A multicomponent decision tool for prioritising the updating of systematic reviews

There is no formal consensus on when to update a systematic review, and updating too frequently can be an inefficient use of resources and introduce bias. A multicomponent tool could help researchers decide when is best to update such reviews.

Yemisi Takwoingi research fellow1, Sally Hopewell senior research fellow2,3, David Tovey editor in chief of the Cochrane Library4, Alex J Sutton professor of medical statistics5

1Public Health, Epidemiology and Biostatistics, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK; 2Centre for Statistics in Medicine, University of Oxford, Oxford, UK; 3INSERM, U738, Paris, France; 4Cochrane Editorial Unit, London, UK; 5Department of Health Sciences, University of Leicester, Leicester, UK

Evidence evolves as new research becomes available, and thus systematic reviews should be kept up to date to maintain their relevance and validity. However, the decision to update a systematic review should be made carefully because updating is potentially resource intensive, and updating too soon could introduce bias.1 In contrast, if reviews are not updated frequently enough, doctors and policy makers could act on evidence that is out of date. There is currently no consensus on when to initiate updating.2

Traditionally, the Cochrane Collaboration used a biennial updating policy,3 yet many Cochrane reviews are out of date (fig 1⇓); only about 20% of reviews are updated every two years. As of February 2013, the Cochrane Database of Systematic Reviews, published in the Cochrane Library, contained 5418 reviews.4 This number implies that 2700 Cochrane reviews will need to be updated per year to comply with the updating policy, in addition to producing new reviews. Despite the emphasis on Cochrane reviews in this paper, updating is a common challenge, and methods are needed to prioritise reviews that are most in need of updating.

A methodological review of different methods for identifying when to update a systematic review found that existing methods were not comprehensive because they only considered either qualitative or quantitative techniques.5 We describe a multicomponent (that is, including both qualitative and quantitative components) updating decision tool that can be used to determine when to prioritise one or more systematic reviews for updating. We provide a brief overview of the initial development and evaluation of the two complementary components, describe how they have been improved and integrated to produce the decision tool, and give illustrative examples of its implementation.

Development and evaluation of the qualitative decision tree

In 2007, the Cochrane Opportunities Fund supported a project to develop a strategic and evidence based approach to guide decisions of whether and when to update Cochrane reviews.6 An international steering committee was established to provide guidance and support to the project. A decision tree comprising a series of qualitative signals of the need to update and an updating checklist were developed. A pilot study evaluated their validity and reliability, and the results of the pilot and input from users and methodological experts were used to refine the decision tree. The full report of the project is available online.7

Development and evaluation of the quantitative tool

The initial quantitative approach4 adapted work on sample size methods for designing new studies to update existing meta-analyses.8 9 The scope of the approach was broadened by compiling a candidate list of quantitative signals (web fig 1), based on existing11 and new signals.12 The signals were categorised into those that can be evaluated “immediately” and those that require simulation methods to compute. Many signals on this list are self explanatory, or are based on quantities described elsewhere (that is, Barrowman’s new participant ratio13 and the I2 statistic14). The web appendix provides further details, and additional information is available elsewhere.15 Although in a small evaluation two signals performed well,6 we were
aware that a comprehensive approach including qualitative signals relevant to the entire review was essential.

Development and evaluation of the multicomponent updating decision tool

The two approaches outlined above were developed independently. Subsequently, a collaboration began between both development groups and the Cochrane Editorial Unit. The quantitative tool prioritised a portfolio of systematic reviews, on the basis of the signals or the total number of signals triggered, and further work was required to establish an optimal combination of signals. Thus, a principal task in developing the multicomponent tool involved developing a statistical prediction tool based on a prediction equation (box 1), derived from using signals from the candidate list. We developed the prediction tool using a sample of Cochrane reviews (web fig 2). The web appendix describes the development of the prediction tool with further details available elsewhere. We modified the qualitative decision tree and integrated the prediction tool to produce the first version of the multicomponent updating decision tool. The Cochrane Airways Group piloted this decision tool on 21 of their most cited systematic reviews or reviews with the most website hits (Welsh E, Karner C, Stovold E, Cates C, 20th Cochrane Colloquium, Auckland, New Zealand, 2012). It took five days to evaluate the 21 reviews, and eight reviews were identified that required updating (that is, with conclusions most likely to change). Consequently, resources that would have been wasted updating reviews less likely to benefit from updating were saved. The tool was considered easy to use, and it provided a structured and transparent way to assess and prioritise Cochrane reviews for updating. Challenges were encountered: for example, screening records for complex interventions, screening topics outside the screener’s comfort zone, and evaluating the possibility that updating the risk of bias would change the conclusions of the review. Nonetheless, the tool gave an opportunity to have these discussions and focused the debate. Specialists can be consulted on discreet questions if needed. The evaluation highlighted the need for clear and detailed guidance to enhance consistency in the assessment of subjective elements and appropriate application of the statistical prediction element of the tool. Therefore, we revised the guidance developed for Cochrane Review Groups, and the document is available on request from the editor in chief’s office, known as the Cochrane Editorial Unit (ceu@cochrane.org). Following feedback, we further refined the tool to create the updating decision tool summarised in fig 2).

