrive at similar conclusions (Aromataris *et al.*, 2014; Aromataris *et al.*, 2015; Baker *et al.*, 2014; Conn and Coon Sells, 2014; Cooper and Koenka, 2012; Pieper *et al.*, 2014a; Pieper *et al.*, 2012; Santaguida *et al.*, 2013; Smith *et al.*, 2011). As noted above, this process can be guided by Jadad *et al.*'s (1997) algorithm to identify and resolve causes of discordance between systematic reviews.

While overviews are often framed as tools to facilitate broad comparisons (e.g. different interventions for the same condition, the same intervention for different conditions etc.), this function highlights that they can also be used to assess “differences between reviews on the *same* topic” (Baker *et al.*, 2014).

This aim is, however, in tension with the third proposed function summarising current evidence. When examining research syntheses focused on the same research problem or hypothesis, a high degree of overlap is likely. As explained above, when included systematic reviews are meta-analysed, using data from individual studies more than once without accounting for overlap will overestimate statistical power and thus risk producing a misleading, overly precise estimate (Pieper *et al.*, 2014c; Smith *et al.*, 2011). When systematic reviews are narratively summarized, a similar problem occurs if steps are not taken to ensure certain primary studies are not overrepresented (Thomson *et al.*, 2010; Pieper *et al.*, 2014c; Caird *et al.*, 2015).

While it is possible to explore heterogeneity and summarize results in the same overview (by, for example, pre-specifying how overlapping systematic reviews explored for heterogeneity will be handled in the synthesis), it is worth distinguishing between overviews intended to facilitate broad comparisons and those intended to resolve discrepancies between systematic reviews on the same research problem.

(iii) Summarise existing systematic reviews

Two separate but related ways of achieving the summary function have been proposed. In one, overviews serve as a “friendly front end,” compiling an exhaustive list of systematic reviews relevant to a specific decision and presenting results of included systematic reviews on a review-by-review basis (Cochrane Comparing Multiple Interventions Methods Group, 2013). In the second, overviews integrate or synthesize evidence from existing systematic reviews narratively and/or quantitatively.

Regardless of how the results are presented, the aim of such a summary or synthesis is to provide “an overall examination” of “the evidence base” available for a given topic (Aromataris *et al.*, 2014; Piso *et al.*, 2015). As explained above, however, overviews may fail to find a systematic review for every relevant comparison and in such cases will not accurately reflect the current evidence base. This raises the question of whether overviews should ever be conducted in the absence of such a search, and, if not, what would differentiate such an overview from a broad systematic review with appropriate subgroupings.

The ability of the overview to provide valuable decision support thus depends on the research question and realistic expectations towards the method. If authors of overviews are confident that a systematic review has been produced for all relevant comparisons they should consider an overview, with the understanding that the only way to test this assumption is to do a systematic search for primary trials.

When considering whether summarising a body of evidence using an overview is likely to produce useful and accurate results, this scoping review suggests that four further conditions ought to be met. After screening search results against eligibility criteria but before undertaking data extraction or analysis, prospective overviews should ensure identified systematic reviews:

1. Do not substantially overlap.
2. Focus on the precise question asked by the overview.
3. Are of high quality.
4. Are up-to-date.

While operationalizing this checklist would require undertaking the full process of guideline development (c.f. Grant *et al.*, 2013), an explanation of how each was derived from the literature and some suggested cut-offs are offered below. Whereas items 1 and 4 are mentioned in the Cochrane Comparing Multiple Interventions Methods Group (2013) editorial decision tree for overviews, this discussion adds further detail, offers two additional conditions, and contextualizes the checklist in terms of the intended purpose of the overview.

1. The systematic reviews do not substantially overlap

When multiple systematic reviews focus on the same question, there will likely be overlap. Yet there is currently only a way to measure the degree of overlap in an overview—the corrected covered area (CCA) (Pieper *et al.*, 2014c) —not a way to correct for it if it is found to be high. Problems arise if systematic reviews substantially overlap (i.e. get moderate, high, or very high CCA scores) because either, as suggested by Cooper and Koenka (2012), some systematic reviews are excluded—resulting in potential loss of valuable information



—or the CCA is noted but all systematic reviews are included—resulting in a misleading, overly precise estimate.

