ELSEVIER

CrossMark

Journal of Clinical Epidemiology 67 (2014) 1302–1308

Journal of Clinical Epidemiology

REVIEW ARTICLES

Up-to-dateness of reviews is often neglected in overviews: a systematic review

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

Institute for Research in Operative Medicine, Faculty of Health, School of Medicine, Witten/Herdecke University, Ostmerheimer Str. 200, Building 38, D-51109 Cologne, Germany

Accepted 25 August 2014; Published online 1 October 2014

Abstract

Background and Objective: As systematic reviews may run out of date, it might be necessary to update them. Out-of-date reviews may jeopardize the comparability when used in the context of overviews (review of reviews).

Methods: Seven electronic databases were searched for overviews up to November 2012. We first aimed to analyze whether the authors of overviews additionally searched for primary studies or alternatively explained why they did not. Second, we sought to analyze the actual publication lag (publication date of the overview − publication date of the review) in overviews and to develop recommendations for authors of overviews.

Results: We included 147 overviews. The mean publication lag in overviews was more than 5 years. A median of 36% of the reviews were published more than 6 years ago. Only one in four reviews considered up-to-dateness. The methods for updating reviews were heterogeneous. We found no overview that systematically investigated whether an update was necessary.

Conclusion: The issue of up-to-dateness when conducting overviews seems to be neglected by most authors of overviews. Authors should assess the quality of evidence, based on their included reviews first.

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

1. Introduction

Keeping current with the scientific literature is a very challenging task for researchers but even more so for health professionals as the amount of published literature in medical science is rapidly rising. Eleven systematic reviews (SRs) and 75 trials need to be read every day to keep up-to-date, when just considering the publications listed in MEDLINE [1].

This huge amount of literature has led reviewers to perform evidence syntheses on reviews instead of primary studies that are often called overviews (of reviews), review of reviews, and umbrella reviews [2]. Overviews enable the comparison of findings across conditions or interventions and thus potentially provide answers to questions for which a trial would never be performed. At the same time, searching for reviews might be time saving as there are fewer hits to screen and less data to extract compared with a synthesis of primary studies [3]. Although the number of published overviews is steadily increasing, there is only little methodological guidance for conducting sound overviews [4]. Many methodological issues have still not been addressed.

If there is a need for updating the included reviews, one of the methodological questions is related to updating literature searches to keep them up-to-date [4]. There is evidence that approximately 15%, 25%, and 50% of published reviews are out-of-date after 1, 2, and 5.5 years, respectively [5]. Including multiple SRs in an overview inevitably poses the question of whether they are up-to-date. This issue requires consideration especially when they are not up-to-date as physicians and decision makers are rather interested in the most current evidence. However, our knowledge on how and when to update SRs is scarce [6]. The ongoing scientific debate seems to concentrate on methods for identifying out-of-date reviews [7–9]. Only 5% of overviews update their included SRs by searching for additional primary studies [4]. The Cochrane Collaboration states that it is not necessary to search for additional primary studies [10]. But as Cochrane reviews are usually updated on a regular basis,

Conflict of interest: None.

Funding: None.

* Corresponding author. Tel.: +49 (221) 98957-42; fax: +49 (0)221 98957-30.

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

0895-4356
http://dx.doi.org/10.1016/j.jclinepi.2014.08.008

What is new?

Key findings

- The mean publication lag per review was more than 5 years.
- Only one in four overviews considered up-to-dateness.
- No overview systematically investigated whether an update was necessary.

What this adds to what was known?

- This is the first systematic analysis of up-to-dateness in overviews.
- We developed recommendations to produce up-to-date overviews.

What is the implication and what should change now?

- Authors should analyze whether the underlying evidence of systematic reviews (SRs) is still up-to-date when conducting overviews.
- Authors should search for primary studies not included in SRs, if needed

the problem of out-of-date reviews is reduced. Some health technology assessment (HTA) agencies showed that a number of agencies emphasized the need of additional searches for primary studies, albeit this may be not necessary if there is already enough evidence. Consequently, in the case that not enough evidence exists, authors should search for additional primary studies.

