Reviewer’s report

Title: Pragmatic Trials and Implementation Science: Grounds for Divorce?

Version: 0 Date: 21 Jan 2019

Reviewer: Peter Bower

Reviewer’s report:

As a pragmatic trialist, I found this piece on the relationship between pragmatic trials and implementation research both intriguing and annoying in equal measure, which I would consider a success in terms of its aim of generating debate.

The piece starts by considering the function and interpretation of pragmatic trials. This was generally appropriate, although I was not entirely convinced that the function of pragmatic trials was primarily to be generalizable across settings? Surely, the primary function was to encourage trials which better represented particular contexts into which results were designed to be implemented (even though they would face the usual interpretative challenges when taken to new contexts)?.

For example, pragmatic trials of psychological therapy in the UK (where I cut my own research teeth) were designed to avoid many of the supposed strengths 'explanatory' trials (such as highly selected populations, specialist therapists, and highly standardised treatments), so that outcomes would be more easily applied to the hurly-burly world of routine clinical practice in the UK (which would involve a heterogeneous mix of patients, variable therapist quality, and variation in adherence to and delivery of treatment protocols). The idea was that policy makers could be reasonably confident that whatever benefits might be demonstrated in the pragmatic trial might be found if the same service was widely commissioned. I am not sure there was ever an assumption that such results would be more widely generalised to ALL contexts, such as therapy delivered in the United States (although the interpretative gap would still be smaller than with a traditional efficacy trial). Therefore, I was not entirely convinced of the point being made.

The other point which caused a little concern is the repeated statement that pRCTs are somehow restricted to application 'in which exactly the same contextual conditions apply'. I am not sure critics of trial methodology have fully explained the core dimensions of context by which we judge whether the 'same configuration of contextual conditions' applies (although I know there are models out there), nor does it seem to me that the empirical case has been made that the impact of context is so strong as to make slight variation in context somehow a critical driver of trial outcomes. A lot of the language here is fairly extreme - is it really true that the EXACT same contextual conditions have to apply? Again, on page 7, the statement about 'widespread applicability' shrinking to 'case to case isomorphism' seems extreme - surely there is a reasonable middle ground?

The Zwarenstein example is useful in helping the reader grasp some of the issues, although in some ways it is an odd example - since the intervention is SO simple, it would not necessarily
seem something that would be impacted by different contexts in the same way as a 'complex intervention'. I wondered about the choice of that exemplar when there would be so many others that would seem to be more relevant to the argument being presented.

Part 2 explores the 'case for reconciliation' and suggests a methodological approach to overcome some of the issues that the author has identified, which goes beyond the general call to include qualitative research alongside trials. Part 2 is useful and detailed and adds significant value to the paper.

I was interested that the author focussed on characteristics of the practitioners as the exemplar here, with a nod to 'heterogeneity of treatment effect' (HTE) analysis. All the proceeding discussion had been around context, but here we are suddenly looking at individual difference subgroups, which is not the norm in this area, at least to my knowledge. This may reflect the example he has chosen, and is not problematic per se, but I thought it was worth noting. Three of the four examples at the end of the piece also focus on professional subgroups.

The general point about the need to see variation as the natural focus of analysis of trials is coherent, although not without problems. Smith et al's (1980) discussion of one of the first meta-analyses ever done deals with this point in detail, and highlights the fact that although there may be interest in variation and 'what works for whom' to a variety of decision makers, the general impact of a programme across multiple settings and types of providers is still a highly relevant question - irrespective of whether it hides important variation in effects - and often the question of relevance to policy makers.

The idea that HTE are a valid perspective for the analysis of trials is also not unproblematic.

One of the issues is that we do not necessarily know in what way to break up groups of professionals and patients - there are so many potentially relevant variables. Of course, the same applies to contexts. This does not seem like a trivial issue. The examples given are all plausible and coherent, but how those are judged over a wealth of competing approaches is not clear to me.

Secondly, it needs to be acknowledged that our ability to assess HTE is very constrained by mundane logistical and methodological issues - we may be unable to determine reliably whether such effects are legitimate, and may either dismiss important effects, or highlight chance findings. This needs to be acknowledged if there is an assumption that this sort of 'programme theory' will have a quantitative test at some point in the process. However, this was not clear to me, and the author needs to be more explicit on this issue. Is there an expectation that the methodological process here will see quantitative analysis being done differently? At present, there is a vague statement that 'data is generated' - but he may be deliberately leaving that open, although the comment on page 12 about referral rates implies a quantitative test. That is fine, but if a quantitative test is planned, there needs to be at least a nod to the substantial methodological and logistical challenges, particularly as they may be so acute as to render this model impractical. 'What works for whom in what context?' is a well-known phrase, but actually delivering on that level of specificity has proven difficult.
In summary, this is a well written paper which raises a number of important issues in an engaging style, presenting a clear model for how some of the problems could be overcome. I have identified some areas where I think arguments are perhaps a little overstated, and where some of the challenges have been minimised, and it might be useful if those were acknowledged.

Minor issues

The statement on page 6 about the opportunity to mount an efficacy trial is an odd one, as I am not aware that anyone is suggesting that is what we need to do in knowledge transfer.

Figure 2 did not seem to add significant value - I found the text description entirely adequate.

Page 11 typo 'standerised care movement'

Page 13 - HTE should refer to 'heterogeneity of treatment effects' rather than 'outcomes'

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

Unable to assess

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Unable to assess

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

Yes

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

Not relevant to this manuscript

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

Acceptable

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