Reviewer’s report

Title: Online Randomized Controlled Experiments at Scale: Lessons and Extensions to Medicine

Version: 0 Date: 14 Jul 2019

Reviewer: Munyaradzi Dimairo

Reviewer's report:

This is an interesting manuscript and I applaud the authors for their effort. However, I believe this could be improved to add its value to readers. For instance, lessons learned from tech companies should be prominent beyond just describing what is being done by these companies. Furthermore, some discussion on implications should be covered (e.g., ethics and regulatory issues).

1. The background does not acknowledge or reflect that there are several ongoing initiatives to make RCTs more efficient (less costly, etc) and more generalizable. For instance, embedding RCTs within healthcare systems, the use of routinely collected data, trials within trials, etc. This point is ignored and should be reflected. Of course, there is more to learn and obstacles to address - and this is where this manuscript comes into play.

2. I feel that ethical implications should be reflected throughout. For example, in the example given on page 4, it would be very helpful to highlight if there are any ethical implications (e.g., consent issues etc) when applied in biomedical RCTs including how these companies address these issues (if they do).

3. The structure of the paper could be improved to enhance readability and flow. For instance, the section on statistical power has got a number of subthemes (some of which have nothing to do with statistical power) which can be distilled to help readers. Itemised information under "Several changes typically happen:" could be presented differently (e.g using a box, etc)

4. Are there examples of changes that were made as informed by an experiment but results were disappointing when they were rolled out? If so, it is worth highlighting as a risk trade-off.

5. This statement raises a key issue that could be reflected. One could just increase the sample size to get a very small p-value for a negligible or tiny effect. As stated by the authors, these experiments are being conducted on very large sample size "The rest of the metrics are used for debugging and understanding why something happened, and these are marked interesting when the p-value is low, such as &lt;0.001"

6. I am struggling to understand the point being communicated here because the allocation ratio does not need to be strictly one for an experiment to give valid results. Is this about potential issues with randomisation code or something else? "For example, suppose the experiment design calls for equal assignment to control treatment, and the actual number of users is Control: 821,588 users, Treatment:
815,482 users, and thus the ratio is 50.2% instead of 50%. The system would flag this as a sample-ratio-mismatch and declare the experiment result invalid, as the p-value for such a split is 1.8e-6.

7. There is a missing reference here "Some interesting topics are covered in Error! Reference source not found."

8. This section needs expanding to be of value to readers "Some additional technical issues". For instance, you could summarise the issues in a table or box, then signpost readers for additional details.

9. I am not sure this statement is entirely correct "There is less emphasis in medical RCTs on performing multiple RCTs at the same time and engaging the same participants in multiple concurrent RCTs". I could argue otherwise and that is why (adaptive) platform trials have gained traction.

10. This important section needs expanding "Similarities and dissimilarities with medical RCTs". Perhaps a tabular presentation would help readers. In addition, ethics and regulatory aspects should be reflected. Readers should be informed regarding the implications.

11. I suggest you include a few sentences on how patients are randomised in A/B trials being conducted by tech companies. In addition, are the other principles of RCTs upheld in these trials beyond just the randomisation of participants?

12. Worth highlighting key points of criticisms here? "Recently, criticism of randomized trials in medicine seems to be growing."

13. Box 1 should be referenced in the text.

14. I feel that the manuscript did well in describing the perspectives of tech companies. However, there seems to be little information on lessons learned and how these could be adopted in biomedical trials research. For instance, in the summary section, lessons learned are key to readers so they should be presented in a manner that is easier for readers.

Minor/pedantic comments

- MVP has been defined a number of times and alternating use of acronym and full definition. The same thing with OEC. Perhaps you just need to define it when first used then use acronym thereafter
Level of interest
Please indicate how interesting you found the manuscript:

An article of importance in its field

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Quality of figures
All images and figures within the manuscript should be genuine i.e. without evidence of manipulation. No specific feature within an image may be enhanced, obscured, moved, removed, or introduced. If you have concerns about the veracity of the figures you should choose the first option below.

Statistical review
Is it essential that this manuscript is seen by an expert statistician? If so, please give your reasons in your report.

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.
I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal

Were you mentored through this peer review?

No