Journal of Clinical Epidemiology 108 (2019) 10−16

ORIGINAL ARTICLE

Living network meta-analysis was feasible when considering the pace of evidence generation

Perrine Créquit[a,b,c,*,1], Tania Martin-Montoya[a,b,1], Nassima Attiche[a,b], Ludovic Trinquart[d], Alexandre Vivot[a,b], Philippe Ravaud[a,b,c,e]

[a]*Centre d'Epidémiologie Clinique, Assistance Publique-Hôpitaux de Paris, Hôpital Hôtel-Dieu, Paris, France*
[b]*Centre de Recherche Epidémiologie et Statistique Paris Sorbonne Cité, INSERM U1153, Université Paris Descartes, Sorbonne Paris Cité, Paris, France*
[c]*Cochrane France, Paris, France*
[d]*Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA*
[e]*Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA*

Accepted 7 December 2018; Published online 12 December 2018

Abstract

Objectives: The aim of the study was to assess the feasibility of living network meta-analysis (NMA) taking into account the pace of evidence generation across different medical areas.

Study Design and Setting: We performed a systematic review to identify published NMAs. For each NMA, we calculated the cumulative number of new trials. To assess the feasibility of living NMA, we considered different update frequencies (4, 6, and 12 months), then evaluated the number of new trials to be included at each update in the NMA and the workload percentage for an update relative to the initial NMA.

Results: We identified 77 NMAs covering 17 different medical areas; 60 (78%) had fewer than four new trials included per year, on average, and 5 (7%) had more than seven trials. With an update frequency of 4, 6, and 12 months, the median number of new trials to be included in the NMA was 0 (interquartile range, 0−1), 1 (0−2), and 2 (1−4), respectively, with mean of 4%, 5%, and 11% workload per update, respectively.

Conclusion: The workload associated with updating a living NMA represents about one-tenth of the initial workload; therefore, living NMA is manageable.

Keywords: Live cumulative network meta-analysis; Living network meta-analysis; Living systematic review; Update; Systematic review; Evidence synthesis

1. Introduction

New evidence is continually appearing, and results from evidence synthesis can become rapidly out-of-date, sometimes even by the time they are published [1]. To be user-driven, evidence syntheses need to be up-to-date and incorporate recent data. The Cochrane collaboration advocated first updating systematic reviews on a regular basis, at least every 2 years [2]. Several authors have criticized the fixed updating frequency, considering it a "too basic approach for what appears to be a complex issue influenced by several factors" [3]. They mention that some fields evolve at an extremely rapid pace and others progress more gradually [4−7]. Updating evidence synthesis remains time- and resource-consuming. Even the Cochrane organization has had to adapt and consider an update based on need and priority [8].

A new approach has been proposed, called "living systematic review," which involves a continual surveillance of

Funding: This study was supported by a grant from the French National Cancer Institute (Institut National du Cancer, INCa; no. 2016-020/058/AB--KA). The funding source had no role in the design of this study, its execution, analyses, interpretation of the data, and decision to submit results.

Authors' contributions: P.C. was involved in the study conception, selection of trials, data extraction, data analysis, interpretation of results, drafting the article, and the revision of the article. T.M. was involved in the study conception, selection of trials, data extraction, data analysis, and drafting the article. N.A. was involved in the data analysis. L.T. was involved in the study conception, selection of trials, and data extraction. A.V. was involved in the data analysis, interpretation of results, and drafting the article. P.R. was involved in the study conception, interpretation of results, and drafting the article. All authors read and approved the final manuscript and the revised manuscript.

[1] These authors equally contributed to this study.

* Corresponding author. Centre d'Epidémiologie Clinique, Hôpital Hôtel-Dieu, 1 place du Parvis Notre Dame, 75004 Paris, France. Tel.: +33 1 42 34 89 32; Fax: +33 1 42 34 89 87

E-mail address: France.perrine.crequit@aphp.fr (P. Créquit).

https://doi.org/10.1016/j.jclinepi.2018.12.008
0895-4356/

What is new?

Key findings

- Many network meta-analysis (NMAs) had fewer than four new trials included per year.

What this adds to what was known?

- Whatever the update frequency (4, 6, or 12 months), the median number of new trials to be included in an NMA at each update was $\leq$2.
- Considering different update scenarios, the workload to conduct an update ranged from 4% to 14% of the initial workload.

