Author's response to reviews

Title: LOST to follow-up Information in Trials (LOST-IT): a protocol on the potential impact

Authors:

Elie A Akl (elieakl@buffalo.edu)
Matthias Briel (mbiel@uhbs.ch)
John J You (jyou@mcmaster.ca)
Francois Lamontagne (francois.lamontgne@usherbrooke.ca)
Azim Gangji (gangji@mcmaster.ca)
Tali Cukierman-Yaffe (cukierm@mcmaster.ca)
Mohamed Alshurafa (alshurm@mcmaster.ca)
Xin Sun (sunx26@mcmaster.ca)
Kara A Nerenberg (nerenbk@mcmaster.ca)
Bradley C Johnston (bjohnston@med.ualberta.ca)
Claudio Vera (cmverapg@med.uc.cl)
Ed Mills (emills@ccnm.edu)
Dirk Bassler (dirk.bassler@med.uni-tuebingen.de)
Arturo Salazar (esalazar@med.wayne.edu)
Neera Bhatnagar (bhatnag@mcmaster.ca)
Jason W Busse (j.busse@utoronto.ca)
Zara Khalid (khalidz@mcmaster.ca)
Stephen D Walter (walter@mcmaster.ca)
Deborah J Cook (debcook@mcmaster.ca)
Holger J Schunemann (hjs@buffalo.edu)
Douglas G Altman (doug.altman@cancer.org.uk)
Gordon H Guyatt (guyatt@mcmaster.ca)

Version: 2 Date: 2 June 2009

Author's response to reviews: see over
May 23, 2009

Editors-in-Chief, Trials

RE: “LOST to follow-up Information in Trials (LOST-IT): a protocol on the potential impact”

Dear Editors,

We thank you for the opportunity to revise our manuscript for Trials. The reviewer’s comments were helpful in improving the quality of our manuscript. Please find on the following pages our detailed point-by-point responses.

With kind regards,

Gordon Guyatt, MD, MSc
Departments of Medicine and Biostatistics & Clinical Epidemiology
McMaster University
The reviewers’ comments are in **bold** font and our replies in regular font. Extracts from the text are in *italic* fonts with changes to existing sections *underlined*.

**Reviewer: Andrew Vickers**

**Reviewer’s report:**

a. I have previously written several methodologic reviews and do believe that they are valuable. I have some concerns about this proposal, however.

b. In a general sense, the protocol left a bit of a bad taste in my mouth: it felt a little snarky and superior, and I could almost feel the investigators chomping at the bit to publish a paper on just how badly trials were reported and conducted, even in the very best journals. Perhaps this started from the title, which really sets the tone. “Lost it” has a negative connotation, and I could almost hear the methodologists thinking, “those trialists have really lost it”.

We are sorry the reviewer experienced the protocol as snarky or superior; our intent was quite different. As stated in the protocol, our primary objective is to assess the potential impact of loss to follow-up on the estimates of treatment effect of positive studies published in the top 5 journals. It is possible that our findings will demonstrate that the treatment effect in most of these studies is robust even with extreme and unlikely assumptions about patients lost to follow-up. Our discussion presents the possible outcomes as follows:

“Our findings will uncover the potential impact of plausible assumptions about the outcome of participants LTFU on the results of those positive studies that are most likely to affect clinical practice. They might provide reassurance that these results are usually – at least for reports in five prestigious journals - robust or, on the contrary, suggest that many high-profile trials are vulnerable.”

The name LOST-IT is not based on any sense of sarcasm; it is based on (our attempt at) a sense of humor. Our group has been working on a project called Study Of Trial Policy Of Interim Truncation (STOP-IT) (see: http://jama.ama-assn.org/cgi/content/abstract/294/17/2203). We thought it would be humorous to give a similar name to our current project.

Many of the authors of this study are trialists themselves. We have made a number of changes to the text in order to eliminate any hint of superior tone. Many of these changes are detailed below.

c. Something else that contribute to what I felt was a rather inappropriate tone was the review of previous studies on missing data. The authors seemed to be saying: “gosh, there is a lot of missing data (who’d of thought it?) and investigators don’t do much about it”. I think that the previous studies were a little silly. It makes no sense whatsoever to give an estimate of the proportion of studies with missing data because, in my view, a study is either going to have missing data or not, and there isn’t much you can do about it. A study looking at overall survival will not have missing data, neither
should, say, a trial in migraine headache assessing the immediate effect of an analgesic. Any longer term study with patient-reported outcomes will have missing data and you can’t avoid it (I say this as someone who has published research on how to reduce missing data: yes, you can reduce rates of missing data, you can’t eliminate missing data). An estimate of the number of studies with missing data simply reflects the type of studies in the sample.

Our review of previously published methodological studies is an attempt to put ours in context. We strived to report this literature review in an objective and non judgmental way.

