Author’s response to reviews

Title: Protocol for a Systematic Review of N-of-1 trial protocol guidelines and protocol reporting guidelines

Authors:
Antony Porcino (aporcino@touchoffreedom.com)
Salima Punja (punja@ualberta.ca)
An-Wen Chan (anwen.chan@utoronto.ca)
Richard Kravitz (rlkravitz@ucdavis.edu)
Aaron Orkin (aaron.orkin@mail.utoronto.ca)
Philippe Ravaud (philippe.ravaud@htd.aphp.fr)
Christopher Schmid (christopher_schmid@brown.edu)
Sunita Vohra (svohra@ualberta.ca)

Version: 1 Date: 21 Mar 2017

Author’s response to reviews:

NOTE: The PRISMA-P checklist initially submitted was to show compliance with PRISMA-P (pages on which each item was addressed), given our mention of writing this protocol using PRISMA-P in the manuscript. It was not meant as a supplemental file for publication, and has been removed.

Reviewer #1: This strikes me as being a fair proposal.

I have one concern. N of 1 trials are a minefield for statistical analysis. This is partly because the original developers of the concept, who deserve recognition for their pioneering enthusiasm, were not very knowledgeable about statistics. Hence some of the recommendations for analysis that were made originally are inappropriate. When this is added to the fact that data-structures can be very complicated (for example episodes within cycles within patients within a collection of patients or, as an alternative, designs based on Latin squares), this means that appropriate reporting of n-of-1 trials will require great care and considerable detail. I think that for your project to achieve success it will need complete up to date mastery of the statistical methodology of an evolving field with a fair amount of disagreement. You will need to make sure that you are equipped to do this. I have a strong suspicion that some of the published recommendations you will come across will be inappropriate and, indeed, that recommendations will contradict each other. You will need to exercise insight and judgment to know what to recommend. On the other
hand, of course, the complexity and controversy in the field means that your review has the potential to be particularly welcome and useful.

RESPONSE: Thank you – we agree. We think there are two issues to be addressed: 1) the complexity and current thinking about N-of-1 data analysis, and 2) assessing what is found in this systematic review.


With regards to #2, all of the authors have participated in N-of-1 trials or other methodological papers, and appreciate the comments regarding the complexity of the statistical issues. To address these, we will use an international Delphi process to develop an N-of-1 protocol reporting guideline. However, since our systematic review is focused on published guidelines and recommendations regarding the reporting of N-of-1 trial protocols (including analysis plans), we will use these to identify relevant issues for design such as assessing heterogeneity between participants, or how correlation, carryover, and period effects may be addressed.

Associate Editor's comments:

EDITOR: Line 76 to 78 – not clear why limited in terms of conditions and treatments

RESPONSE: N-of-1 trials are best suited for chronic conditions, when evaluating therapies with relatively quick onset/offset. Because N-of-1 evaluation takes time to conduct, they are more limited in their applicability to acute conditions or treatments that are irreversible or have a very long half-life. We have changed the wording, lines 77 to 83, to clarify this: “Due to the design’s multiple cross-overs within a single participant, some aspects of an N-of-1 trial require particular consideration for feasibility, including: i) the types of conditions (e.g., chronic and stable; if episodic, then frequently occurring) or treatments (e.g., reversible, preferably with quick onset and quick offset) that may be evaluated through the N-of-1 design; ii) some design constituents (e.g., shorter treatment period lengths preferable, presence or absence of washout)...

EDITOR: Line 80 to 83 – not entirely clear why or in what way treatment length would be different in N of 1 compared to RCTs – shorter?

RESPONSE: Treatment length would not be different; in an N-of-1 context, multiple periods are planned, and may affect feasibility. We have re-worded this to clarify (lines 77 to 83): “N-of-1
trial require particular consideration for feasibility, including:...ii) some design constituents (e.g., shorter treatment period lengths preferable, presence or absence of washout);...”

