ELSEVIER

Journal of Clinical Epidemiology 67 (2014) 368–375

Journal of Clinical Epidemiology

REVIEW ARTICLES

Systematic review finds overlapping reviews were not mentioned in every other overview

Dawid Pieper*, Sunya-Lee Antoine, Tim Mathes, Edmund A.M. Neugebauer, Michaela Eikermann

Institute for Research in Operative Medicine (IFOM), Witten/Herdecke University, Ostmerheimer Str. 200, Building 38, D-51109, Cologne, Germany

Accepted 11 November 2013

Abstract

Objectives: The objective of this study was to determine if the authors mention overlapping reviews in overviews (reviews of reviews). In addition, we aimed to calculate the actual overlap in published overviews using newly introduced, validated measures.

Study Design and Settings: We systematically searched for overviews from 2009 to 2011. Reviews included in the overviews were obtained. Tables (review × primary publication) were generated for each overview. The first occurrence of a primary publication is defined as the index publication. We calculated the ''corrected covered area'' (CCA) as a measure of overlap by dividing the frequency of repeated occurrences of the index publication in other reviews by the product of index publications and reviews, reduced by the number of index publications. Subgroup analyses were performed to investigate further differences in the overviews.

Results: Only 32 of 60 overviews mentioned overlaps. The median CCA was 4.0. Validation of the CCA and other overlap measures was in accordance with our predefined hypotheses. The degree of overlap tended to be higher in health technology assessment reports than in journal publications and was higher with increasing numbers of publications.

Conclusions: Overlaps must be reported in well-conducted overviews, and this can comprehensively be accomplished using the CCA method.

Keywords: Systematic review; Methods; Research design; Meta-analysis; Evidence-based medicine; Information science

1. Introduction

Overviews (reviews of reviews), as a new type of evidence synthesis, have recently gained more interest, such that the number of published overviews is steadily increasing [1]. It is possible that overviews are becoming more prevalent because overviews have potential advantages over systematic reviews (SRs). For example, overviews enable data obtained from different interventions or conditions to be compared, which provides decision makers with a broader summary of the current information available. This is a limitation of SRs, which may be overcome by using overviews [2]. Furthermore, overviews can compare the findings of several reviews and determine the reasons for conflicting results. By identifying the reasons for discordance, users are able to base their decisions on the most current, reliable, and suitable data for their situation [3,4].

It has been stated that many of the methodological standards for SRs can also be applied to overviews [5]. However, little guidance is available for authors on how to conduct methodologically sound overviews. Interestingly, a descriptive analysis concluded that overviews often have limited rigor [4].

Decisions in health care should be based on all of the available evidence to draw reliable conclusions and to support policy making. Therefore, we most often rely on the SRs [6]. When conducting an overview, one might argue that a decision should be based on an enormous body of evidence. Even if this holds to be true, it may be difficult and challenging to survey all of the available evidence that is gained from primary studies mainly because they are often included in more than one review. Additionally, a meta-analysis of meta-analyses may also be difficult to conduct because many of the primary studies will usually be included in more than one meta-analysis. Therefore,

The authors of this article received no financial support for this research and/or authorship of this article.

All of the authors declare that they have no conflicts of interest to disclose.

* Corresponding author. Evidence-Based Health Services Research, Institute for Research in Operative Medicine, Witten/Herdecke University, Ostmerheimer Str. 200, Building 38, D-51109 Cologne, Germany. Tel.: +49-221-989-5742; fax: +49-221-989-5730.

E-mail address: dawid.pieper@uni-wh.de (D. Pieper).

0895-4356/$ - see front matter
http://dx.doi.org/10.1016/j.jclinepi.2013.11.007

What is new?

Key findings

- Almost half of the overviews (53.3%) mentioned overlaps, whereas the remaining overviews did not.

What this adds to what was known?

- This is the first systematic analysis of overlaps in reviews of reviews.
- Development and validation of a measure (corrected covered area [CCA]) to calculate the actual degree of overlap in overviews.

What is the implication and what should change now?