To tackle this issue, we first aimed to analyze whether the authors of overviews additionally searched for primary studies or alternatively explained why they did not. Second, we sought to analyze the actual publication lag in overviews and to develop recommendations for authors of overviews.

2. Methods

2.1. Data sources and searches

To identify overviews, MEDLINE (via EMBASE), EMBASE (via EMBASE), Cumulative Index to Nursing and Allied Health Literature (via EBSCO), Physiotherapy Evidence Database, and all databases of the Cochrane library (via the Cochrane library) were searched (see Appendix A at www.jclinepi.com, for search strategies). The last search was carried out in the beginning of November 2012.

2.2. Inclusion criteria

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. The authors had to name at least one database and explicitly state that they searched for reviews. We included overviews, irrespective of whether they also included primary studies additionally. For inclusion, the evidence synthesis had to rely at least in part on reviews (eg, combining primary studies and reviews in the evidence synthesis). Furthermore, all included literature (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 quality of SRs in a specific field). The articles had to be written in English or German.

2.3. Study selection

All titles and abstracts were screened independently. The full texts of potentially eligible articles were then obtained and assessed for eligibility against the review inclusion criteria. All steps were performed independently by two researchers. Any disagreements were resolved by discussion.

2.4. Data extraction and analysis

The data were extracted by one reviewer in structured summary tables and checked for accuracy by a second reviewer. Any disagreements were resolved by consensus. For each overview, we extracted data on

- The name of the first author
- The publication year of the overview and of all included reviews
- The objective
- The number of the included reviews
- The number of the additionally included primary studies (if any)
- Whether the authors searched for additional primary studies
- Considering up-to-dateness of reviews (at least the authors should acknowledge that the included reviews might be outdated; in the case that authors searched also for primary studies, we assumed that they considered up-to-dateness; an explanation that more research is unlikely to change the result in a significant manner; a systematic assessment on whether an update of the underlying reviews is necessary would be a welcomed strategy. An evidence grading system could be used to achieve this, for example.)
- The way they searched for primary studies [in parallel/simultaneously (eg, one search strategy to identify reviews and primary studies at the same time) or additional searches for primary studies (eg, one search

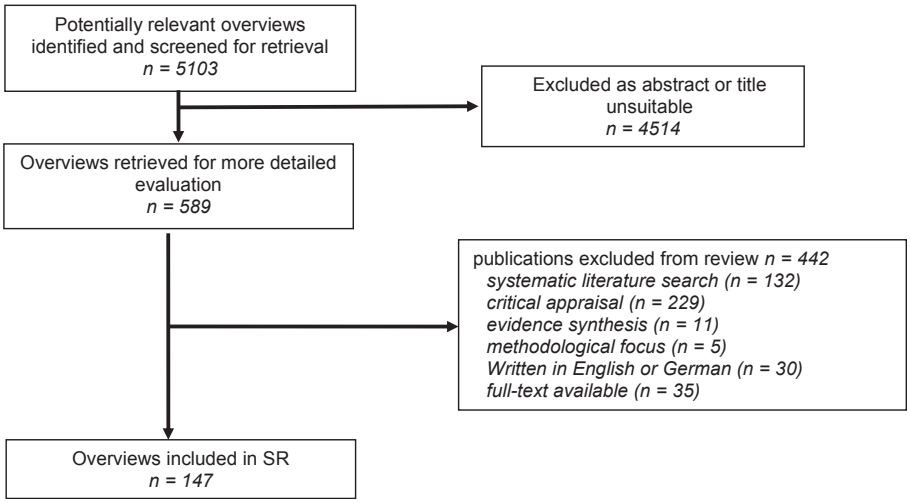

Fig. 1. Flow chart. SR, systematic review.

strategy for reviews and a second independent search strategy for primary studies)]

We generated the publication lag by subtracting the publication year of an included review from the publication year of the overview. The mean publication lag was calculated for each overview. Furthermore, following Shojania et al. [5], we assumed that the probability for a review to be out-of-date after 6 years is more than 50%. Thus, we calculated the proportion of reviews in overviews that were published 6 or more years before the overview.