EDITOR: I’m a little concerned that your review will be an empty review, given that you state “an initial scan indicated no explicit reporting guideline exists for N-of-1 protocols” and that your inclusion criteria are for [sic]

RESPONSE: We would like to clarify: this initial scan confirmed that there is a need for a reporting guideline for N-of-1 protocols (one does not already exist). We do not anticipate an empty review, as a number of publications provide methodological recommendations for N-of-1 trials [1,17-23], including the Consolidated Standards of Reporting Trials (CONSORT) Extension for N-of-1 Trials (CENT) [8].” See lines 101 to 111 for clarification.

EDITOR: Does “an explicit, itemized guide detailing the content or headings for producing or reporting a complete N-of-1 trial protocol [14] or aspects of a such a protocol” mean you will include guidelines on how to design N of 1 trials? Perhaps make clearer? This is made explicit in the discussion, but I think it could be clearer here.

RESPONSE: Thank you. We have changed the wording from “producing or reporting a complete N-of-1 trial...” to “designing or reporting...” (line 136) to clarify this.

EDITOR: I’m not sure I understand what “referenced for guidance during SPENT development” means. Are you going to use these articles? And if so, why are they not just included? Excluding guidelines that are specific to a condition/population/treatment is sensible if the review is expected to be large, but if your review is empty/nearly empty, these items could prove useful. I do wonder if they should be included regardless of size of the review, as they may highlight areas that are problematic and dependent on the trial specifics, and as such need special attention or flexibility in a generic protocol reporting guideline.

RESPONSE: Thank you – we agree. We have removed that sentence from our exclusion criteria. The next sentence then became irrelevant and was also removed. (lines 137 to 142)

EDITOR: Might guidelines be contained within book chapters and grey literature reports, e.g. for research and/or funding bodies? Have you a strategy for finding such literature?

RESPONSE (a): Book Chapters: We are not including book chapters, as we would like to focus on the peer-reviewed literature. Textbooks as references will be included, as per #2, line 148.

RESPONSE (b): Grey Lit reports: We are contacting research & funding bodies (31 identified). This detail has been added to the methods, as per #3, lines 148 to 150, “Requests for guidelines specific to N-of-1 trials will be sent to 31 large private and public research and funding organizations in North America, Europe, and Australia.”

EDITOR: The phrase “(SP) will spot-check for extraction accuracy” is not clear. I think you mean you are just checking a sample? If just a sample, what proportion? Why have you chosen 20% as an acceptable level of disagreement?
RESPONSE: Thank you – we have clarified (lines 193 to 195): One reviewer (AP) will extract, and a second (SP) will verify data quality by checking a 15% sample from each paper for extraction accuracy.”

We based our process on “Shamseer L, Stevens A, Skidmore B, Turner L, Altman DG, Hirst A, Hoey J, Palepu A, Simera I, Schulz K, and Moher D. Does journal endorsement of reporting guidelines influence the completeness of reporting of health research? A systematic review protocol. Systematic Reviews. 2012, 1:24”, where their process was verification of 10% of studies by the second reviewer, and 20% of specific items; they required 50% discrepancies for a full review.

EDITOR: The analysis section seems incomplete. What will be done with the long list of items? How will the review become a reporting guideline? Whilst these aspects are not specifically the output of the review, it would be useful for readers to know what the next planned steps are. E.g. are you going to conduct a Delphi exercise? Indeed, contact with experts is notably lacking in this protocol – your search strategy at the least would benefit from such contact.

RESPONSE: Thank you. Additional information on the use of the results has been added, lines 209 to 212, “All protocol-specific results and extracted qualitative-identified topics will form the starting list for the development of SPENT. This list will be assessed and refined using a sample of published and unpublished N-of-1 trial protocols, followed by an international Delphi process.” and lines 242 to 243 the discussion summary, "Specifically, the results will form the basis for the Delphi process needed to assess and develop possible items in SPENT."

Additionally, the author team is largely comprised of experienced N-of-1 trial and reporting guideline development authors/leads; missing information has been added to the Authors’ Information, lines 285 to 287. As mentioned previously, we are contacting international research and funding agencies for guidelines, and other N-of-1 protocol authors for sample protocols.