- Insufficient reporting of the quality of systematic reviews complicates the production of overviews, in particular with respect to overlaps.
- Overlaps must be reported in well-conducted overviews, and this can comprehensively be done using the CCA method.

pooling the results of all of the reviews would give disproportionate statistical power to multiple primary studies [2]. An informal analysis that sums the results of the reviews could also introduce significant overlap and result in many primary studies being included more than once, which would lead to biased results. This problem should be addressed by developing standard methods for authors of overviews to follow [7].

However, to the best of our knowledge, the degree of overlap in overviews has not been examined systematically. Thus, the amount of overlapping data in overviews remains unknown. First, we aimed to determine whether authors mentioned overlaps in their overviews, and if so, we examined how the authors dealt with these overlaps. Second, we examined the actual overlap in published overviews and suggested potential measures for handling these overlaps.

2. Methods

2.1. Literature search

In November 2011, we performed a systematic search for overviews of reviews. An update was performed in May 2012. We searched the databases of Medline, Embase, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Physiotherapy Evidence Database (PEDro), and all of the databases of the Cochrane library. Because no subsets, subject headings, or search filters for overviews were available thus far, we searched the databases using text words. We adapted this search strategy based on prior work from our group [4]. Furthermore, we hand searched the web sites of health technology assessment (HTA) agencies for published reports. A list of HTA agencies was derived by investigating member lists from the International Network of Agencies for Health Technology Assessment, Health Technology Assessment international, the European network for Health Technology Assessment, and supplemental HTA agencies. Overall, we searched the web sites of 127 HTA agencies.

The search was restricted to articles published in English or German. All reviews had to be published between 2009 and 2011. The study protocols for overviews such as those published by the Cochrane Collaboration were excluded. If updates were published, only the most recent publication was included. Two members of the research team screened all titles and abstracts independently. The full texts of potentially eligible articles were then obtained including relevant supplements or appendices. Two reviewers assessed the eligibility of the full texts against the review inclusion criteria. Any disagreements were resolved by discussion.

2.2. Selection

For the inclusion and the analysis of overlaps, we developed a two-step approach. In the first step, we included all overviews that synthesized reviews on the same or a similar topic and/or intervention that were derived through a systematic literature search. To be included, the authors had to name at least one database and explicitly state that they searched for reviews. We did not exclude overviews that also included primary studies. For inclusion, the evidence synthesis had to rely at least in part on reviews (eg, combining primary studies and reviews in evidence synthesis). In addition, all of the literature that was included (either secondary or primary research) must have been critically appraised. We excluded all overviews with a methodological focus (eg, reviews dealing with the reporting characteristics or the quality of SRs in a specific field) and those that included clinical practice guidelines (CPGs).

2.3. Data collection

In the second step, we obtained reviews that were included for evidence synthesis in the overviews based on the reference list. We excluded all overviews that had inadequately reported the data. This problem may also occur in reviews of primary research. Irrespective of whether the problem arose in an overview or review, we defined a cutoff point for each overview such that not more than 10% of the reviews (or the primary studies included in it) would be missing. There were many reasons for missing reviews, such as the full text was not obtainable, language restrictions (in addition to German and English, we also considered reviews in French), or bad reporting resulting. If there were any doubts or inconsistencies, we made every

attempt to contact the authors of the overviews or reviews for clarification.

Data were extracted by one reviewer in structured summary tables. A second reviewer verified the accuracy of the results. We did not extract any information regarding the results. Any disagreements were resolved by consensus. For each overview, we obtained the following information:

- Name of the first author,
- Publication year,
- Publication type (journal or HTA report),
- Objective,
- Number of included reviews,
- Number of included primary studies (if any), and
- If and how the authors mentioned overlaps,

2.4. Analysis

The HTA reports may include more than one question for a given topic. We treated each objective in an HTA report or a journal publication containing a literature search as a unique overview. Thus, the unit of analysis was not necessarily the whole publication.

We analyzed the data descriptively as frequencies or medians and interquartile ranges or ranges. We generated citation matrices that cross-linked individual reviews (column) with all of the primary publications that were included (row). If one publication was included in more than one review, there was only one row for that publication, and we included a tick mark in the where this publication was included (Fig. 1). For primary publications, we relied on the citations given by the authors of the review after accounting for any misspellings or transposed digits. Citation matrices were generated by one reviewer and checked by a second reviewer for accuracy.