What is the implication and what should change now?

- Living NMA is feasible, and analysis of the pace of evidence generation could help in determining the optimal update frequency.

the literature, keeping evidence syntheses as up-to-date as possible [9,10]. In the context of multiple competing treatments available for an increasing number of conditions, this concept has been extended to provide a complete and updated evidence synthesis comparing all available treatments for a given condition via "live cumulative network meta-analysis (NMA)," also called living NMA [11–13].

In this framework, the first step is to perform the initial NMA, considered the baseline up-to-date evidence [12]. Then iterations are repeated at regular intervals to keep the NMA up-to-date over time. Each iteration consists of six methodological steps: (1) adaptive search for treatments and trials, (2) screening of reports and selection of trials, (3) data extraction, (4) assessment of risk of bias, (5) update of the network of trials and synthesis, and (6) dissemination. This new form of synthesis lends itself to online posting of findings on an accompanying Web site to disseminate the updates and publications in journals at regular time points. In our study, we considered the posting of the results on a Web site as the final product of the living NMA to focus on its practicality. The crucial issue is to determine the frequency of the search and review updating. If research data emerge quickly, current evidence is uncertain, and end users need the latest evidence to make decisions.

The implementation of a continuous update scheme remains challenging, and determining an a priori optimal update frequency is difficult. A trade-off for the update frequency needs to be found to ensure that the NMA stays up-to-date without undue burden. The optimal update frequency would imply a feasible workload. Contrary to a standard review updating, which requires the intermittent deployment of substantial resources, maintenance of a living NMA should require modest resources. The screening and selection steps of a systematic review remain the most time-consuming tasks and the most cumbersome, largely for the initial NMA but also for updates. The workload of these steps is related to the pace of evidence generation for the research question of interest.

This study aimed to explore the pace of evidence generation across a wide range of medical areas to evaluate the feasibility of living NMA and help determine the update frequency.

2. Methods

After identifying published NMAs covering a wide range of medical areas, we first assessed the pace of evidence generation in terms of the cumulative numbers of new trials that appeared each year for each NMA. Then, we assessed the feasibility of living NMA by evaluating (1) the number of new trials to be included at each update in the NMA with different update scenarios and (2) the workload to conduct an update relative to the workload to complete the initial NMA.

2.1. Systematic review of NMAs

2.1.1. Search and selection

We performed a systematic search strategy to identify eligible NMAs. First, we screened the list of NMAs included in two previous methodological systematic reviews of NMAs that used a comprehensive search strategy [14,15] (Appendix A). Second, we searched for NMAs published in 2015 and indexed in MEDLINE using the strategy reported in Appendix A; the search was conducted on February 18, 2016. Eligible studies were NMAs of randomized controlled trials reporting the date of last search, the flowchart describing the study selection, and the citations of included trials. We excluded NMAs including fewer than 20 trials, updated NMAs, and overlapping NMAs. We chose to study NMAs with more than 20 trials (20 being the median number of trials included in NMAs [14]) to focus our demonstration on the "worst-case scenario"—NMAs with a priori the greatest pace of evidence—and to ensure that we had sufficient data. In cases of overlapping NMAs, we selected the NMA that included the largest number of trials. Titles and abstracts were independently screened, and the full-text articles were selected by two reviewers. Disagreements were discussed to achieve consensus.

2.1.2. Data extraction

For each NMA, two reviewers extracted the publication date, medical area, type of intervention, and treatment arms for all trials included in the network of trials (regardless of outcome).

2.2. Pace of evidence generation

In an exploratory phase, for each NMA, we plotted the cumulative number of trials included and the network of trials. We then focused on the evidence generation over the last 10 years, fixing as a time origin the date of the last search of the NMA. We decided to consider this last 10-year period because (1) the pace of evidence generation is not constant over time, (2) the time lag between the first and the last trial included in each NMA is highly variable, and (3) it seems appropriate to consider recent changes in the pace before starting a living evidence synthesis process.

2.3. Feasibility assessment

2.3.1. Number of new trials to be included

We calculated for each NMA the number of new trials to be included at each update according to three different update scenarios (every 4, 6, and 12 months).