We completely agree with Dr. Vickers that trialists can reduce loss to follow-up, but in most instances cannot eliminate it. This is exactly why we take this issue to the next level and, as it is reflected in our primary objective, we aim to assess the potential impact of loss to follow-up on the estimates of treatment effect. We have added the following text to the background section:

“Investigators can reduce the amount of missing data, but in most instances cannot eliminate it [2]. How to best deal with missing data remains controversial;”

Dr. Vickers makes an excellent point about the association between the type of an outcome and the timing of its assessment on one hand, and loss to follow-up on the other. We have integrated it in our discussion section as follows:

“The type of outcome and the timing of its assessment are likely to be associated with the extent of loss to follow-up. Loss to follow-up is, for instance, likely to be lower for survival compared to a measurement of quality of life. Similarly, in a trial of an analgesic, loss to follow-up is likely to be lower for pain at day 2 compared to pain at 1 year. We are thus categorizing the type of outcome and recording the timing of its assessment.”

d. Second, the complaint that only one in 4 trials adjusted for missing data is misplaced. I have published numerous randomized trials with missing data, but have only adjusted for missing data occasionally. This was sometimes because studies had obviously negative results and other times because the rates of missing data were pretty low.

We assume that Dr. Vickers is referring to the following statement related to the study by Baron et al.:

“Baron et al. examined the rate of missing data in superiority trials published between 1994 and 2003 that assessed structural outcomes in rheumatic diseases (n= 81)... Only 24% reported statistical methods for handling missing data.”

This statement simply reports the results of a previous investigation. However, we agree with Dr. Vickers regarding the “negative results”. We have added the following text to the above statement:
“Although missing data is not as problematic in interpreting the results of an individual trial with statistically non-significant effect estimates, none of the three aforementioned studies distinguished between negative and positive studies when reporting about handling missing data.”

In addition, a major exclusion criterion in our study is: “primary outcome is not statistically significant”. We have expanded our rationale for this decision as follows:

“We will focus on reports of trials with statistically significant effect estimates published in major general medical journals because these studies are more likely to influence clinical practice than studies with non-significant effect estimates or studies published in lower profile journals. In addition missing data is not as problematic in interpreting the results of an individual trial with statistically non-significant effect estimates.”

On the other hand, we have not defined any cutoffs to categorize the extent of loss to follow-up (e.g. low vs. high) for 2 reasons: (1) we believe that the gravity of loss to follow-up relates mainly to the extent it jeopardizes the inferences one might appropriately draw from a study, and not to absolute numbers of loss to follow-up, and (2) we aim to check the potential impact of loss to follow-up on the estimates of treatment effect among all positive studies published in the top 5 medical journal, irrespective of the extent of loss to follow-up. For study results that remain robust given even extreme assumptions, we will conclude that statistical methods for handling missing data are irrelevant.

e. How will the authors deal with the fact that both the proportion of studies with missing data, and the proportion that corrected for missing data, are essentially uninformative statistics that reflect the types of study in their sample and their results?

We believe these proportions are valuable because they inform us on the extent of the problem of missing data in positive trials published in major general medical journals that are likely to influence clinical practice. However, we would like to re-emphasize that our primary objective is not to estimate the proportion of studies with missing data or the proportion that corrected for missing data. It is to assess the potential impact of loss to follow-up on the estimates of treatment effect among positive studies published in the top 5 medical journal.

f. On a more methodologic note, I think the investigators are painting themselves into a corner. They are only looking at binary endpoints so they can do some simple imputations (about which, see below). But perhaps the most interesting areas of missing data concern patient-reported outcomes, and this are often continuous (e.g. pain, depression). I see no reason to exclude these studies from the descriptive analysis (e.g. proportion of trials that discussed implications of loss to follow up) just because there are not susceptible to some additional analyses.

We agree with Dr. Vickers that exploring our research question in studies reporting continuous outcomes would be interesting and important. As he suggests, the analytical approach to continuous outcomes is different from that to binary outcomes. We thus hope to
eventually conduct a separate study for continuous outcomes. We are identifying those studies in the process of selecting studies for the current project. We have added the following text to the discussion section:

“We have specifically chosen not to include RCTs reporting primary outcomes that are continuous variables or expressed as rates, because of specific challenges related to analyzing and reporting LTFU information in those trials, e.g. use of last value carried forward with continuous outcomes [3]. We hope to conduct a separate study for continuous variables.”

g. On which point, the imputations suggested by the authors are mainly just silly. Assuming that all patients lost to follow-up had the event, or didn’t have the event, or had the event in one group and didn’t in the other, is extremely unsophisticated and well outside accepted practice in missing data analysis. As a simple example, imagine that there was a trial with 500 patients per group, follow-up data on 480 in each group, and 25 and 10 events in control and treatment respectively (5 vs 2%, p<0.01). On what possible grounds would it be realistic to say that the event rate in those lost to follow-up was 50 fold higher than those analyzed?