To date, there are no measures available for displaying the degree of overlap. We presumed that a single measure would be easier to understand than citation matrices and this would also allow for interoverview comparisons. Furthermore, it can provide useful information for developers and consumers of overviews. It might be a simpler approach to calculate the percentage of primary publications included in more than one review in the overview (abbreviated as "% overlaps"), and we calculated the "% overlaps" for each overview in this way. Although intuitively easy to understand, this measure will not account for primary publications included in more than two reviews, as a primary publication can only be counted once. This is an important difference when compared with our newly developed measures. Furthermore, for each overview, we calculated the measure of the "covered area" (CA) by dividing the sum of the included publications (irrespective of overlaps) by the product of the rows and columns. This measure increases with each primary publication included in more than one review as it adjoins a ticked box while not changing the number of rows and columns equally, which holds the area constant. A potential drawback of the CA measure is that it is capable of being influenced by a single review containing a large number of primary publications when compared with other included reviews. These types of reviews do add area but not necessarily additional overlaps. Imagine a hypothetical overview with four included reviews. Three reviews included exactly the same five studies. The fourth review included 50 primary publications, and none of these studies were included in the other three overlapping reviews. In this case, the CA would be 29.5% (65 [5 + 5 + 5 + 50] primary publications divided by 220 [4 columns × 55 rows]), whereas it would have been 100% if the fourth review was eliminated. The CA decreases to 27.4, 26.2, and 25.2 if the fourth review would have included 100; 200; and 1,000 primary publications, respectively. Based on this it can be concluded that the CA may decrease slightly by simply adding primary publications from a huge review. A CA of 25.2 could still be interpreted highly overlapping, although the fourth review adds much more data to the analysis. Another drawback of the CA is that it has an indefinite range as the numerator will never be zero. Imagine two reviews each with five different studies. The CA would be 50% (10/20), although no overlap is present (Fig. 2).

To take this into consideration, we introduced a measure of "corrected covered area" (CCA). The first occurrence of a primary publication is defined as the index publication (index case). We calculated the CCA as a measure of overlap by dividing the frequency of repeated occurrences of the index publication in other reviews by the product of index publications and reviews, and this product is reduced by the number of index publications. The reduction of the denominator by the number of index publications (rows) results in a possible range of 0–100% for the CCA for each review. The numerator constitutes the real overlap as it only

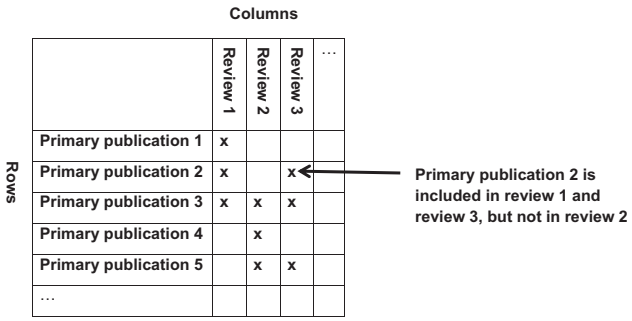

Fig. 1. Citation matrix.

$$CA\ (covered\ area) = \frac{N}{r\,c}$$

$$CCA\ (Corrected\ CA) = \frac{N-r}{r\,c-r}$$

where N is the number of included publications (including double counting) in evidence synthesis (this is the sum of the ticked boxes in the citation matrix); where r is the number of rows (number of index publications) and c is the number of columns (number of reviews).

Fig. 2. Calculation formulas. CA, covered area; CCA, corrected covered area.

counts primary publications included in more than one review and diminishes the impact of large reviews. Using our aforementioned example, the CCA would be 6.1% ([65 − 50 − 5]/[220 − 50 − 5]) in the baseline scenario. This can be interpreted as the area that is covered after eliminating the inclusion of all primary publications the first time they are counted. In other words, the CCA is the degree of overlap. In the case that each review include exactly all same studies, the CCA would be 100%, indicating a complete overlap. The CCA decreases faster than the CA because the numerator is constant. After repeating the sensitivity analysis previously mentioned, the CCA would be 3.2, 1.6, and 0.3 for 100; 200; and 1,000 primary publications, respectively. A further example is provided in Fig. 3. A CCA value lower than 5 can be considered as a slight overlap, whereas values greater than or equal to 15 can be considered as a very high overlap (Table 1).