2.3.2. Workload for an update

For each NMA, we estimated the workload involved to perform an update relative to the workload needed to complete the whole NMA. For the three update time scenarios, we divided the number of trials to be included in the NMA at each update by the total number of trials published up to the time of the update—considering that the latter should be greater than 2 to consider an update scheme. We represented this workload as a percentage of the entire workload needed to complete the whole NMA. All analyses involved using R v3.1.0 (R Core Team, Vienna, Austria) [16].

2.4. Sensitivity analyses

We performed sensitivity analyses (1) considering different periods (5 years and 1 year instead of 10 years) to calculate the number of new trials to be included at each update; (2) considering a sample of 50 NMAs including fewer than 20 trials (this sample was randomly selected among the NMAs including fewer than 20 trials identified by the selection process); and (3) considering the different medical areas.

3. Results

3.1. Systematic review of NMAs

We included 77 NMAs covering 17 different medical areas (flow chart provided in Appendix B). Most NMAs assessed a pharmacologic intervention (n = 62, 81%; Table 1) and cardiovascular disease (n = 17, 22%), then five medical areas (anesthesia/intensive care, oncology/supportive care, psychiatry/psychology, rheumatology, and hepato-gastroenterology). The NMAs included a median of 37 (interquartile range [IQR], 24–63) randomized trials and assessed a median of 10 (IQR, 7–15) treatments. The median delay between the first and the last included trials was 21 (IQR, 15–28) years. The individual characteristics and list of NMAs are in Appendices C and D.

3.2. Pace of evidence generation

Plots of exploratory analyses are in Appendix E. We did not observe any link between the trend of pace of evidence generation and medical area (Appendix F). Fig. 1 shows the pace of evidence generation for the 77 NMAs over the last 10-year period: each NMA had three new trials included per year, on average; 60 (78%) had fewer than four new trials included per year, 12 (16%) had four to seven, and 5 (6%) had more than seven.

3.3. Feasibility assessment

3.3.1. Number of new trials to be included

The median number of new trials to be included in the NMA was 0 (IQR, 0–1) with updates at 4 months, 1 (0–2) with updates at 6 months, and 2 (1–4) with updates at 12 months (Fig. 2). For the whole set of NMAs, the number of new trials to be included in the NMA was very close for updates at 4 and 6 months but was quite different for updates at 12 months.

To illustrate fast paces of evidence generation, we chose two examples: our NMA assessing second-line treatments

Table 1. Characteristics of the 77 included NMAs (n = 77)

Item	n (%)
Type of intervention studied in NMAs, number (%)	
Pharmacological	62 (81)
Nonpharmacological	11 (14)
Both	4 (5)
Medical area, number (%)	
Cardiovascular disease	17 (22)
Anesthesia and intensive care	8 (10)
Oncology/supportive care	9 (12)
Psychiatry/psychology	8 (10)
Rheumatology	7 (9)
Hepato-gastroenterology	8 (10)
Endocrinology	5 (7)
Respiratory diseases	4 (5)
Other	11 (15)
Number of trials included in the network, median (Q1–Q3)	37 (24–63)
Number of treatments included in the network, median (Q1–Q3)	10 (7–15)
Publication year of the first trial included, median (Q1–Q3)	1991 (1983–1996)
Publication year of the last trial included, median (Q1–Q3)	2013 (2010–2014)
Delay between the first and last included trial (years), median (Q1–Q3)	21 (15–28)

Abbreviation: NMAs, network meta-analyses.