3. Results

We identified 5,103 citations and included 147 overviews in our review (Fig. 1). Appendix B at www.jclinepi.com lists the included and excluded studies, along with the reasons for exclusion. Table 1 summarizes the general overview characteristics.

The number of included reviews in overviews ranged from 2 to 153 (median, 12; interquartile range, 6–26.75). The mean publication lag per review was more than 5 years. This is close to the 6-year range where more than half of the reviews would be expected to be out of date [5]. The median

Table 1. Overview characteristics

Category	Characteristic	Number (%), *N* = 147
Publication year	2012 (up to November 2012)	25 (17)
	2011	30 (20)
	2010	30 (20)
	2009	20 (14)
	2008	11 (7)
	2007	8 (5)
	2006	6 (4)
	2005	7 (5)
	2004	3 (2)
	2003	2 (1)
	2002	0 (0)
	2001	2 (1)
	2000	1 (1)
	1999	0 (0)
	1998	1 (1)
	Not reported	1 (1)
Number of included reviews	Median (IQR)	12 (6–27)
Considering up-to-dateness of reviews if not searching for primary studies	*n* (%)	30 (26)[a]
Searches for primary studies	*n* (%)	30 (20)
Method of searching primary studies	Parallel, *n* (%)	24 (16)
	Updates, *n* (%)	6 (4)
Publication lag (yr)	Mean (SD)	5.18 (2.06)
Mean percentage of out-of-date reviews (≥6 yr)	Median (IQR)	36 (19–55)

Abbreviations: IQR, interquartile range; SD, standard deviation.

[a] Calculated from 30 of 117; reviews searching for primary studies were excluded from the denominator.

proportion of the included potentially out-of-date reviews was 35.72%. In addition to the inclusion of reviews, 30 overviews (20.41%) also searched for primary studies. Two different ways of searching for primary studies in overviews can be distinguished.

In the first approach (6 of 30), the authors identified reviews in a first step and then searched for additional primary studies.

In the second approach, the authors searched for secondary and primary literature simultaneously. This was the predominant way of searching for primary studies in our analyzed overviews (24 of 30). However, there were many ways to operationalize this. One used approach within parallel searches allows for the inclusion of primary studies only if they were published after the search period of a corresponding review. In a second approach, authors performed a "hierarchical evidence synthesis." In a hierarchical evidence synthesis, the authors begin with the highest level of evidence according to the evidence pyramid (ie, searching for reviews first) going down to lower ranked levels of evidence placed below in the evidence pyramid [eg, randomized controlled trials (RCTs), cohort studies] until the best available evidence can be identified. This could also mean that publications based on different levels of evidence are included depending on the intervention, for example. This might be the case if there is evidence from SRs for one intervention, but there is only evidence from RCTs for a second intervention for which no SR exists.

Subtracting the 30 overviews that also included primary studies in addition to reviews, 30 of the remaining 117 overviews considered up-to-dateness (we assumed that including additional primary studies is a strategy to consider up-to-dateness). In contrast, overviews that also searched for primary studies seldomly explained why they updated the search, as well. Collectively, we found no overview that systematically investigated whether an update was necessary. In the case of considering up-to-dateness, we found authors mostly ($n = 16$) stating the absence of updating reviews to be a limitation of their work. An explanation that more research is unlikely to change the result in a significant manner was very rarely mentioned [11,12]. Almost one-tenth (14 of 148) of our analyzed overviews used evidence grading systems. Anyway, we found no author using the grading system to assess whether updating reviews by searching for primary studies would be necessary, albeit five (four in parallel and one updating) of them searched for primary studies.

4. Discussion

Our analyses showed that currently, there seems to be no way to systematically investigate whether an update in the context of overviews is necessary. More than one-third of the included reviews in overviews are older than 5 years, indicating a high risk of being not up-to-date. The issue of up-to-dateness of reviews was only considered in every fourth review based on our analysis of 147 included overviews.