Correlation coefficients (Kendall τ b) were used for hypotheses testing. We hypothesized that the CA should have a strong (0.60–0.80) negative correlation with the number of included reviews and, compared to this, a lower negative correlation with the number of included primary publications. In contrast, we assumed that the CCA should have a very weak (0.00–0.20) or weak (0.20–0.40) negative correlation with the number of included reviews and the primary publications.

Table 1. Interpretation of CCA

CCA	Overlap
0–5	Slight
6–10	Moderate
11–15	High
>15	Very high

Abbreviation: CCA, corrected covered area.

Subgroup analyses were performed for the number of reviews, the number of primary publications, the overlap %, the CA, and the CCA using the Mann–Whitney *U* statistic. A Pearson χ^2 test was used for the comparison of overviews mentioning overlaps vs. publication type. Significance was assumed as *P*-value lower than 0.05.

All calculations and statistical analyses were performed using Microsoft Excel 2010 (Microsoft Corporation, Richmond, WA) and SPSS version 20.0.0 (SPSS, Inc., Chicago, IL).

	Review 1	Review 2	Review 3	Review 4
Primary publication 1	x			
Primary publication 2	x			x
Primary publication 3	x	x	x	x
Primary publication 4	x	x	x	
Primary publication 5	x	x		x
Primary publication 6	x			
Primary publication 7	x	x	x	
Primary publication 8	x			
Primary publication 9		x		
Primary publication 10		x		
Primary publication 11				x
Primary publication 12				x

% Overlap: 5/12=41.7

CA: (8 + 6 + 3 + 5)/(4 × 12) = 22/48 = 0.46

CCA: (22 - 12)/(48 − 12) = 10/36 = 0.28

Fig. 3. Illustrative example for the calculation of the overlap measures. CA, covered area; CCA, corrected covered area.

3. Results

3.1. Search results

We identified 1,993 potentially relevant studies. After checking the titles and abstracts, we retrieved 273 full-text articles for further consideration. Of these, 200 were excluded for various reasons, namely 44 did not explicitly search systematically for reviews and named at least one database, 100 did not critically assess the included literature, 6 did not use the reviews in evidence synthesis, 4 had a methodological focus, 10 were neither written in English nor in German, 3 were duplicates, and the full text was not available for 33 overviews primarily because they were protocols. Thus, 73 overviews were eligible for analysis. In the second step, 26 of these overviews were excluded for various reasons, namely 2 had an irreproducible evidence synthesis (eg, inconsistent data), more than 10% of included reviews were not obtainable in 13 overviews, and the primary studies were not identifiable (>10% of reviews) in 11 overviews. A search of the HTA agencies web sites resulted in 13 overviews that were included in this analysis. Thus, altogether 60 overviews (58 publications) were eligible for analysis (Fig. 4).

3.2. Descriptive characteristics

Of the 60 overviews that were eligible for analysis, 45 (75%) overviews were published in journals and 15 (25%) were part of an HTA report. Two HTA reports contained two overviews each. The overviews included a median of 6.5 reviews (range, 2–45) and 93 primary studies (range, 11–639), not accounting for primary publications that were also searched. Almost half of the overviews (53.3%) mentioned overlaps. In most cases, the authors provided