2. The systematic reviews focus on the precise question asked by the overview

Problems likewise arise with systematic reviews that do not focus on the precise question asked by the overview, but contain relevant trials. While the systematic reviews may sometimes run a separate meta-analysis with studies relevant to the overviews, they may sometimes lump all studies together. In the latter case, extracting which studies are relevant will likely require going back to trial level. When systematic reviews that do not focus on the precise question asked by the overview but contain relevant trials not found in other included systematic reviews are found, an overview may not be the best way of synthesizing the relevant evidence.

3. The systematic reviews are of high quality

Though a diversity of instruments to assess systematic review quality in recent overviews, there is a reliable and valid measure of systematic review quality (Moher *et al.*, 2010). The way in which overview findings should be presented, interpreted, and applied if systematic review quality is moderate or poor (i.e. most systematic reviews do not achieve a score of 8 or higher on AMSTAR), however, is less clear. While poor quality trials can be excluded from systematic reviews, poor quality systematic reviews may include high quality trials and so ought not to be excluded from overviews. That said, poor quality reviews may not be comprehensive or may have used inappropriate methods to combine study findings and so, when included, produce biased results.

Moreover, as explained above, there is currently no guidance on appropriate use and interpretation of the GRADE tool when assessing the quality of evidence based on systematic reviews rather than primary studies for which the tool was designed.

4. The systematic reviews are up-to-date.

If the systematic review is out of date, important evidence may be excluded. While the likelihood of a systematic review being out of date was calculated based on signals of the need to update (defined above) and not definitive judgments about actual changes in evidence, such signals are important for two reasons. As outlined Shojania *et al.* (2007), either a formal update that incorporates new evidence might yield conclusions that differ substantially from those of the previous systematic review or in the case that it does not, such an update would explicitly address the new evidence by placing it in context. The latter outcome is particularly important if, for example, a previous systematic review concluded that a treatment was effective and a trial in a high-impact journal concluded that the treatment had no benefit. Because practitioners might preferentially act on the trial's conclusions because it was more recent or appeared in a high-impact journal, it would be important to reassert the findings of the original systematic review. Unfortunately, there is currently little empirical evidence on when and how to conduct updates of systematic reviews in the context of overviews; as such, if a substantial portion of the systematic reviews appear to be out of date, it may be misleading to proceed with the overview.

Given that the “speed” with which overviews can be conducted is cited as one of their primary advantages (Aromataris *et al.*, 2014; Caird *et al.*, 2015) it should not be onerous to assess these conditions after conducting searches but before full data extraction and analysis. An example of how these conditions might be applied is presented in Appendix C of the supporting file. In the event that the identified systematic reviews do not meet these criteria the overview should not proceed and a broad systematic review with appropriate subgroupings—in which identified systematic reviews could be used as a source of primary trials—ought to be conducted instead.

#### 4.2. Strengths and weaknesses

Given the rising number of overviews produced (Pieper *et al.*, 2012) and their role as a highly-accessed tool for decision support (Thomson, 2014), addressing risk of bias in this publication type is a priority. This paper is the first to use rigorous, systematic methods to comprehensively summarise current methodological guidance for overviews and delineate areas highlighted as outstanding challenges or where methodological recommendations conflict. It is also the first to consider these methodological debates in terms of the proposed functions of overviews. The result is a decision tool that unpicks the methodological implications of these functions, demonstrating the ways in which they may be unachievable, in tension with each other, or possible only under certain conditions. This tool provides clarity on when an overview ought to be considered and its conditions can quickly be applied before data extraction and synthesis to ascertain whether a proposed overview ought to proceed.

That said, this work has some limitations that should be taken into account. This tool was constructed solely on the basis of existing literature. Subjecting this decision tool to revision through a consensus process involving experts in the field of overviews, together with producers and consumers of overviews could (1) help identify important methodological issues not mentioned in the literature (e.g. aggregation bias) (Pieper *et al.*, 2014b), (2) refine the conditions and cutoffs, and (3) ensure its relevance to and uptake by potential authors of overviews and those who commission them (Grant *et al.*, 2013; Moher *et al.*, 2010).