Two different ways of searching for primary studies in overviews were identified. On the one hand, authors searched in parallel for reviews and primary studies. On the other hand, authors performed additional searches for primary studies after having inspected the included reviews. Within the first approach (parallel searches), the idea of "hierarchical evidence synthesis" will not update any SR, but it might be worth investigating in overviews comparing different interventions for a given condition if there is no SR for one of the analyzed interventions. In general, the idea of parallel searches allows for the inclusion of primary studies only if they were published after the search period of a corresponding SR. This approach will ensure the up-to-dateness of the included reviews and provide an actual impression of the evidence in the field of investigation provided that the updated review is of high quality and is based on a comprehensive literature search. Updating low-quality reviews that did not perform a comprehensive search strategy by searching only the time span not covered by the review (ie, searching after the search period of a corresponding SR) might be pointless if many studies were not retrieved by the original search strategy. In this case, it would probably be necessary to perform a new search that is performed without any time restrictions.

It may also occur that more than one review for a specific question exists. However, it remains unclear how often this issue might arise as slight variations in search strategies, methods, populations, study designs, and outcomes can make similar looking reviews diverse. Addressing the issue of discordance, defined as conflicting results of reviews on the same review question, may be one of the major advantages of overviews, as it could help clinicians, patients, and policy makers base their decisions on the evidence most reliable and suitable to their situation. Discordance is rarely dealt with in overviews [4]. Jadad et al. [13] provide an algorithm that can help the authors of overviews to resolve issues of discordance. Probably to overcome this, some authors of our analyzed overviews include only the most recent review in the case of multiple reviews for a given question. In proceeding this way, they would limit the period in which potentially new primary studies were published. But this approach does not take the quality of the reviews into account and thus cannot be recommended as the most recent review might be of lower quality than an older review. Producers of HTA reports recommend only the inclusion of high-quality reviews when conducting an overview [14–16].

In the second approach, the authors identified SRs in a first step and then searched for additional primary studies. Methods for updating SRs varied between the overviews. A Cochrane overview used the search strategies of each included SRs to update it [17], although the Cochrane Collaboration states that it is not necessary to search for

additional primary studies. The precondition for this is that all authors of SRs report their search strategy. This was fulfilled in this case as only Cochrane reviews were included. Furthermore, a limited access to literature databases or search interfaces (eg, OVID) will often impede authors of overviews in proceeding this way to update SRs. If technically feasible, authors should assess the search strategies with appropriate methods before using them [18,19].

An overview on early-life determinants of obesity did not update all included SRs, but only the most recent SR for each relevant risk factor [20]. Another overview explored salt reduction in patients with hypertension [21]. Therein, evidence was updated for the time frame after the most recent SR. Arising problems related to this approach were already mentioned. However, a possible bias was minimized here as only high-quality SRs were included (scoring at least five points on the Overview Quality Assessment Questionnaire [22,23]), although this does not necessarily mean that these were based on a comprehensive search strategy.

If authors decided not to search for primary studies, we would expect an explanation for this decision or at least a reference on the fact that not updating SRs in overviews may lead to out-of-date results. This was fulfilled by every fourth overview.

An explanation that more research (ie, more primary studies) is unlikely to change the result in an evidence synthesis in a significant manner was very rarely mentioned. Such a statement is highly welcome as it clarifies if important findings from new primary research might have been missed. If doing so, authors should clearly state how they came to their conclusion. Such a conclusion could be derived by investigating the whole body of evidence. Evidence grading systems might be of high value here and were already used in some of the included overviews. However, some grading systems are based on the study design alone without explicit consideration of other important factors that play a major role in assessing the whole body of evidence based on secondary research. This includes in particular the up-to-dateness of research findings.

Up-to-dateness of SRs is a prerequisite for sound overviews. Updating the included SRs is time consuming but will yield a current picture of the evidence. Our analysis showed that authors of overviews do not systematically deal with up-to-dateness. This is probably due to the missing methodological guidance on how to update SRs in the context of overviews. Thus, we provide some recommendations for authors of overviews on what they should consider and how they should proceed when conducting high-quality overviews (Fig. 2).

4.1. Assess quality of evidence

Assessing the quality of evidence (QoE) is the crucial issue in all kinds of evidence syntheses. However, there are some important points that should be taken into consideration in the context of overviews. The critical appraisal of the included SRs is an important part in conducting an overview.