One of the assumptions the reviewer refers to is the worst case scenario. This is an extreme, and in most cases unrealistic, assumption. However, it is valuable in demonstrating the robustness of study results when one is able to show that the effect estimate from a positive study remains statistically significant even under the assumption of a worst-case scenario. We have added the following text to the discussion section:

“The 3rd assumption (worst case scenario) is an extreme and in most cases unrealistic assumption. However, it is valuable in demonstrating the robustness of trial results when the effect estimate remains statistically significant even under the assumption of a worst-case scenario.”

Many of the other scenarios (e.g. assuming that all patients lost to follow-up had the event or didn’t have the event) are commonly recommended in the textbooks of clinical epidemiology or used by trial authors (as we found in a preliminary review of a number of eligible trials) to either impute missing data or test the robustness of their results.

While other imputation methods such as regression models and multiple imputations might be preferable, they are not feasible in the context of our study because they would require raw data from each included trial.

h. The normal method of adjusting for loss to follow-up is to build regression models. For example, imagine that in the hypothetical study, the event rate in older patients was 8% and in younger patients 0.5%. One could use the distribution of ages in the patients lost to follow-up to make some guess about likely event rates. Multiple imputation takes this type of method a step further by using a simulation approach. The point is, it is not possible to use these methods on summary data, you need the raw data.
We agree with Dr. Vickers. It would be ideal to use regression models and multiple imputations. It is unrealistic, however, to obtain raw data from eligible studies in order to conduct these analyses. We have added the following text to the discussion section:

“While other imputation methods such as regression models and multiple imputations might be preferable, they are not feasible because they would require raw data from each included study.”

i. Some other thoughts, problems:
1. There are surely some cases where there are missing data, these are unlikely to influence the results of the trial, and the authors ignore these (e.g. a small amount of missing data in an unambiguously negative trial). Do the authors of the review really want to count such trials as a “problem” (no account taken of missing data)?

We agree with Dr. Vickers. Please refer to our response to point d.

2. On page 19: how is “appropriate” and “inappropriate” post-randomization exclusion defined?

The definitions of appropriate and inappropriate exclusions are included in the methods section under “Categories of loss to follow-up” as follows:

“Mistakenly randomized subjects are those that were ineligible at the time of randomization. We will consider their post-randomization exclusion inappropriate if at least one the two following conditions is not satisfied: (1) the information about ineligibility was available at randomization; and (2) the individual making the exclusion decision was blinded to allocation. For participants who did not receive the intervention we will consider their post-randomization exclusion inappropriate if patients were not blinded to their allocation. We will consider all post randomization exclusions related to study center exclusion as appropriate.”

These definitions were also included in table 3 and its footnotes.

3. On page 10, the comments about survival analysis are misplaced. The 2 x 2 table would indeed give you a risk, however, the confidence intervals could not be calculated from the 2 x 2 table.

Our options to deal with studies reporting survival data were either to include them or exclude them. Given the common use of survival analysis in high-profile trials, we opted to include them and use the 2x2 table approach. Acknowledging the limitation of this approach, we planned to conduct sensitivity analyses excluding these studies.

In order to calculate the 95% CI for the relative risk, we will use the following standard formulas:
- lower limit= the exponential of (log(Rel risk)-(1.96*SElogR))
It is true that the resulting confidence interval, similarly to the point estimate, is not going to be the same as if we had the survival data. We have modified the following statement to account for this:

“We acknowledge that the proportion we are calculating for time to event data (number of subjects with events divided by number randomized) as well as the upper and lower values of its confidence interval are not exactly risks and plan a sensitivity analysis excluding these trials.”

4. The start of the article is very odd, starting with a “note” about a definition.

Thank you for the advice. We have changed the text as follows:

“In this paper, we define loss to follow-up as…”

5. Table 1 is misformatted

Thank you for noting this. The table is appropriately formatted in the Word file we submitted. We will make sure that it doesn’t become misformatted when it is converted to a PDF file on submission.

6. Page 21: how is “type of intervention” categorized?

As shown in the data abstraction form (additional file 3), the categories for type of intervention are the following: pharmacological, surgery/invasive procedure, rehabilitation, behavioral intervention, complementary and alternative medicine, diagnostic test, other

We have now clarified in the text that in the regression analysis, we will use type of intervention as an independent variable with 3 categories: pharmacological vs. surgery/invasive procedure vs. other.

7. Page 18: Loss to follow-up is a strange concept in a survival study, “censoring” occurs because some patients are followed longer than others. However, throughout the protocol, it is as though a patient not followed through, say, 5 years, would be counted as “lost to follow up”.

Thank you for noting this unclear point. We are distinguishing between those patients who are censored because of the planned termination of study follow-up and those patients who are lost to follow-up for reasons unrelated to the planned termination of study follow-up. We have clarified this in the methods section as follows:
“For survival data, we will not consider as LTFU those subjects that are censored because of the planned termination of study.”