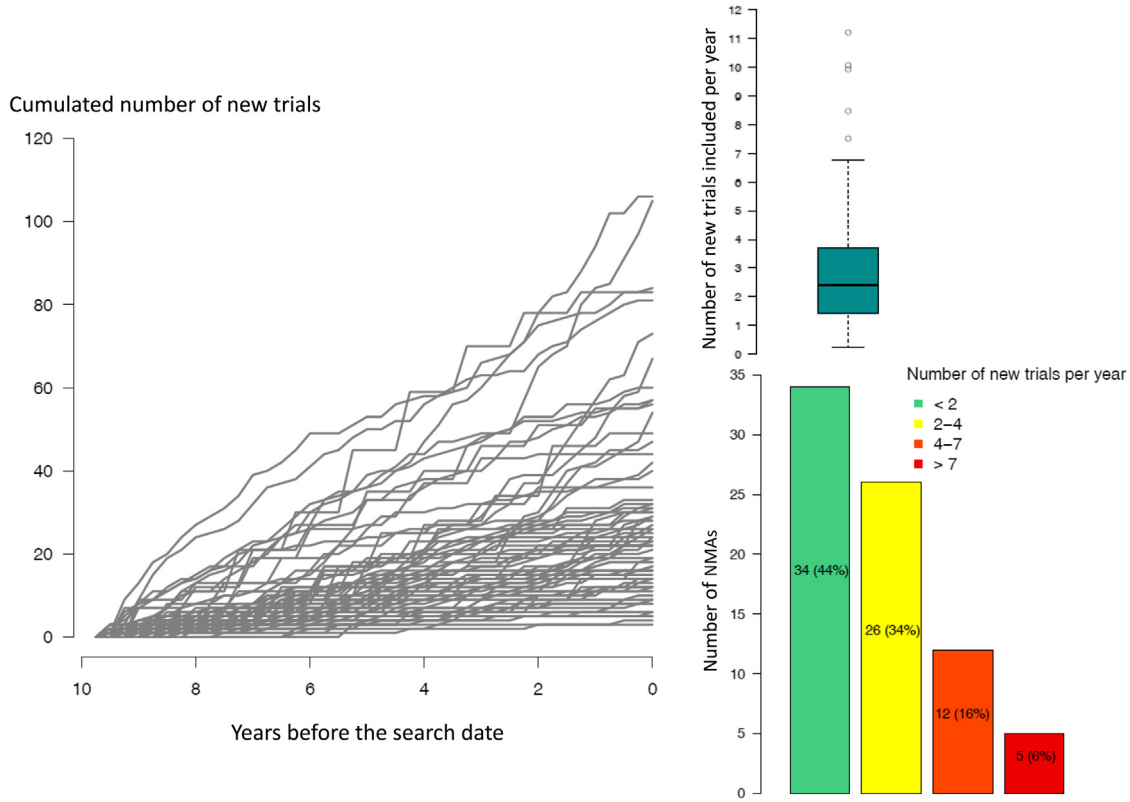

Fig. 1. Pace of evidence generation for the 77 network meta-analyses (NMAs) over the last 10 years. On the left figure, each line represents an NMA, the origin time corresponds to the last search date for each NMA. The right boxplot median and interquartile range represents the distribution of the number of new trials included per year for the 77 NMAs. The bottom-right histogram represents the distribution of NMAs according to different paces of evidence generation.

of non–small-cell lung cancer [11] and the NMA with a faster pace of evidence, which evaluated first-line treatments for glaucoma [17]. In the first example, the median number of new trials to be included in the NMA was 1.5 (IQR, 1–4.75) with updates at 4 months, 2.5 (1–5.25) with updates at 6 months, and 5.5 (2.25–9.5) with updates at 12 months. In the second, the medians were 4 (2–5), 5 (4–7.25), and 10.5 (8.25–12.75), respectively.

3.3.2. Workload for an update

The average workload for an update was 4%, 5%, and 11% for updates at 4, 6, and 12 months, respectively (Fig. 3). The third quartile of the workload reached 5%, 8%, and 16% for the three different update frequencies.

3.4. Sensitivity analyses

Sensitivity analyses of different periods (5 years and 1 year instead of 10 years) to calculate the number of new trials to be included at each update gave consistent results (Appendix G).

For the 50 NMAs including fewer than 20 trials (Appendix H), the three main medical areas were cardiovascular disease (n = 10, 20%), infectiology (n = 10, 20%), and surgery (n = 8, 16%). The median number of included randomized trials was 12 (IQR, 10–16), treatments assessed 5 (IQR, 4–7), and median delay between the first and last included trials 14 years (IQR, 9–18). Each NMA included one new trial per year, on average, with only seven NMAs (14%) including two to three new trials per year (Appendix I). The median number of new trials to be included in the NMA was 0 (IQR, 0–1) with updates at 4 and 6 months and 1 (0–2) with updates at 12 months (Appendix J). The average workload for an update was 5%, 7%, and 14% for updates at 4, 6, and 12 months, respectively (Appendix K).

Considering the medical area, we observed a large variability in the pace of evidence generation within the same medical area, which confirms that the pace of evidence generation is specific to each clinical question (Appendices F, J, and L).