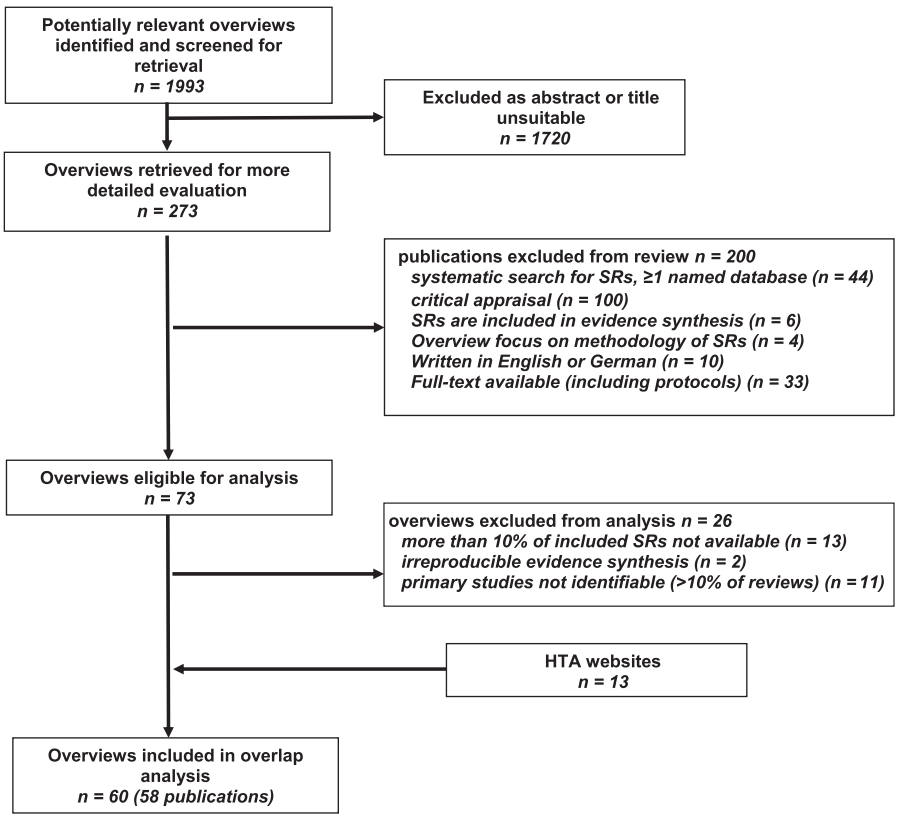

Fig. 4. Flow chart. SR, systematic review; HTA, health technology assessment.

either a citation matrix ($n = 13$) or described the number of studies that were included in one, two, three, four, and so on reviews ($n = 11$). One overview illustrated the extent of overlaps by using Venn diagrams, whereas another one calculated the degree of overlaps but failed to report the underlying methodology. Six overviews simply mentioned the problem of overlaps without further investigation. A median of 25% (range, 0.0–66.7%) primary publications was included in more than one review in the overviews. The overviews had a median of 22.0 CA (range, 3.3–85.2) and 4.0 CCA (range, 0.0–70.4). Descriptive characteristics of the overviews are summarized in Table 2.

3.3. Validation

Both of our measures for evaluating overlaps, CA and CCA, showed a moderate correlation (Kendall τ b: 0.520, $P < 0.01$). The CA revealed a very strong correlation with the number of included reviews (Kendall τ b: −0.814, $P < 0.01$), whereas this correlation was weak for the CCA (Kendall τ b −0.323, $P < 0.01$). The correlation coefficients converged more when both measures were compared with the number of primary publications. The percentage of primary publications that were included in more than one review showed a strong correlation with the CCA (Kendall τ b: 0.643, $P < 0.01$) but only a very weak correlation with the CA (Kendall τ b 0.192, $P < 0.05$). The correlation coefficients are summarized in Table 3.

3.4. Differences in mentioning overlaps

Almost half of the overviews did not mention overlaps. Overviews that did not mention overlaps included more reviews and primary publications than overviews that did mention overlaps, although this difference was not statistically significant. Our overlap measures (CA, CCA, and % overlaps) tended to be slightly higher in overviews mentioning overlaps, but only the CCA measure was statistically significant ($P = 0.037$), although the means were nearly equal. A description of the overviews and whether they mentioned overlaps is summarized in Table 4.

3.5. Differences by publication type

The descriptive characteristics differed widely by publication type (journal publication vs. HTA report). However, the publication type did not influence whether or not the authors mentioned overlaps. In journal publications and HTA reports, 23 of 45 (51.1%) and 9 of 15 (60.0%) mentioned overlaps ($P = 0.550$), respectively. There were significant differences in the mentioning of overlaps depending on the number of included reviews or primary publications. Journal publications tended to include double the research findings of

Table 2. Descriptive characteristics

Category	Number
Year of publication, *n*	
2011	23
2010	23
2009	14
Publication type, *n* (%)	
Journal	45 (75)
HTA report	15 (25)
Number of included reviews, median (IQR)	6.5 (4.75–12.5)
Number of primary publications in reviews, median (IQR)	93 (42.25–191)
Additional searches for primary publications, *n* (%)	25 (41.7)
Dealt with overlaps, *n* (%)	32 (53.3)
% Primary publications included in more than one review, median (IQR)	25.0 (11.8–35.7)
CA, median (IQR)	22.0 (10.5–30.4)
CCA, median (IQR)	4.0 (1.5–11.4)

Abbreviations: HTA, health technology assessment report; IQR, interquartile range; CA, covered area; CCA, corrected covered area.