Overview of the multicomponent updating decision tool

Our decision tool has three steps (fig 2), and an assessment is required at each step.

Step 1: Is the clinical question already answered by the available evidence or is the clinical question deemed no longer relevant?

If it is expected that there will never be any further information that could change the findings of the review, the current evidence is deemed conclusive, or that the clinical question is deemed no longer relevant, this should be noted. A decision can be made to flag the review as “Current question; no longer being updated” or “Historical question; no longer being updated” as appropriate.

Step 2: Are there any new factors relevant to the existing review?

These might include:

- Information from existing included studies—for example, information about new treatment regimens, population subgroups, harms, economic data, or outcome measures, including data from ongoing studies or previously missing data
- New methodology—for example, new statistical techniques, or changes in methodological guidance
- Response to feedback from users of the review
- Inclusion in policy decision making or clinical practice guidelines—for example, it might be important to update a review to include it in a new clinical guideline. If any such factors (termed updating signals) are identified, then a judgment is made on whether a signal for updating is likely or unlikely to change the results or conclusions of the review.

This step will involve a degree of subjectivity and should involve all members of the review team or editorial team (or both).

Step 3: Are there new studies?

If new studies are identified that are relevant to the primary outcome on which the conclusions of the review are based—and thus could be included in the main meta-analysis of this outcome in the review—then the statistical prediction tool can be applied. The probability (given as a percentage) of this new evidence changing the conclusions of the systematic review is based on the size and number of new studies added. Our results (web fig 3 and web appendix) indicated that a threshold probability of around 50% is suggestive of the need to update a systematic review, but any threshold can be chosen. If the new studies identified are not eligible to be included in a meta-analysis but still provide new information, then a judgment will need to be made about their likely effect on the conclusions of the review.

Documentation and presentation of decisions

To provide clarity and transparency for readers, the decision made at each step of the tool should be clearly documented together with reasons for the decisions. If an updating signal is deemed unlikely to change the conclusions of a systematic review, then the decision can be made not to update the review. For Cochrane reviews, such decisions (citing any new studies) can be documented in the “What’s new” table that shows readers the current status of the review, and the review can be flagged as “Current question; considered to be up to date.” For other systematic reviews, authors will need to explore avenues of dissemination with journals or publishers. Alternatively, if a signal is deemed likely to change the conclusions of a systematic review, and there is a review team available, the review should be updated as soon as possible. If a review team is not currently available then the review should be flagged as a “Priority for updating.” We retrospectively applied the tool to two Cochrane reviews to illustrate its use (boxes 3 and 2; tables 1 and 2 compare the out of date and updated reviews considered).

Discussion

The unification of two approaches has produced a comprehensive and practical tool that can aid updating systematic reviews at the appropriate time by identifying priority updates. We believe that the decision tool has a role in
improving the efficiency of the systematic review process; the task of updating can be managed methodically and limited resources used more efficiently. The tool can also be used to objectively compare systematic reviews that have new studies to determine which review(s) should have the highest updating priority. The statistical prediction element of the decision tool can rank reviews in order of the probability that conclusions would change of 96%. The new studies also provided data for additional analyses. These included the secondary outcome, which was the use of rescue treatment (proportion of participants using rescue treatment; time to use of rescue treatment), and sensitivity analyses (pain model, dental vs other types of surgery; dose response in dental studies, 200 mg vs 400 mg; salt preparation, standard ibuprofen v ibuprofen lysine, astringe, and “soluble”).

Decision: Flag the review as “High priority for updating.”

Current status and abridged history of the review in the Cochrane Library

Review content assessed as up to date: 11 May 2009.

Publication status and date: Stable (no update expected for reasons given in ‘What’s new’), published in issue 6, 2012.

‘What’s new’ table entry

Date: 25 April 2012.
Event: Review declared as stable.
Description: Although new studies on ibuprofen may be published, they are unlikely to affect the results of this review; therefore, the authors suggest that there should be no need to update this review for at least five years.