4. Discussion

In this study, we assessed the feasibility of living NMA taking into account the pace of evidence generation across a wide range of medical areas. Many NMAs (n = 60, 78%) had fewer than four new trials included per year. The amount of work required to keep a living NMA up-to-date was manageable compared with the considerable investment required to complete the initial NMA (about 10%). To help determine the update frequency, the pace

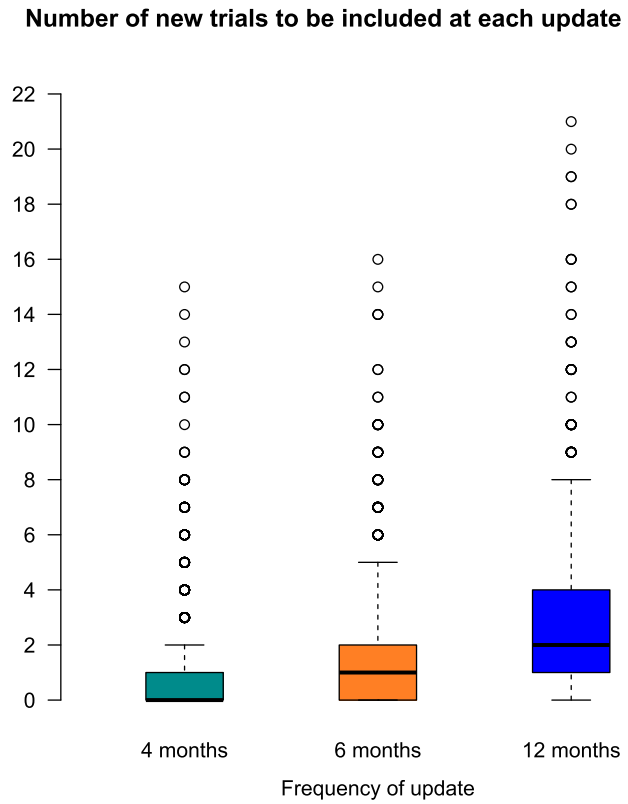

Fig. 2. Number of new trials to be included at each update. Box plots of median and interquartile range.

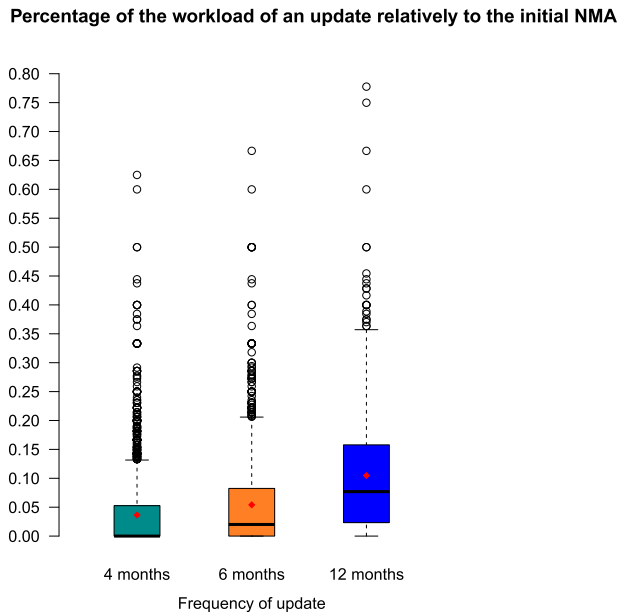

Fig. 3. Workload of an update. The workload of an update is represented as a percentage of the initial workload. Box plots of median and interquartile range. The red diamond represents the mean. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

of evidence generation—which is directly related to the workload—can be assessed. The frequency of updating depends on the clinical question, in most cases, but an update frequency of 12 months seems reasonable. For the small number of NMAs with a fast pace of evidence generation, updates every 4 or 6 months should be preferred.

To our knowledge, this is the first study exploring the pace of evidence generation to assess the feasibility of living NMA and help determine the updating frequency. The pace of evidence generation can be assessed by the number of new trial reports retrieved by search equations over time or the number of new trials selected at each update. The workload of most steps of a living NMA (data extraction, assessment of risk of bias, update of the network, and synthesis) is associated with the number of new trials included at each update. Moreover, the search and screening steps can be automated in the future [18,19]. Therefore, we decided to consider the pace of evidence generation as the cumulative number of new trials included in the NMA at each update.