## 5. Conclusion

This scoping review highlights the lack of guidance regarding many critical facets of overview conduct and delineates five emerging debates. These debates have implications for the potential risk of bias of prospective overviews and can be elucidated by clarifying the intended purpose of overviews. While methods papers routinely list three functions for overviews—identify gaps, explore heterogeneity, summarize evidence—overviews can only perform the latter two. They are unlikely to achieve the second and third simultaneously, and can only perform the third with a low risk of bias when four conditions are met.

Overviews are always susceptible to the interpretive, conceptual, and reporting biases of previous reviewers (Hartling *et al.*, 2014); however, when the four conditions are not met, the overview outcomes risk being biased and misleading. In these cases, a broad systematic review with multiple comparisons would likely produce a more precise result than an overview. Though there may be substantial heterogeneity in the included studies of broad systematic reviews—a reflection of the diversity of issues considered by the overview and the myriad ways they have been implemented—this can be addressed by careful grouping of studies and selecting for meta-analysis only those subgroups in which the intervention, context and outcomes were sufficiently similar. Such a broad systematic review, while more time consuming than an overview, serves the same essential purpose—summarizing a broad evidence base and/or indicating the relative efficacy of different interventions—in a way that is less likely to produce a biased outcome.

While future work is required to fully develop consensus-based guidelines (Moher *et al.*, 2010) these preliminary thoughts are offered for consideration by the field. Consideration of the intended function of proposed overviews along with the corresponding methodological conditions may improve the quality of this burgeoning publication type.

## Acknowledgements

We thank Janina Steinert and Juliana Gaertner for their assistance with German translation and Rachel Woodlee for her assistance with the Mandarin translation. MB was supported by The Rhodes Trust.

## References

- Aromataris E, Fernandez R, Godfrey C, Holly C, Khalil H, Tungpunkom P 2014. Methodology for JBI umbrella reviews. *The Joanna Briggs Institute Reviewers Manual* (Suppl): 5–34.
- Aromataris E, Fernandez R, Godfrey CM, Holly C, Khalil H, Tungpunkom P 2015. Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. *International Journal of Evidence-Based Healthcare* **13**: 132–40.
- Baker PR, Costello JT, Dobbins M, Waters EB 2014. The benefits and challenges of conducting an overview of systematic reviews in public health: a focus on physical activity. *Journal of Public Health* **36**: 517–21.
- Becker L, Oxman A 2011. Chapter 22: Overviews of reviews. In Higgins J, Green S (eds.). *Cochrane Handbook for Systematic Reviews of Interventions* www.cochrane-handbook.org. The Cochrane Collaboration.
- Caird J, Sutcliffe K, Kwan I, Dickson K, Thomas J 2015. Mediating policy-relevant evidence at speed: are systematic reviews of systematic reviews a useful approach? *Evidence and Policy* **11**: 81–97.
- Caldwell DM, Welton NJ, Ades AE 2010. Mixed treatment comparison analysis provides internally coherent treatment effect estimates based on overviews of reviews and can reveal inconsistency. *Journal of Clinical Epidemiology* **63**: 875–82.
- Cochrane Comparing Multiple Interventions Methods Group. 2012 Review type & methodological considerations: background paper for the first part of the Paris CMIMG discussion. Paris CMIMG Discussion, Paris, France.
- Cochrane Comparing Multiple Interventions Methods Group 2013. Undertaking, publishing and maintaining Cochrane Reviews that compare multiple interventions.
- Cochrane Empty Reviews Project. 2013. The Cochrane empty reviews project [online]. Available from: <https://empty-reviews.org> [Accessed May 10 2016].
- Conn VS, Coon Sells TG 2014. WJNR welcomes umbrella reviews. *Western Journal of Nursing Research* **36**: 147–151.
- Cooper H, Koenka AC 2012. The overview of reviews: unique challenges and opportunities when research syntheses are the principal elements of new integrative scholarship. *The American Psychologist* **67**: 446–462.
- Flodgren G, Eccles Martin P, Shepperd S, Scott A, Parmelli E, Beyer Fiona R 2011. An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes. *Cochrane Database of Systematic Reviews*[Online]. Available from: <http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD009255/abstract>.
- Grant SP, Mayo-Wilson E, Melendez-Torres G, Montgomery P 2013. Reporting quality of social and psychological intervention trials: a systematic review of reporting guidelines and trial publications. *PLoS One* **8**: e65442.