History

Date: 11 May 2009.
Event: New citation required and conclusions changed.
Description: Information from 37 new studies with 5595 participants was added, giving a total of 72 studies and 9186 participants. Numbers needed to treat for at least 50% pain relief over 4-6 h were not significantly changed. Additional information are provided on the proportion of participants requiring rescue treatment and median or mean time to use of rescue treatment, with higher doses giving slightly better results. Pain model and ibuprofen formulation could both affect the result, with dental impaction models and soluble ibuprofen salts producing better efficacy estimates. A dose response was shown in dental pain.

availability of new studies, and popularity of the review with their readers. The updating decision tool has limitations. The prediction component should not be used if studies included in the main meta-analysis in the earlier version of a review are excluded in the updated version. Prediction models are also typically overoptimistic when developed and the tool will benefit from evaluation in a wider cohort of reviews to assess the accuracy of predictions of the need to update. Updating is an iterative process, and we have not yet developed criteria to establish when enough evidence has accrued on a given topic and systematic review that it is deemed decisive and not worth conducting further primary research. This factor should be investigated in a future study.

The use of the tool depends on monitoring the literature for new evidence. Tsafnat and colleagues have suggested the possibility of automation of systematic reviews leading to best currently available evidence at the push of a button. Therefore, we can foresee a seamless and “on demand” approach to updating if the tool is linked with innovative methods, such as machine
learning techniques for identifying relevant new studies. We also anticipate a role for the quantitative component of the decision tool beyond updating. The quantitative component can be extended to calculate the likelihood of future studies overturning the conclusions of a new systematic review. This assessment will take into account ongoing studies and provide explicit information on further research needed, such as the number and size of new studies. This information can be added to the review to show the robustness of its conclusions and potential “shelf life.” It can also provide justification for new clinical trials.

Systematic reviews are not equal—the rate at which evidence accumulates and the effect of new evidence on conclusions may differ. The decision tool can provide channeling limited resources into updating systematic reviews that are most sensitive to change, thus improving the quality and reliability of healthcare decisions made on the basis of current evidence. We believe that this priority-based approach that minimises unnecessary updating is more sustainable than an approach based on an arbitrary time period.

We thank Rachel Marshall from the Cochrane Editorial Unit for the helpful advice given during this project; Kirsty Loudon, Mike Clarke, David Moher, Simon French, Anne Eisinga, and Rob Scholten for their valuable input into the earlier development of the quantitative component of the decision tool; and Emma Welsh, managing editor of the Cochrane Airways Group and who led the pilot study of the decision tool, and her team for sharing their experience of using the tool. The views expressed in this article are those of the authors and do not necessarily represent the views of the Cochrane Collaboration.

Contributors: All authors contributed to the idea and design of the study, drafted, and edited the manuscript. DT and SH obtained funding for the study. SH was part of the team that developed the original qualitative component. YT and AJS developed the original quantitative component. SH extracted the data, and YT conducted analyses under supervision by AJS. All authors are guarantors.

Funding: This project was funded by the National Institute for Health Research (NIHR) Cochrane-NHS Engagement Award Scheme (project no 10/4000/01). The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Department of Health.

Competing interests: All authors read and understood the BMJ policy on declaration of interests and declare financial support from the NIHR Cochrane-NHS Engagement Award Scheme.

Data sharing: No additonal data available.

Box 3: Applying the decision tool to a Cochrane review with a low priority for updating

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Is the clinical question already answered by the available evidence or is the clinical question deemed no longer relevant?</td>
</tr>
<tr>
<td>2</td>
<td>Are there any new factors relevant to the existing review to consider?</td>
</tr>
<tr>
<td>3</td>
<td>Are there new studies?</td>
</tr>
</tbody>
</table>

There were six new studies, as well as existing meta-analyses of the primary outcome. When three of the new studies were included in the meta-analysis of the primary outcome shown above, the prediction tool gave a probability that conclusions would change of 45%. Only one of the other three new studies contributed data (328 participants) to any of the other meta-analyses in the review.

Decision: Do not update yet. Flag review as “Current question: considered to be up to date.”

Current status and abridged history of the review in the Cochrane Library

Review content assessed as up to date: 12 July 2009.

Publication status and date: New search for studies and content updated (no change to conclusions), published in issue 4, 2009.

“What’s new” table entry

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>9 July 2009</td>
<td>New search has been performed.</td>
<td>Description: The review has been updated to include six new trials involving 4209 participants, bringing the total number of included trials to 10 involving 26865 participants. The background and discussion sections have been updated to include new material of relevance. The main conclusions are essentially unchanged from the previous version of the review.</td>
</tr>
</tbody>
</table>

History

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>18 August 2008</td>
<td>Amended.</td>
<td>Description: Converted to new review format.</td>
</tr>
</tbody>
</table>

Summary points

There is no consensus on appropriate methods for deciding when to update systematic reviews.