HTA reports. All overlap measures (CA, CCA, and % overlaps) were higher in HTA reports than in journal publications, but only the CA measure reached statistical significance. Descriptions of the characteristics of the overviews by publication type are summarized in Table 3.

4. Discussion

There is a substantial overlap of reviews being included in overviews. However, only half of all authors of overviews acknowledge this problem. The manner in which overlaps are dealt with in the literature does not depend on the degree of overlap. The CCA measure showed a statistically significant difference in the degree of overlap, but the absolute difference was small. The degree of overlap was higher in HTA reports than in journal publications. This is perhaps because the purpose of HTA reports being different from journal publications. The HTA reports intend to support the process of decision making in health care at the policy level [8]. Therefore, the HTA reports have a more narrow research question, which results in a smaller number of relevant and identifiable reviews and an increase in the amount of overlap.

A high degree of overlaps more than likely reflects an unnecessary duplication of reviews. The SRs should only be performed in the case of an out-of-date review or a significantly different research objective [9]. Prospective registration of SRs might be helpful for avoiding unnecessary duplications [10].

Despite being aware of the degree of an overlap, the task of synthesizing reviews in overviews remains challenging. This challenge results from having to rely on already-existing and highly aggregated data in reviews, whereas it is easier to modify more parameters, mainly the inclusion criteria, when conducting evidence synthesis based on primary studies.

We introduced and validated two new measures (CA and CCA) of overlap. Our hypotheses were confirmed by our analysis of correlation coefficients. There is only a moderate correlation between the two measures, which indicates that they measure two different concepts, although they cannot be independent by definition. The CCA measure depends more on the number of primary publications than on the number of reviews. However, the correlations only range from weak to moderate. In contrast, the CA is more dependent on the number of included reviews than the number of primary publications, which results in much higher correlations when compared with the CCA.

The method of calculating the CCA will always be a citation matrix, which could potentially become time consuming depending on the number of reviews and primary publications included in the overview. Citation matrices contain all of the information to calculate the CCA and are often used to illustrate overlaps. However, they require a lot of space that is not always available in scientific journals. Furthermore, from the reader's perspective, it remains unclear how useful citation matrices are, because it may be difficult to keep track of them.

Because our inclusion criteria were stricter than former SRs of overviews [1,4], we only analyzed 60 overviews. We were particularly strict when including only literature that was previously critically appraised. By doing so, we actually included only "systematic overviews," although this term has been rare in prior research. Thus, we had to exclude 26 overviews because of the inability to identify the included studies in SR, irreproducible evidence synthesis, or the inability to obtain the included reviews. Had we included these 26 overviews, our samples size would have increased by nearly 50%. Nevertheless, this is a very important finding as it reflects the problems authors of overviews face when trying to account for overlaps. By contacting the authors of overviews or reviews for clarification, we used the same method potential overview authors would have used. However, contacting the authors was seldom successful. All things considered, the characteristics of our overviews are very similar to a recent work [1] with respect

Table 3. Correlation coefficients

Coefficients	CCA	No. of reviews	No. of primary publications	Overlap %
CA	0.520**	−0.814**	−0.600**	0.192*
CCA		−0.323**	−0.373**	0.643**
Number of reviews			0.601**	−0.002
Number of primary publications				−0.157

Abbreviations: HTA, health technology assessment; IQR, interquartile range; CA, covered area; CCA, corrected covered area.
*$P < 0.05$; **$P < 0.01$.