- Hartling L, Chisholm A, Thomson D, Dryden DM 2012. A descriptive analysis of overviews of reviews published between 2000 and 2011. *PLoS ONE* **7**: e49667.
- Hartling L, Vandermeer B, Fernandes RM 2014. Systematic reviews, overviews of reviews and comparative effectiveness reviews: a discussion of approaches to knowledge synthesis. *Evidence-Based Child Health a Cochrane Review Journal* **9**: 486–94.
- Higgins JPT, Green S 2011. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. Available from: [http://handbook.cochrane.org/0\\_2\\_how\\_to\\_cite\\_the\\_handbook.htm](http://handbook.cochrane.org/0_2_how_to_cite_the_handbook.htm)
- Hill S, Ryan R 2014. Cochrane review—interventions to improve safe and effective medicines use by consumers: an overview of systematic review. *Journal of Evidence-Based Medicine* **7**: 231–231.
- Ioannidis JP 2009. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. *CMAJ* **181**: 488–93.
- Jadad AR, Cook DJ, Browman GP 1997. A guide to interpreting discordant systematic reviews. *CMAJ* **156**: 1411–6.
- Jones L, Othman M, Dowswell T, Alfirevic Z, Gates S, Newburn M, Jordan S, Lavender T, Neilson James P 2012. Pain management for women in labour: an overview of systematic reviews. *Cochrane Database of Systematic Reviews* [Online]. Available from: <http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD009234.pub2/abstract>
- Jørgensen AW, Hilden J, Gøtzsche PC 2006. Cochrane reviews compared with industry supported meta-analyses and other meta-analyses of the same drugs: systematic review. *BMJ* **333**: 782.
- Lamming L, Mason D, Morton K, Pears S, Bijker M, Sutton S, Hardeman W 2014. What do we know about brief interventions to promote physical activity? Systematic review of reviews. *European Health Psychologist* **16**: 808.
- Li L, Tian J, Tian H, Sun R, Liu Y, Yang K 2012. Quality and transparency of overviews of systematic reviews. *Journal of Evidence-Based Medicine* **5**: 166–173.
- Martínez-González NA, Berchtold P, Ullman K, Busato A, Egger M 2014. Integrated care programmes for adults with chronic conditions: a meta-review. *International Journal for Quality in Health Care* **26**: 561–570.
- Meltzer H 1972. Review of reviews in industrial psychology, 1950–1969. *Personnel Psychology* **25**: 201–222.
- Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF, Group Q 1999. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *The Lancet* **354**: 1896–1900.
- Moher D, Schulz KF, Simera I, Altman DG 2010. Guidance for developers of health research reporting guidelines. *PLoS Medicine* **7**: e1000217.
- Moseley AM, Elkins MR, Herbert RD, Maher CG, Sherrington C 2009. Cochrane reviews used more rigorous methods than non-Cochrane reviews: survey of systematic reviews in physiotherapy. *Journal of Clinical Epidemiology* **62**: 1021–1030.
- O'donnell A, Anderson P, Newbury-Birch D, Schulte B, Schmidt C, Reimer J, Kaner E 2014. The impact of brief alcohol interventions in primary healthcare: a systematic review of reviews. *Alcohol and Alcoholism* **49**: 66–78.
- Olsen O, Middleton P, Ezzo J, Gøtzsche PC, Hadhazy V, Herxheimer A, Kleijnen J, McIntosh H 2001. Quality of Cochrane reviews: assessment of sample from 1998. *BMJ* **323**: 829–832.
- Oxman AD, Guyatt GH 1991. Validation of an index of the quality of review articles. *Journal of Clinical Epidemiology* **44**: 1271–8.
- Pieper D, Antoine S-L, Mathes T, Neugebauer EAM, Eikermann M 2014a. Systematic review finds overlapping reviews were not mentioned in every other overview. *Journal of Clinical Epidemiology* **67**: 368–375.
- Pieper D, Antoine SL, Morfeld JC, Mathes T, Eikermann M 2014b. Methodological approaches in conducting overviews: current state in HTA agencies. *Research Synthesis Methods* **5**: 187–99.
- Pieper D, Antoine SL, Neugebauer EAM, Eikermann M 2014c. Up-to-dateness of reviews is often neglected in overviews: a systematic review. *Journal of Clinical Epidemiology* **67**: 1302–1308.
- Pieper D, Buechter R, Jerinic P, Eikermann M 2012. Overviews of reviews often have limited rigor: a systematic review. *Journal of Clinical Epidemiology* **65**: 1267–73.
- Piso B, Semlitsch T, Reinsperger I, Breuer J, Kaminski-Hartenthaler A, Kien C, Thaler K, Siebenhofer A 2015. Practical experience with overviews of reviews—valuable decision aid or academic exercise? *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen* **109**: 300–8.
- Ryan RE, Kaufman CA, Hill SJ 2009. Building blocks for meta-synthesis: data integration tables for summarising, mapping, and synthesising evidence on interventions for communicating with health consumers. *BMC Medical Research Methodology* **9**: 16.
- Santaguida PL, Keshavarz H, Carlesso LC, Lomotan M, Gross A, Macdermid JC, Walton DM, Group IW 2013. A description of the methodology used in an overview of reviews to evaluate evidence on the treatment, harms, diagnosis/classification, prognosis and outcomes used in the management of neck pain. *The Open Orthopaedics Journal* **7**: 461–72.
- Shea B, Moher D, Graham I, Pham B, Tugwell P 2002. A comparison of the quality of Cochrane reviews and systematic reviews published in paper-based journals. *Evaluation & the Health Professions* **25**: 116–129.
- Shea BJ, Bouter LM, Peterson J, Boers M, Andersson N, Ortiz Z, Ramsay T, Bai A, Shukla VK, Grimshaw JM 2007. External validation of a measurement tool to assess systematic reviews (AMSTAR). *PLoS ONE* **2**: e1350.

