Author’s response to reviews

Title: The impact on healthcare, policy and practice from thirty-six multi-project research programmes: findings from two reviews

Authors:

Stephen Hanney (stephen.hanney@brunel.ac.uk)
Trish Greenhalgh (trish.greenhalgh@phc.ox.ac.uk)
Amanda Blatch-Jones (A.J.Blatch-Jones@soton.ac.uk)
Matthew Glover (Matthew.Glover@brunel.ac.uk)
James Raftery (j.p.raftery@soton.ac.uk)

Version: 1 Date: 06 Mar 2017

Author’s response to reviews:

The impact on healthcare, policy and practice from thirty-six multi-project research programmes: findings from two reviews: Response to reviewers’ comments

HRPS-D-16-00177

We are most grateful to both reviewers for their very positive comments and helpful suggestions, all of which we have been pleased to incorporate to further strengthen the article. These changes are shown on the revised text in tracked changes, which look quite substantial but which include the two suggested re-ordering of parts of the text, and a few minor corrections/amendments that we have added. Below we explain how we have addressed each point.

Reviewer 1: Simon Innvær

Recommendations for the discussion section: 1. Discussions with potential users and mechanisms: In the results section, page 7, the authors refer to an "extremely high impact …on policy" in Canada (100 percent) and Austria (97 percent), while in the Netherlands the impact was only 29 percent. This is one of the most remarkable results in the paper, but I find no
discussion on this. On what, how, why, when and for whom do the results depend? Are they due to differences in "discussions with potential users" or "mechanisms" referred to by the authors at page 10? If this is not so, or if it is impossible to find such "mechanisms", the value of the results needs to be discussed in the well written and argued sections about the limitations of the review (page 8-9). Since "discussions with potential users" and "mechanisms" seem to be of high importance, the authors should include more about these results in the studies. How and why are discussions important, and what are the mechanisms? What might the results be if the only included studies were those where mechanisms were described in an explicit way? RESPONSE: We have addressed these important points by adding the following text in the part of the Discussion where the relevance of the theoretical models about the importance of discussions with users and of mechanisms is described and the importance of the type of research considered:

‘These points can be illustrated by some of the more notable examples from Table One. For example, in the instance of the 100% of the HTA reports from the HTA unit of McGill University Health Centre in Quebec, Canada, the impact was said to be because of ‘(i) relevance (selection of topics by administration with on-site production of HTAs allowing them to incorporate local data and reflect local needs), (ii) timeliness, and (iii) formulation of policy reflecting community values by a local representative committee’ [52, p.263]. In the case of 97% of the assessments from the Austrian HTA programme being classified as making at least some impact on coverage policies [66], there were features of the Austrian policy making structures that facilitated the use of HTA reports. The authors explained that to be used, the HTA reports ‘need primarily to be in German language and they have to be produced within a time period that is strongly linked to the decision-making process.’ [66, p.77] By contrast, and as noted above, while the Health Care Efficiency Research programme from the Netherlands was also classified as an HTA programme, it included a large responsive mode element and most studies were prospective clinical trials rather than the technology appraisal reports that are the main element of many HTA programmes [55]. The lower proportion of projects in these studies (29%) demonstrating a policy impact illustrates that variations in levels of impact achieved can be linked to the type of research conducted, even in the same overall field, which in this case was further exacerbated by the impact assessment occurring soon after the end of the trials [55].

Overall, as Jacob and McGregor reported for the HTAs conducted in Canada by the Quebec CETS, ‘The best insurance for impact is a request by a decider that an evaluation be made’ [48, p.78]. Furthermore, for those programmes (or parts of wider programmes) for which there were explicit mechanisms such as formal committees to receive and use the findings from technology appraisal reports in coverage decisions about investment or disinvestment, the proportion of projects making an impact was very high.’
2. The impact frameworks: It is not perfectly clear whether the impact frameworks, referred to on page 9ff, come from the studies included in the review or from a previous paper from four out of five of the authors of this review (reference number 8). Maybe it's both? RESPONSE: The answer is slightly complicated in that the six framework described are the key ones featured in the article by Greenhalgh et al, but also they all have some link to studies included in this analysis. We are pleased, therefore, to clarify this by briefly explaining the point in the following text added at the point where the frameworks are introduced:

‘Each of the first five of the six frameworks highlighted by Greenhalgh et al [8] helped inform at least one of the 36 studies in this current analysis, and, according to the Higher Education Funding Council for England (HEFCE), the sixth, ie the REF, was itself partly informed by studies applying the Payback Framework [72].’

3. Standardisation (page 11-12): The authors seem to argue for more "standardisation in terms of [the research impact assessments’] timing and methods" (page 11), but the argumentation for this is thin, since "such standardisation can itself generate further difficulties" (page 12). What "methodological inconsistencies" and "strategic decisions" can be solved by standardisation? RESPONSE: We are grateful for this observation which has led to us strengthening the discussion of standardisation by the addition of the following text:

‘Any standardisation of methods might attempt to reduce the current diversity on items such as the categories of impact to include and their definition, and the timing of data collection and its presentation. Such moves towards standardisation might facilitate comparisons between processes used in different programmes and in that way inform strategic decisions that funding organisations will always need to make as to how best to use resources.’