Table 4. Overview characteristics by publication type and dealing with overlaps (mean [SD])

Characteristics	Journal ($n = 45$)	HTA ($n = 15$)	*P*-value
Number of reviews	12.0 (10.1)	4.9 (2.60)	0.003
Number of primary publications	200.0 (232.5)	66.5 (61.8)	0.006
Overlap %	22.6 (13.5)	31.4 (21.0)	0.140
CA	19.6 (14.3)	38.9 (20.7)	0.001
CCA	6.4 (7.7)	16.6 (19.3)	0.053
	Dealt with overlaps ($n = 32$)	**Not dealt with overlaps ($n = 28$)**	***P*-value**
Number of reviews	8.25 (6.9)	12.54 (11.1)	0.288
Number of primary publications	124.6 (141.5)	214.7 (264.7)	0.310
Overlap %	27.2 (14.5)	21.9 (17.4)	0.121
CA	25.4 (14.5)	23.4 (21.6)	0.133
CCA	9.0 (8.5)	9.0 (15.8)	0.037

Abbreviations: SD, standard deviation; CA, covered area; CCA, corrected covered area.
Mann-Whitney-U-statistic.

to the number of included reviews in overviews. Thus, we believe that our sample is representative.

However, we acknowledge that our study has some limitations. We chose primary publications as the unit of analysis. It is irrefutable that our analysis would have been more precise if we would have accounted for multiple publications. However, our approach seems to be unavoidable. Authors of overviews would need to obtain the full texts of all included primary publications to preclude the double counting that overstates the evidence [11]. Only in this way, authors can analyze the evidence and avoid the problem of including multiple publications [12,13]. Although all primary publications could be obtained, it would consume a lot of time and that would negate the time saved by conducting an overview instead of an SR. A similar problem can arise when including conference abstracts because the results of studies are usually presented on more than one occasion. In particular, this was an issue in some groups of the Cochrane Collaboration. Furthermore, authors of overviews might face problems obtaining all of the primary publications included in the overview. However, by not recognizing double publications in our situation will result in underestimating the degree of overlaps in our analysis.

Conversely, there may also be some pitfalls that might result in an overestimation of the degree of overlap. According to the Cochrane Handbook, overviews can be helpful for comparing different indications [14]. This turns out to be problematic if studies contain more than one indication. For instance, a study investigating the volume–outcome relationship in surgery analyzed several procedures [15]. Many indication or procedure-related SRs have been published to explore the volume–outcome relationship [16]. Of these, each review might have included the aforementioned study extracting only the data for one indication or procedure within the scope of the review. Thus, this study might be included in several reviews with different outcomes. Of course, there is really no overlap in data; but when conducting an overview of these reviews, we will find the single study in multiple reviews declaring it misleadingly as an overlap. We assume that this issue might be more prominent in reviews of observational studies and is only relevant to a small number of studies. Thus, we believe that our results rather underestimate than overestimate the degree of overlap.

For example, the broader the research question, such as investigating different interventions in different populations for a given condition, the more reviews will be included in an overview. In this case, the calculated overlap is expected to decrease with the number of reviews because it is improbable that reviews comparing different interventions or populations rely on the same primary studies. This is in accordance with our finding of a significantly high negative correlation between the CA and the number of reviews. As expected, the corresponding correlation of the CCA with the number or reviews is much weaker.

The inclusion of different study types throughout reviews may artificially minimize the overlap. Two reviews investigating the same research question might include different study types. For example, one review may only include randomized controlled trials (RCTs) and the second review included both cohort studies and RCTs. Both reviews may include exactly the same RCTs suggesting a full overlap between both reviews, but with the inclusion of each additional cohort study the overlap will decrease. Even if this is fully justifiable from a statistical point of view, it disregards the levels of evidence. In our example, the additional inclusion of cohort studies makes no statement about their weight in the evidence synthesis. However, it might be necessary for the analysis to include cohort studies to explore adverse effects [17]. Furthermore, the weight of study in the evidence synthesis does not only rely on the study design but also on the number of patients included and effect estimates, among others. Our analysis does not account for this problem, but it is an important problem to be aware of when conducting overviews and not a limitation of our findings. This problem is simply owing to the level of analysis in overviews, namely the use of reviews instead of primary studies.