A decision tool was developed to replace an approach based on an arbitrary and rigid time period with a priority-based approach.

The tool broadly consists of three criteria: clinical question answered or no longer relevant, new relevant factors to consider, and availability of new studies.

The decision tool can help identify reviews most sensitive to change and thus minimise unnecessary updating and waste of resources.
### Tables

#### Table 1: Applying the decision tool to a Cochrane review with a high priority for updating: summary of out of date and updated reviews

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Out of date version</th>
<th>Updated version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Included studies</td>
<td>35</td>
<td>72</td>
</tr>
<tr>
<td>Primary outcome</td>
<td>Proportion with ≥50% pain relief over 4-6 h</td>
<td>Proportion with ≥50% pain relief over 4-6 h</td>
</tr>
<tr>
<td>Comparisons (no of studies)</td>
<td>Ibuprofen v placebo (n=35); diclofenac v placebo (n=6); diclofenac v ibuprofen (n=2)</td>
<td>Ibuprofen v placebo (n=72); comparisons involving diclofenac are no longer in the review</td>
</tr>
<tr>
<td>Ibuprofen dose (no of studies)</td>
<td>Mainly 200 mg (n=6) and 400 mg (n=31)</td>
<td>Mainly 200 mg (n=20) and 400 mg (n=61)</td>
</tr>
<tr>
<td>Participants</td>
<td>Adults (n=2214 on ibuprofen; n=1377 on placebo)</td>
<td>Adults (n=5804 on ibuprofen; n=3382 on placebo)</td>
</tr>
<tr>
<td>Findings (risk ratio (95% CI))</td>
<td>200 mg (707 participants): 4.7 (3.2 to 6.9); 400 mg (2817 participants): 3.4 (3.0 to 3.9)</td>
<td>200 mg (2690 participants): 4.6 (3.9 to 5.6); 400 mg (6475 participants): 3.9 (3.6 to 4.4)</td>
</tr>
<tr>
<td>Authors’ conclusions</td>
<td>Both drugs work well. Choosing between them is an issue of dose, safety, and cost.</td>
<td>The substantial amount of high quality evidence shows that ibuprofen is an effective analgesic in treating postoperative pain. Numbers needed to treat for 200 mg and 400 mg ibuprofen did not change significantly from the previous review even when a substantial amount of new information was added. New information is provided on retreatment.</td>
</tr>
</tbody>
</table>
Table 2 Applying the decision tool to a Cochrane review with a low priority for updating: summary of out of date and updated reviews

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Out of date version</th>
<th>Updated version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Included studies</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>Primary outcome</td>
<td>Measure of effectiveness: composite outcome of stroke, myocardial infarction, or death from a vascular cause</td>
<td>Measure of effectiveness: composite outcome of stroke, myocardial infarction, or death from a vascular cause</td>
</tr>
<tr>
<td>Comparison (no of studies)</td>
<td>Thienopyridines (ticlopidine, clopidogrel) v aspirin (n=4)</td>
<td>Thienopyridines (ticlopidine, clopidogrel) v aspirin (n=7)</td>
</tr>
<tr>
<td>Participants</td>
<td>Patients at high risk of vascular disease (n=11 329 on thienopyridines; n=11 327 on aspirin)</td>
<td>Patients at high risk of vascular disease (n=13 119 on thienopyridines; n=13 136 on aspirin)</td>
</tr>
<tr>
<td>Findings (odds ratio (95% CI))</td>
<td>0.91 (0.84 to 0.98); 22 656 participants</td>
<td>0.92 (0.85 to 0.99); 26 255 participants</td>
</tr>
<tr>
<td>Authors’ conclusions</td>
<td>The available randomised evidence shows that the thienopyridine derivatives are modestly but significantly more effective than aspirin in preventing serious vascular events in patients at high risk (and specifically in patients with a transient ischaemic attack or ischaemic stroke), but there is uncertainty about the size of the additional benefit.</td>
<td>The thienopyridine derivatives are at least as effective as aspirin in preventing serious vascular events in patients at high risk, and possibly somewhat more so. However, the size of any additional benefit is uncertain and could be negligible.</td>
</tr>
</tbody>
</table>
Figures

**Fig 1** Number of active and updated reviews published in the *Cochrane Library* from 1995 to 2012. Active reviews are the total number of published reviews excluding reviews that have been withdrawn. Therefore, it is expected that these reviews will be considered for updating within two years of the last date of publication.

**Fig 2** Multicomponent updating decision tool for prioritising systematic review updates. The decision made at each step of the tool should be carefully documented together with reasons for the decisions, thus ensuring transparency and reproducibility. The decision nodes have been numbered to aid the documentation process.