Reviewer's report

Title: A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results

Version: 1  Date: 27 May 2015

Reviewer: Ian Saldanha

Reviewer's report:

General Comments

The manuscript deals with an interesting issue and is, for the most part, clearly written. I was mainly concerned about the novelty of your findings. It’s unclear what’s new about your work. You even reference a previous systematic review that has similar findings to yours. What are your key messages? I needed to be convinced that there is value to readers reading your manuscript beyond what is already out there. Sadly, I wasn’t.

Major Compulsory Revisions

Background

• In the first paragraph, you say that randomization provides consistent estimates of treatment effect. I don’t see how that is true. Consistency is a matter of reliability and might not be observed if the source populations in studies are different, even if randomization is done. Randomization doesn’t help cross that hurdle.
• I disagree that it is “impossible” to optimize internal and external validity. Consider studies where (1) you can randomize the entire target population; or (2) the treatment effect can be generalized to external populations. Such studies (and there may be other examples), if done well of course, optimize both internal and external validity.
• In this paragraph, and elsewhere in the manuscript, I don’t see why pragmatic trials are separated from RCTs. These are not mutually exclusive designs. A key feature of RCTs is the randomization, which can be perfectly well achieved in so called “pragmatic” trials. Pragmatism in clinical trials, as you state, is a spectrum (with plenty of overlap).
• At various locations, you use the word “representativeness”. This is a survey research term that gets at whether a study population “represents” the target population. In the context of RCTs, what matters is not whether the population is representative. Rather, of relevance is whether the treatment effect can be applicable to populations beyond those studied (i.e., the idea of “generalizability”).

Methods

• Your search was run in September 2013. You probably should update it.
• Why did you restrict to RCTs of pharmaceutical interventions? You need to discuss the reasoning behind this choice, its implications, and the limitations arising therefrom.

• Why did you exclude conference abstracts? Only 61% of conference abstracts of RCTs make it to full publication, and those that do are a biased subset (Scherer RW et al. Cochrane Database of Systematic Reviews 2007). You should discuss the implications of this.

Results

• I did not find the separation of the three fields (cardiology, mental health, and oncology) useful or meaningful. Generally, the findings were similar, and differences among fields do not appear to be your main message. At least, not main enough to give them so much airtime. It’s fine in a table, but you could summarize any differences among fields briefly in a couple of sentences. Also, they tended to lengthen your manuscript beyond what would probably be sufficient to convey your main messages. Or, maybe this was one of your main messages, and I just didn’t catch on.

• I liked your paragraph under “Explicit factors”. It seems to be the one of your main messages, and you reasoned through them nicely!

• Implicit factors: I found this vague. I did not feel like I learned about how these factors might affect the external validity. For many of them, my hypotheses could go both ways, and I would have benefitted with some hand-holding.

• Study recommendations: OK, I see that you only collated the recommendations. But, you could go further. Which ones do you think make sense? Which make sense but might not be feasible? Which are absolute must-dos? Why? Why might some of these recommendations have not caught on yet?

Discussion

• Buried on page 11, first paragraph is the point I raised earlier regarding external validity being possible even in the absence of generalizability. I’d give this idea more prominence, and even say it to the Background.

• I don’t understand what you are trying to say in the top paragraph on page 12. Please clarify.

• Page 12 – last paragraph

  o How do meta-analyses enhance external validity? If they are different populations (high clinical heterogeneity), they should not be meta-analyzed in the first place. In other words, meta-analysis does not take care of heterogeneity across individual studies.

  o You should define the methods that you are suggesting people use. It isn’t sufficient to just refer to them.

  o How do you suggest that the methods you list be used in the trial planning stage? Do you have examples?

  o What is “generalizability bias”? If I understand it correctly, how is it a “bias”? Even if it is a bias, how does predicting “the extent of generalizability bias”
improve the design of a trial?
o You say that CONSORT does not state that RCTs should include details of the background population. Do you suggest that CONSORT states this?

• Page 13 – second paragraph
o Certainly there can be “wrong” decisions regarding trial design!
o The language in this paragraph suggests that healthcare databases might replace RCTs. This is a highly debated issue in the literature and you should allude to writings regarding this issue if you are going to bring it up. Healthcare databases have challenges; some of which are: they were not designed for research, selection bias is a problem, confounding is a problem. After all, the databases are only a source of data, they are not a design!

Minor Essential Revisions

Background
• Last paragraph – are you sure patients in “observational cohorts” are those in “routine clinical practice settings”?

Methods
• First paragraph
o I don’t understand the first sentence.
o I’m confused about why you said that your review is not a systematic review because you did not have “outcome parameters a priori”. You did, right? I’m referring to the specific parameters you calculated and summarized under methods A and B.

• Page 5 – last paragraph
o How big was the “sample” of studies that you used to test your data extraction table?
o The data extracted can be more clearly described, especially (i), (ii), (iii), and (vi). It is unclear what was exactly extracted.

• Page 6 – first paragraph
o Explain what you mean by “statistical comparison”.

Discussion
• Please consider using sub-headers to structure your discussion. They help the reader take a break and know when you are switching from one idea to another.

Conclusions
• What are your “two ways” based on?

Level of interest: An article of limited interest

Quality of written English: Acceptable
Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:
I declare that I have no competing interests.