Primary studies might be included in more than one review. Informal analysis by summing up results of reviews, as we do in SRs of primary studies, could introduce significant overlap and double counting of evidence (primary studies) leading to biased overview results. Methods for calculating the degree of overlap have been recently published [24]. For example, take into account two SRs on the volume–outcome relationship in bariatric surgery [25,26] that were investigated by us in an overview on the hospital volume–outcome relationship in surgery

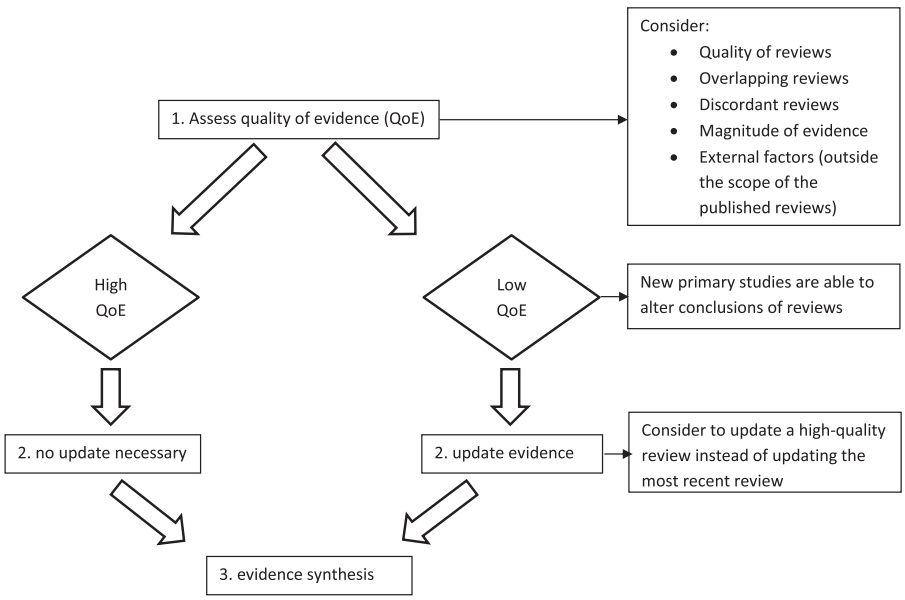

Fig. 2. Making overviews up-to-date.

[27]. The comparison of the search periods revealed that the review by Padwal is 2 years older than the review by Zevin. All studies in the review by Padwal were also included in the review by Zevin. Certainly, this is a glaring case of overlapping reviews. We could only rely on the review by Zevin and simply disregard the review by Padwal. But when investigating the quality of both reviews, it becomes apparent that the latter, the one that is probably more out of date, is of higher methodological quality when applying the Assessment of Multiple Systematic Reviews criteria [28].

It is important to note that even a number of additional low-quality SRs for the same question may have no impact on the QoE if there is high-quality evidence from at least one SR.

It may occur that an update is necessary for some but not for all aspects of an overview, depending on the research question. A rationale for an overview might be to summarize the evidence of different interventions for the same condition, for instance [10]. This can lead to a situation where we find a high QoE for one intervention but not for the second. In this case, evidence should be only updated for the second intervention.

Shekelle et al. [29] proposed a number of questions for the assessment of the evidence basis. They could also be taken into account when assessing QoE: Were any new interventions or procedures introduced (including approvals for new indications) that are possibly not covered by an SR? Were there any alerts (ie, pharmaceuticals or medical devices) or even market withdrawals? Has there been a change in scope with respect to the outcomes or population? Have new methodologies (eg, statistics) been introduced that may change the conclusions of the SR or the underlying studies? Authors should also consider whether slight modifications in interventions or even more in products might have an influence on the evidence synthesis. This is especially relevant for medical devices if new devices are only slightly modified from a device that has already been proven to be safe, for example. For example, take into account the case of hip implants. Hip implants are not required to undergo premarket testing. Manufacturers simply have to state that their hip implants are similar to hip implants that have already been proven to be safe. However, slight modification problems with new hip implants still exist. Thus, it is very important to watch out for new interventions or procedures when updating reviews.