- Shojania K, Sampson MJ, Ansari MT, Ji J, Doucette S, Moher D 2007. How quickly do systematic reviews go out of date? A survival analysis. *Annals of Internal Medicine* **147**: 224–233.
- Singh Jasvinder A, Wells George A, Christensen R, Tanjong Ghogomu E, Maxwell L, Macdonald John K, Filippini G, Skoetz N, Francis Damian K, Lopes Luciane C, Guyatt Gordon H, Schmitt J, La Mantia L, Weberschock T, Roos Juliana F, Siebert H, Hershan S, Cameron C, Lunn Michael PT, Tugwell P, Buchbinder R 2011. Adverse effects of biologics: a network meta-analysis and Cochrane overview. *Cochrane Database of Systematic Reviews*[Online]. Available from: <http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD008794.pub2/abstract> .
- Smith V, Devane D, Begley CM, Clarke M 2011. Methodology in conducting a systematic review of systematic reviews of healthcare interventions. *BMC Medical Research Methodology* **11**: 15.
- Stultjens EM, Dekker J, Bouter LM, Leemrijse CJ, Van Den Ende CH 2005. Evidence of the efficacy of occupational therapy in different conditions: an overview of systematic reviews. *Clinical Rehabilitation* **19**: 247–254.
- Theodoratou E, Tzoulaki I, Zgaga L, Ioannidis JP 2014. Vitamin D and multiple health outcomes: umbrella review of systematic reviews and meta-analyses of observational studies and randomised trials. *BMJ* **348**: g2035.
- Thomson D 2014. Evidence synthesis in child health: overviews of reviews. *Evidence-Based Child Health* **9**: 1–2.
- Thomson D, Foisy M, Oleszczuk M, Wingert A, Chisholm A, Hartling L 2013. Overview of reviews in child health: evidence synthesis and the knowledge base for a specific population. *Evidence-Based Child Health: A Cochrane Review Journal* **8**: 3–10.
- Thomson D, Russell K, Becker L, Klassen T, Hartling L 2010. The evolution of a new publication type: steps and challenges of producing overviews of reviews. *Research Synthesis Methods* **1**: 198–211.
- Wang, X., Lindsley, K. & Li, T. 2012. Is there agreement in outcomes among Cochrane reviews to support 'Overviews' of reviews? A case study within the Cochrane Eyes and Vision Group (CEVG). The Cochrane Library. 2012 Supplement.
- Yuan JQ, Liu YL, Yang KH, Chen YL, Yang M 2012. Survey of quality assessment methods of included studies in overviews of reviews. *Chinese Journal of Evidence-Based Medicine* **12**: 238–242.

## Supporting information

Additional supporting information may be found in the online version of this paper at the publisher's website.