The living systematic review network was recently proposed "to incorporate relevant new information within a maximum of 6 months of the information becoming available" [9]. However, the information will not become available at the same time for different research questions. Therefore, we propose assessing the pace of evidence generation for each research question. Indeed, our findings showed how the pace of evidence generation varies between the different NMAs and within an NMA itself. To manage the variability within NMAs (intravariability), we focused on the last 10-year period of the pace of evidence generation. Regarding the variability between NMAs (intervariability), only one-fifth of NMAs had a fast pace of evidence generation (more than four new trials per year) and 6% a very fast pace (more than seven new trials per year).

In this study, we assumed that all NMAs should be updated, rather than considered that some NMAs should be updated and others not because (1) determining which NMAs need to be updated based on needs and priorities without accepting a degree of arbitrariness is not easy [20], and (2) NMAs performed by a Bayesian approach are not affected by repeated updates. Our aim was to assess the feasibility of living NMAs by focusing on its practical aspect—all the steps up to posted results on a Web site—excluding any considerations regarding the need for the update. We considered that if the process was feasible, the thorny and unsolved issue of the need for updating did not need to be considered. Therefore, we show that, after the substantial cost of producing the initial NMA, maintaining the evidence up-to-date over time seems manageable with a reasonable investment over time (about 10%). We also acknowledge that we did not take into account the question of publication, nor the transfer into guidelines, because these remain important unsolved questions requiring a specific answer for each clinical condition.

Our work may help authors determine the update frequency in a living evidence synthesis process. We advise analyzing the pace of evidence generation for the research question and focusing on the last 10-year period to

determine a fast or slow pace of evidence generation. In the first case, authors should consider an update frequency of every 4−6 months, and in the second, every 6−12 months.

Our sample of NMAs covered a wide range of medical areas (17 different medical specialties). Our NMA characteristics are congruent with previous studies [14,15,21,22]. We found great variability in the pace of evidence generation within the same medical area, which confirms that the pace of evidence generation is specific to each clinical question and must be analyzed at the level of each clinical question. By selecting NMAs including more than 20 trials, we focused our demonstration on large NMAs with a priori the greatest pace of evidence and to ensure that we had sufficient data. However, the sensitivity analysis assessing NMAs including fewer than 20 trials showed a lower median number of new trials per year, close to 1, and a slightly higher relative workload for an update compared with NMAs including more than 20 trials (5%, 7%, and 14% vs. 4%, 5%, and 11% for updates at 4, 6, and 12 months, respectively). So, whatever the size of the NMA, the amount of work required to maintain a living NMA up-to-date is manageable, representing globally less than 15% of the initial workload.

We acknowledge that our study has several limitations. First, we performed a retrospective study to analyze the pace of evidence generation. If the pace of evidence generation is suddenly modified, our proposals for the choice of update frequency cannot be applied. Indeed, we chose to focus on the last 10-year period to analyze the pace of evidence generation, which is probably oversimplified. Because the pace of evidence generation for topics is prone to even out over time, the update frequencies suggested in this study might also need to be adapted over time. Second, we assessed the workload considering only the number of new trials to be included in the NMA at each update, not the number of new trial reports retrieved by the search equation, which may have underestimated or overestimated the overall workload in some cases. However, the number of new trial reports should be proportional to the number of new included trials, and therefore, the workload of an update should be similar. Third, our evaluation of the workload is not optimal: we used a proxy to have an idea of the workload for an update compared with the initial workload. We considered the workload up to the posting of results on the Web site and not up to a publication. Indeed, the interpretation phase for an update is reduced to the presentation of results with tables and figures, excluding the time required to update the written version of the article and format the submission. However, we acknowledged that our evaluation mainly reflects the time spent identifying studies, extracting the data, assessing the quality, and performing statistical synthesis—steps depending on the number of new trials contrary to interpretation phase, which always requires the same time. As a concrete example, for our NMA [11], we spent about 2 months for a 1-year update, representing 11% of the initial workload (18 months).

To conclude, our findings suggest that living NMA seems feasible. The updating frequency may be adapted to the pace of evidence generation. For a similar workload, the clinical question with a faster pace of evidence could be updated every 6 months, whereas those with a slower pace of evidence could be updated once a year.