Unfortunately, to the best of our knowledge, there is no established grading scheme for assessing the QoE based on SRs instead of on primary studies, so it is not our intention to condemn authors for not using a grading scheme. For example, the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach is one of the most prominent grading schemes where the quality of body of evidence, and not only individual studies, is assessed [30]. However, this approach is not transferable to overviews as some of the criteria are only applicable on the level of primary studies when assessing the body of evidence. For example, the QoE can be uprated in the case of large effect sizes. This could only be applied to SRs with meta-analyses as narratively synthesized SRs do not offer quantitative measures. Furthermore, publication bias will be hard to assess at the level of SRs. There is a need to develop a grading system that can also (or solely) be applied to SRs in overviews. In accordance with GRADE, the guiding idea for grading schemes should be whether new research on a given topic is likely to change the confidence in the estimate of the effect. This is the main question authors should allow for when deciding whether an update of the evidence is necessary in their overview. A modification of GRADE so that it can be applied in overviews as well would be pleasing and probably does make more sense as many authors are already familiar with the system. If authors do not want to use a grading system, they should clearly report their reasons and state the reasons for not analyzing the necessity of an update.

4.2. Limitations of our analysis

Our ad hoc descriptive analysis is explorative and has some limitations, but it brings up a disregarded subject with respect to overviews. Our results may be not comparable to the analysis by Shojania et al. because of our broader inclusion criteria without limiting the clinical area, for instance. Furthermore, our analysis was also not only limited to secondary research on randomized or quasi-randomized, controlled trials. There were also a number of overviews investigating risk factors where observational studies played a major role. There is currently no evidence on whether SRs of risk factors might be longer up-to-date as they are probably not as much subject to new interventions or enhanced procedures.

Because of the very large retrievals, we used a search strategy focusing on words appearing in the title of the article to increase precision, as we already did in a recent article. Thus, we probably have not identified all published overviews. However, it is unlikely that this would alter our conclusions.

It was not our intention to provide the publication lag as a (new) measure of up-to-dateness. It should be rather regarded as a descriptive measure as a part of an explorative analysis as an overview will be as up-to-date as the newest included review for that aspect of the overview.

5. Conclusion

Up-to-dateness of SRs is a prerequisite for sound overviews. We developed some recommendations for overviews against the background that evidence syntheses should be up-to-date as much as they can when they are published. This issue seems to be neglected by most authors of overviews. We suggest to focus on the assessment of the whole body of evidence and not only the single SRs. If there is

more than one SR addressing the same research question, these SRs should be regarded jointly. We encourage the use of grading systems to assess the QoE in the body of evidence. Future research should focus on grading systems (among other methods) easily applicable to overviews. This will help to make overviews more precise and valuable.

Acknowledgments

The authors thank Johannes Morche and Jana-Carina Morfeld for their help in data analysis and manuscript preparation.

Authors' contributions: S-.L.A. screened and searched the literature and analyzed the data. D.P. conceived of the study, participated in its design and coordination, searched and screened the literature, analyzed and interpreted the data, and drafted the manuscript. M.E. participated in the study design and participated in the analysis and interpretation of data. E.A.M.N. participated in the analysis and interpretation of data. All authors read and approved the final manuscript.

Supplementary data

Supplementary data related to this article can be found online at http://dx.doi.org/10.1016/j.jclinepi.2014.08.008.

References

[1] Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? PLoS Med 2010;7:e1000326.

[2] 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;1:198−211.

[3] Whitlock EP, Lin JS, Chou R, Shekelle P, Robinson KA. Using existing systematic reviews in complex systematic reviews. Ann Intern Med 2008;148:776−82.

[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] Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. Ann Intern Med 2007;147:224−33.

[6] Moher D, Tsertsvadze A, Tricco AC, Eccles M, Grimshaw J, Sampson M, et al. When and how to update systematic reviews. Cochrane Database Syst Rev 2008;MR000023.

[7] Chung M, Newberry SJ, Ansari MT, Yu WW, Wu H, Lee J, et al. Two methods provide similar signals for the need to update systematic reviews. J Clin Epidemiol 2012;65:660−8.

[8] Pattanittum P, Laopaiboon M, Moher D, Lumbiganon P, Ngamjarus C. A comparison of statistical methods for identifying out-of-date systematic reviews. PLoS One 2012;7:e48894.

