ELSEVIER

Journal of Clinical Epidemiology ■ (2017) ■

Journal of Clinical Epidemiology

REVIEW ARTICLE

Living systematic reviews: 3. Statistical methods for updating meta-analyses

Mark Simmonds[a,*], Georgia Salanti[b], Joanne McKenzie[c], Julian Elliott[c], On behalf of the Living Systematic Review Network

[a]*Centre for Reviews and Dissemination, University of York, York YO10 5DD, UK*
[b]*Institute of Social and Preventive Medicine (ISPM), University of Bern, Niesenweg 6, Bern 3012, Switzerland*
[c]*Cochrane Australia School of Public Health & Preventive Medicine, Monash University, Level 4, 553 St Kilda Road, Melbourne, Victoria 3004, Australia*

Accepted 18 August 2017; Published online xxxx

Abstract

A living systematic review (LSR) should keep the review current as new research evidence emerges. Any meta-analyses included in the review will also need updating as new material is identified. If the aim of the review is solely to present the best current evidence standard meta-analysis may be sufficient, provided reviewers are aware that results may change at later updates. If the review is used in a decision-making context, more caution may be needed. When using standard meta-analysis methods, the chance of incorrectly concluding that any updated meta-analysis is statistically significant when there is no effect (the type I error) increases rapidly as more updates are performed. Inaccurate estimation of any heterogeneity across studies may also lead to inappropriate conclusions. This paper considers four methods to avoid some of these statistical problems when updating meta-analyses: two methods, that is, law of the iterated logarithm and the Shuster method control primarily for inflation of type I error and two other methods, that is, trial sequential analysis and sequential meta-analysis control for type I and II errors (failing to detect a genuine effect) and take account of heterogeneity. This paper compares the methods and considers how they could be applied to LSRs.

Keywords: Living systematic review; Meta-analysis; Type I error; Type II error; Heterogeneity

1. Background

The key intention of a living systematic review (LSR, see Box 1), which differentiates it from a standard systematic review, is that it will be updated frequently, ideally as soon as any new relevant study is published or identified [1−3]. Over time the information available to be included may increase, requiring the review to be updated to ensure it is presenting the best available evidence. In many updates, this will require updating one or more of the meta-analyses included in the review.

There are two purposes for undertaking an LSR, which while subtly different have implications for the methods used to update meta-analyses. The first purpose is to present a summary of the evidence at the time of the most recent update. For this purpose, simply repeating each meta-analysis (whether fixed or random effects), adding the newly identified studies and presenting new forest plots and summary estimates, may be the most appropriate approach. All other components of the meta-analyses such as assessment of heterogeneity, subgroup analysis, and investigations of reporting bias will also have to be updated and repeated. Provided the meta-analysis methods used are appropriate, this approach will give the best estimate of the effect of interest at that point in time [4]. However, both the reviewers and readers should be aware that the results may change at later updates, and findings may be highly uncertain if there are few studies or participants included in the analysis.

Systematic reviews and meta-analyses are also used for clinical decision-making, guideline development, and reimbursement decisions. Typically, the level of credibility for the meta-analyses of many beneficial and harmful outcomes is considered before making recommendations for

Funding: All the authors of this paper were funded to produce this research by a grant from the Cochrane Methods Innovation Fund. The Living Systematic Review Network is supported by funding from Cochrane and the Australian National Health and Medical Research Council (Partnership Project grant APP1114605). Georgia Salanti is supported by a Marie Skłodowska-Curie fellowship (MSCA-IF-703254).

* Corresponding author. Tel.: 01904 321091.
E-mail address: mark.simmonds@york.ac.uk (M. Simmonds).

http://dx.doi.org/10.1016/j.jclinepi.2017.08.008
0895-4356/

What is new?

- Living systematic reviews will require updating of any included meta-analyses at each review update.
- If a living systematic review is used as part of a decision-making process, the frequent updating of the meta-analysis could lead to inappropriate conclusions being drawn, due to an inflated risk of falsely concluding statistical significance (type I error).
- Four statistical methods exist to avoid type I error inflation, and other statistical problems, that arise in repeated meta-analyses.
- This paper gives an overview of these methods and how meta-analyses should be performed in a living systematic review.

practice. An LSR in particular might be used to support the creation of ''living guidelines'' [5], in which the best available evidence about the benefits and harms of an intervention is used to inform frequently updated recommendations about the use of the intervention. The effect estimate from the meta-analysis and its precision (or confidence interval) is one of the deciding factors in grading the existing evidence, and in this paper, we discuss the implications of continually or frequently updating meta-analyses for the statistical precision of the summary effects.

In a meta-analysis of clinical trials, we may wish to determine if an experimental treatment is superior, inferior, or equivalent to a control treatment. If the review presents assessments of statistical significance with a conventional 95% confidence interval or a *P*-value of 0.05, then updating of the meta-analyses may overestimate the number of meta-analyses considered statistically significant. While each individual analysis has only a 5% chance of finding a statistically significant result when, in fact, there is none (type I error), the chance of finding a false statistical significant result in any one meta-analysis increases as we repeat these analyses with each review update [6].

As an example, consider a sequence of clinical trials of a new intervention compared to a control, with an updated meta-analysis conducted as soon as each new trial is published. Suppose that there is no true difference in effect between intervention groups on a particular outcome. In this circumstance, the type I error rate, of incorrectly getting a statistically significant result, rises rapidly with each new analysis, as shown in Fig. 1. Similarly, the confidence intervals that often accompany the summary effect will be too narrow if calculated using a conventional meta-analysis. Therefore, using assessments of statistical significance at any individual update of a meta-analysis carries a substantial risk of erroneously concluding that the new intervention

Box 1 Living systematic reviews

- A systematic review which is continually updated, incorporating relevant new evidence as it becomes available
- An approach to review updating not a formal review methodology
- Can be applied to any type of review
- Uses standard systematic review methods
- Explicit and a priori commitment to a predetermined frequency of search and review updating

is beneficial (or harmful). More formally, repeating a meta-analysis inflates the type I error.

In an LSR, we may also wish to determine when there is sufficient evidence such that we can be confident there is no meaningful effect to detect (such as no important difference in effect between new intervention and the control). This should be achieved so that a type II error is avoided, that is, the error of failing to detect a genuine effect and so that no future update will detect any evidence of a clinically meaningful effect. In a clinical trial, we might select an effect size to identify, such as a minimal clinically meaningful effect, a statistical power to detect that effect (e.g., 80% or 90%) and calculate the required sample size for the trial. We might conclude that the true effect size is less than the clinically meaningful effect if no statistically significant result is found once the specified sample size has been reached [7]. A similar approach can be taken with meta-analyses, including those in an LSR. However, previous analyses have found that few meta-analyses ever reach a sufficient sample size [8].

When an LSR is used only to summarize the best evidence on a topic over time, using standard meta-analysis methods should be sufficient as the review is updated. However, if the LSR is being used to make decisions or readers will use it to do so, then we may wish to consider approaches to avoid inadvertent type I and II errors. This paper considers four methods that have been proposed to correct for these potential errors when updating a meta-analysis. While this paper focuses on LSRs, the same issues apply to all systematic reviews which may be updated. For example, Cochrane recommends that all Cochrane reviews be kept up to date, with revisions at least every 2 years if new trials have been published.

2. Analysis methods for repeated meta-analyses

Updating a meta-analysis has some similarities with interim analyses of clinical trials [9–11]. Interim analyses are often performed in trials so the trial can be stopped early if there is convincing evidence that the intervention is beneficial or harmful. Methods have been developed to avoid type I and II errors and produce robust conclusions for these trial