Acknowledgments

The authors thank Aida Bafeta and Areti Angeliki Veroniki for the lists of NMAs.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jclinepi.2018.12.008.

References

[1] 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.
[2] Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions 2011. Available at http://handbook.cochrane.org/. Accessed July 11, 2016.
[3] Moher D, Tsertsvadze A. Systematic reviews: when is an update an update? Lancet 2006;367:881−3.
[4] Moher D, Tsertsvadze A, Tricco AC, Eccles M, Grimshaw J, Sampson M, et al. A systematic review identified few methods and strategies describing when and how to update systematic reviews. J Clin Epidemiol 2007;60:1095−104.
[5] Barrowman NJ, Fang M, Sampson M, Moher D. Identifying null meta-analyses that are ripe for updating. BMC Med Res Methodol 2003;3:13.
[6] French SD, McDonald S, McKenzie JE, Green SE. Investing in updating: how do conclusions change when Cochrane systematic reviews are updated? BMC Med Res Methodol 2005;5:33.
[7] Sutton AJ, Donegan S, Takwoingi Y, Garner P, Gamble C, Donald A. An encouraging assessment of methods to inform priorities for updating systematic reviews. J Clin Epidemiol 2009;62:241−51.
[8] Cochrane. Editorial and publishing policy resource 2016. Available at http://CommunityCochraneOrg/Editorial-and-Publishing-Policy-Resource. Accessed March 11, 2018.
[9] Elliott JH, Synnot A, Turner T, Simmonds M, Akl EA, McDonald S, et al. Living systematic review: 1. Introduction-the why, what, when, and how. J Clin Epidemiol 2017;91:23−30.
[10] Elliott JH, Turner T, Clavisi O, Thomas J, Higgins JPT, Mavergames C, et al. Living systematic reviews: an emerging opportunity to narrow the evidence-practice gap. PLoS Med 2014;11:e1001603.
[11] Créquit P, Trinquart L, Yavchitz A, Ravaud P. Wasted research when systematic reviews fail to provide a complete and up-to-date evidence synthesis: the example of lung cancer. BMC Med 2016;14:8.
[12] Créquit P, Trinquart L, Ravaud P. Live cumulative network meta-analysis: protocol for second-line treatments in advanced non-small-cell lung cancer with wild-type or unknown status for epidermal growth factor receptor. BMJ Open 2016;6:e011841.
[13] Vandvik PO, Brignardello-Petersen R, Guyatt GH. Living cumulative network meta-analysis to reduce waste in research: a paradigmatic shift for systematic reviews? BMC Med 2016;14:59.

[14] Bafeta A, Trinquart L, Seror R, Ravaud P. Analysis of the systematic reviews process in reports of network meta-analyses: methodological systematic review. BMJ 2013;347:f3675.

[15] Veroniki AA, Vasiliadis HS, Higgins JPT, Salanti G. Evaluation of inconsistency in networks of interventions. Int J Epidemiol 2013; 42:332−45.

[16] R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2016. Available at https://www.R-project.org/. Accessed May 31, 2016.

[17] Li T, Lindsley K, Rouse B, Hong H, Shi Q, Friedman DS, et al. Comparative effectiveness of first-line medications for primary open-angle glaucoma: a systematic review and network meta-analysis. Ophthalmology 2016;123:129−40.

[18] Tsafnat G, Dunn A, Glasziou P, Coiera E. The automation of systematic reviews. BMJ 2013;346:f139.

[19] Tsafnat G, Glasziou P, Choong MK, Dunn A, Galgani F, Coiera E. Systematic review automation technologies. Syst Rev 2014;3:74.

[20] Garner P, Hopewell S, Chandler J, MacLehose H, Schünemann HJ, Akl EA, et al. When and how to update systematic reviews: consensus and checklist. BMJ 2016;354:i3507.

[21] Nikolakopoulou A, Chaimani A, Veroniki AA, Vasiliadis HS, Schmid CH, Salanti G. Characteristics of networks of interventions: a description of a database of 186 published networks. PLoS One 2014;9:e86754.

[22] Zarin W, Veroniki AA, Nincic V, Vafaei A, Reynen E, Motiwala SS, et al. Characteristics and knowledge synthesis approach for 456 network meta-analyses: a scoping review. BMC Med 2017;15:3.