5. Conclusion

Authors often disregard overlaps rather than address the issue in their work, which may lead to false conclusions. Thus, all producers of overviews should analyze the

overlaps and report their analysis. Reporting should be done even if the amount of overlap is small and unlikely to have an impact on the conclusion. Otherwise consumers will not know whether there is no meaningful overlap or if the authors simply did not account of it. Consequently, overlaps should be reported by default, and this should be taken into consideration when developing reporting guidelines for overviews. Notably, overlaps were missing in a recently developed quality checklist [18].

We introduced and validated two possible measures of overlaps. Based on our analysis, we encourage authors to calculate and report the CCA in future overviews because it is the best approximation of overlap.

The CCA is comprehensive and easy to understand. Nevertheless, more practical examples of the CCA are needed to verify our measure.

High-quality reporting is a prerequisite for conducting overviews and analysis of overlaps. Producers of overviews should consider excluding reviews without a detailed and complete list of the included studies. The same also applies to reviews with a poorly documented evidence synthesis that has confusing or incomprehensible results.

Acknowledgments

The authors are thankful to all the authors who submitted the requested files or answered our clarification questions.

Appendix

Supplementary material

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jclinepi.2013.11.007.

References

[1] Hartling L, Chisholm A, Thomson D, Dryden DM. A descriptive analysis of overviews of reviews published between 2000 and 2011. PLoS One 2012;7(11):e49667.

[2] Smith V, Devane D, Begley CM, Clarke M. Methodology in conducting a systematic review of systematic reviews of healthcare interventions. BMC Med Res Methodol 2011;11:15.

[3] Cooper H, Koenka AC. The overview of reviews: unique challenges and opportunities when research syntheses are the principal elements of new integrative scholarship. Am Psychol 2012;67(6):446–62.

[4] Pieper D, Buechter R, Jerinic P, Eikermann M. Overviews of reviews often have limited rigor: a systematic review. J Clin Epidemiol 2012;65:1267–73.

[5] Higgins J, Green S. Cochrane handbook for systematic reviews of interventions. Hoboken, NJ: The Cochrane Collaboration; 2011.

[6] Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. Ann Intern Med 1997;126:376–80.

[7] Thomson D, Russell K, Becker L, Klassen T, Hartling L. The evolution of a new publication type: steps and challenges of producing overviews of reviews. Res Synth Methods 2010;1(3–4):198–211.

[8] Velasco Garrido M, Gerhardus A, Rottingen JA, Busse R. Developing health technology assessment to address health care system needs. Health Policy 2010;94(3):196–202.

[9] Martel G, Duhaime S, Barkun JS, Boushey RP, Ramsay CR, Fergusson DA. The quality of research synthesis in surgery: the case of laparoscopic surgery for colorectal cancer. Syst Rev 2012;1(1):14.

[10] Stewart L, Moher D, Shekelle P. Why prospective registration of systematic reviews makes sense. Syst Rev 2012;1:7.

[11] Senn SJ. Overstating the evidence: double counting in meta-analysis and related problems. BMC Med Res Methodol 2009;9:10.

[12] Gotzsche PC. Multiple publication of reports of drug trials. Eur J Clin Pharmacol 1989;36:429–32.

[13] Leizorovicz A, Haugh M, Boissel JP. Meta-analysis and multiple publication of clinical trial reports. Lancet 1992;340:1102–3.

[14] Becker L, Oxman A. Chapter 22: Overviews of reviews. In: Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions. Chichester, UK: John Wiley & Sons; 2008.

[15] Birkmeyer JD, Siewers AE, Finlayson EV, Stukel TA, Lucas FL, Batista I, et al. Hospital volume and surgical mortality in the United States. N Engl J Med 2002;346:1128–37.

[16] Davoli M, Amato L, Minozzi S, Bargagli AM, Vecchi S, Perucci CA. Volume di attività ed esito delle cure: revisione sistematica della letteratura. [Volume and health outcomes: an overview of systematic reviews]. Epidemiol Prev 2005;29(3–4 Suppl):3–63.

[17] Loke YK, Price D, Herxheimer A. Adverse effects. Cochrane Handbook for systematic reviews of interventions. John Wiley & Sons, Ltd; 2008:433–48.

[18] Li L, Tian J, Tian H, Sun R, Liu Y, Yang K. Quality and transparency of overviews of systematic reviews. J Evid Based Med 2012;5(3):166–73.