Reviewer #2: Rita Banzi

Background

1. I would suggest to clarify what this paper adds to the reviews previously published by the same group. I understand the focus is on multi-project assessment exercise and I think the reader could be interested in knowing why a specific summary of studies that analyses this sample deserves a new publication. Some key aspects are already covered in the discussion/conclusion sections for example the issue of developing standard for the research impact evaluation.
RESPONSE: This point has been helpfully addressed by elaborating the text in the final paragraph of the background so that it now provides a fuller picture:

‘For this paper we deliberately sought studies that had assessed the impact of all projects in multi-project programmes, whether coordinated or not. We focused on such multi-project programmes because this approach offered the best opportunities for meaningful comparisons across programmes both of the methods and frameworks most frequently used for impact assessment, and, crucially, of the levels of impact achieved and some of the factors associated with such impact. Furthermore, such an approach focused attention on the desirability of finding ways to introduce greater standardisation in research impact assessment. However, we also discuss the severe limitations on how far this analysis can be taken. Finally, we consider the implications of our findings for investment in health research and development (R&D) and the methodology of research on research impact.’

Methods

2. The methods section would probably become clearer if it starts with the information about the methodology applied in the current review.

I found the paragraph "search strategy of the two original reviews" a bit misplaced. It reports the search strategies the previously published reviews and a brief description of the main findings. I support the idea to shortly describe the two original reviews, but I think this information (e.g., from page 3 line 39 to page 4 line 25) can be included in a dedicated box.

RESPONSE: We have implemented this helpful suggestion.

Results

3. The authors should explain if the "Factors associated with level of impact; comments on methods and use of the findings" reported in table 1 are those defined by the authors of the original reports as key factors that might help explain the level of impact achieved (page 7, line 25-34) or if this evaluation has been done by the review authors.
RESPONSE: We have made this helpful clarification in two ways. First, we have elaborated the relevant text in the brackets referring to the last column of Table One: ‘(see last column in Table One for direct quotes, or comments that in most cases came from the original paper).’ Second, later in the text we have taken a quote from, and referred to, the final column of Table One to illustrate the point about the factors linked to the impact achieved by the three of the four non-HTA studies where over 50% of the projects were thought to have made an impact on policy:

‘for example, Reed et al said that the figure of 53% of projects from a programme of primary care research in Australia making an impact on policy and organisational decisions reflected, ‘a high level of engagement of the researchers with potential users of their research findings.’ [57, p.5] (See Table One for further details).’

4. The authors could extract from the papers included one example of tangible impact for each category of impact they used in their conceptualization. This would help defining and understanding these categories better.

RESPONSE: We have implemented this important suggestion by adding the following text immediately after Table Two:

‘One example from the various studies can be used to illustrate what is included in each of the four types of impact. The 1997 study by Jacob and McGregor [48] reported that 86% of the HTAs conducted in Canada by the Quebec CETS had influenced policy. One of these HTAs found that the likelihood of health benefits from routine preoperative chest radiography were extremely slender; prior to the publication of that HTA report 55 out of 118 hospitals questioned had a policy of using such routine chest radiography, but three years later all but three had abandoned this policy and in 79% of cases the HTA was cited as a reason for the policy change. In terms of impact on practice, in 2007 Kwan et al gave the following as an example of the local impact on provider behaviour made by the health and health services research programme in Hong Kong: ‘improved reporting of unintentional child injury cases and liaison between the Hospital Authority Informatics and Accident and Emergency’ [51, p.8]

Illustrating the combined category, Milat et al [53] used a category called ‘Policy and practice impacts’ in their 2013 assessment of the impact from the research funded in Australia by the New South Wales Health Promotion Demonstration Research Grants Scheme. While the analysis provided overall figures only for this combined category, the few examples were that were given were presented separately for policy impacts and practice impacts. In some, but not all, instances
the accounts covered both dimensions, for example research informed policy planning by identifying areas for investment in tai chi for older people (as a way of preventing falls) and smoking cessation brief interventions. Then in terms of practice the research in those same two areas helped inform professional development for the relevant staff providing the services. An example of health gain comes from one of the NIH trials analysed in the 2006 assessment by Johnston et al [49] described above. Johnston et al estimated that implementation of the findings from the trial of the use of Tissue Plasminogen Activator in cases of acute ischemic stroke, published in 1995, had a projected health gain in the 10 years after funding was completed of 134,066 quality-adjusted life years.’

Discussion

5. The authors correctly acknowledged the difficulties in making a clear distinction between the studies included and excluded in this analysis. Thus, it would be useful if the authors report the list of the studies initially considered relevant and then excluded because the data cannot be "combined" with the others (page 8, line 44-48).

RESPONSE: We have implemented this suggestion by the addition of a new Table Three, which can be seen in the revised text, and is introduced by the following text added to the end of the first paragraph on the limitations (which is now in its new position in line with suggestion 6 below):

‘The seven studies [80-86] assessing the impact of multi-project programmes that were included in the two reviews on which this study was based, but excluded from this current analysis, are listed on Table Three, along with reasons for their exclusion.’

6. I would suggest a change in the paragraph order to report first the main findings of the review (page 9, line 48 to page 10, line 25), then the limitations (page 8, line 43 to page 9, line 46).

RESPONSE: This helpful suggestion has been implemented. Thank you.