[9] Shekelle PG, Newberry SJ, Wu H, Suttorp M, Motala A, Lim YW, et al. Identifying signals for updating systematic reviews: a comparison of two methods. 2011/08/12 ed. Rockville: Agency for Healthcare Research and Quality; 2011.

[10] Becker L, Oxman A. Chapter 22: overviews of reviews. In: Higgins JPTGS, editor. Cochrane handbook for systematic reviews of interventions. Chichester, UK: John Wiley & Sons; 2008.

[11] Kumar A, Galeb S, Djulbegovic B. Treatment of patients with multiple myeloma: an overview of systematic reviews. Acta Haematol 2011;125:8−22.

[12] Moe RH, Kjeken I, Uhlig T, Hagen KB. There is inadequate evidence to determine the effectiveness of nonpharmacological and nonsurgical interventions for hand osteoarthritis: an overview of high-quality systematic reviews. Phys Ther 2009;89:1363−70.

[13] Jadad AR, Cook DJ, Browman GP. A guide to interpreting discordant systematic reviews. CMAJ 1997;156:1411−6.

[14] Cleemput I, Van den Bruel A, Kohn L, Vlayen J, Vinck I, Thiry N, et al. Search for evidence & critical appraisal: health technology assessment (HTA). Brussels: Belgian Federal Health Care Knowledge Centre (KCE); 2007.

[15] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden. Cologne: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG); 2011.

[16] National Institute for Clinical Excellence. Methods for the development of NICE public health guidance (second edition). London: National Institute for Clinical Excellence (NICE); 2009.

[17] Cates CJ, Oleszczuk M, Stovold E, Wieland LS. Safety of regular formoterol or salmeterol in children with asthma: an overview of Cochrane reviews. Cochrane Database Syst Rev 2012;CD010005.

[18] Sampson M, McGowan J, Cogo E, Grimshaw J, Moher D, Lefebvre C. An evidence-based practice guideline for the peer review of electronic search strategies. J Clin Epidemiol 2009;62:944−52.

[19] Bak G, Mierzwinski-Urban M, Fitzsimmons H, Morrison A, Maden-Jenkins M. A pragmatic critical appraisal instrument for search filters: introducing the CADTH CAI. Health Info Libr J 2009;26:211−9.

[20] Monasta L, Batty GD, Cattaneo A, Lutje V, Ronfani L, Van Lenthe FJ, et al. Early-life determinants of overweight and obesity: a review of systematic reviews. Obes Rev 2010;11:695−708.

[21] Matyas E, Jeitler K, Horvath K, Semlitsch T, Hemkens LG, Pignitter N, et al. Benefit assessment of salt reduction in patients with hypertension: systematic overview. J Hypertens 2011;29:821−8.

[22] Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. J Clin Epidemiol 1991;44:1271−8.

[23] Oxman AD, Guyatt GH, Singer J, Goldsmith CH, Hutchison BG, Milner RA, et al. Agreement among reviewers of review articles. J Clin Epidemiol 1991;44:91−8.

[24] Pieper D, Antoine SL, Mathes T, Neugebauer EA, Eikermann M. Systematic review finds overlapping reviews were not mentioned in every other overview. J Clin Epidemiol 2014;67:368−75.

[25] Zevin B, Aggarwal R, Grantcharov TP. Volume-outcome association in bariatric surgery: a systematic review. Ann Surg 2012;256:60−71.

[26] Buchwald H, Buchwald JN. Evolution of operative procedures for the management of morbid obesity 1950-2000. Obes Surg 2002;12:705−17.

[27] Pieper D, Mathes T, Neugebauer E, Eikermann M. State of evidence on the relationship between high-volume hospitals and outcomes in surgery: a systematic review of systematic reviews. J Am Coll Surg 2013;216:1015−1025.e18.

[28] Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10.

[29] Shekelle P, Newberry S, Maglione M, Shanman S, Johnsen B, Carter J, et al. Assessment of the need to update comparative effectiveness reviews: report of an initial rapid program assessment (2005-2009); Rockville: Agency for Healthcare Research and Quality; 2009.

[30] Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011;64:401−6.