Author's response to reviews

Title: The effectiveness of community engagement in public health interventions for disadvantaged groups: A meta-analysis

Authors:

Alison O'Mara-Eves (a.omara-eves@ioe.ac.uk)
Ginny Brunton (g.brunton@ioe.ac.uk)
Josephine Kavanagh (j.kavanagh@ioe.ac.uk)
Sandy Oliver (s.oliver@ioe.ac.uk)
Farah Jamal (f.jamal@uel.ac.uk)
James Thomas (j.thomas@ioe.ac.uk)

Version: 2 Date: 30 September 2014

Author's response to reviews: see over
We would like to thank our three reviewers for their careful and extremely helpful comments. We
detail our responses to each in the table below, and are grateful for their insights which we hope
have helped us to clarify the paper.

Reviewer 1 (Lisa Jones)

(Minor essential)

<table>
<thead>
<tr>
<th>Issue</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I am concerned that the scope of the manuscript is a little too broad and that consequently some of the important findings get rather lost. Some re-organisation and editing of the manuscript (particularly the Results section) to make it more focused would help - for example, there is some repetition of the methods throughout the Results (e.g. 4th paragraph, pg 15) and reporting of effect sizes both in the text and in tables.</td>
<td>We have removed the effect sizes from that paragraph, and hope that the changes we have made in several parts of the results section, now makes this more readable. We have also added more ‘signposting’ in order to help the reader navigate the results.</td>
</tr>
<tr>
<td>2. Description of data/additional analyses. The statistical methods used in the subgroup analyses and additional analyses are not clearly reported in the Methods. Both random effects ANOVA analyses and random effects meta-regression analyses are reported but I could not find any details about the effects meta-regression analyses in the Methods section.</td>
<td>Our methods do cover the ANOVAs and meta-regression, though we would be happy to add more detail if that would be useful: “We conducted random effects model analyses (ANOVAs and multiple regressions) with maximum likelihood estimators, following the methods described in [16]. We used SPSS macros written by David Wilson(^1) to run the models. For the homogeneity analyses, between groups Q-statistic (Q(_B)) indicates the extent to which the categories of studies differ from each other; and within groups Q-statistic (Q(_W)) indicates the extent to which the effect size estimates within a category differ from each other. Analyses were conducted separately for post-test measures and follow-up measures. Analyses were also conducted separately for the different outcome categories (health behaviours, health consequences, self-efficacy, and social support). As such, each study only contributed one effect size estimate to each analytical model.”</td>
</tr>
<tr>
<td>3. Study characteristics. The characteristics of the included studies are only presented in the text, which makes it difficult for the reader to gain an understanding of the types of the studies included in the meta-analysis. As a number of subgroup and additional analyses are presented it would be useful to be able to identify which studies contributed to which</td>
<td>We have added some more detail about the included studies in the results, but are unsure whether it would be sensible to include a full summary of included studies table. (This table is 240 pages long in the NIHR report, and the risk of bias table runs for an additional 15 pages.)</td>
</tr>
</tbody>
</table>

\(^1\) [http://mason.gmu.edu/~dwilsonb/ma.html](http://mason.gmu.edu/~dwilsonb/ma.html)
analyses. This could usefully be provided in a further appendix, perhaps combining the risk of bias information in one table?

Instead, we signpost the report in several places, and now include a specific reference to it in respect of further details of included studies:

We present here a summary of the key characteristics of the studies included in this analysis. Full details of all these studies, with a detailed breakdown of the risk of bias assessment, can be found online at reference [10].

**Discretionary revisions**

1. Further elaboration is required as to why the authors chose a method of study identification primarily via systematic reviews. It is not until the Discussion that this method of identification is discussed in relation to publication bias.

This is now clarified in the methods section, and a reference to the methodology paper that goes into detail has been added. “Here we briefly summarise the process which differed from many systematic reviews, because the concepts that we were searching for (i.e. community engagement and inequalities) were not always central concerns of the studies we were looking for – and hence would not appear systematically in their titles, abstracts or keywords. In order to overcome this, we identified systematic reviews of public health interventions, and utilised the structured information in their evidence tables to find relevant studies for our review. Electronic searches thus focused on the identification of systematic reviews (from which we identified primary studies), and electronic searches for primary studies were less extensive than would usually be the case. .We estimate that more than a quarter of the studies we included would have been missed using traditional search techniques.[14]”

2. Whilst further information are reported to be available in a longer report the authors should report the time frame for the literature search in the main sections of the manuscript (or perhaps in the study flowchart?).

We state the dates searched in the searches specified in the appendix. We have also added the following to the results:

“Electronic searches were carried out during July and August, 2011, with supplementary searching continuing during the autumn of 2011.”
**Reviewer 2 (Suzanne Richards)**

### Major Compulsory revisions

1. Although the authors have provided a PRISMA checklist, they have not adhered fully to PRISMA guidance [http://www.prisma-statement.org/](http://www.prisma-statement.org/) for the reporting of their review. Although reference 10 provides source documentation, the omission of some methodological detail is a major limitation that needs to be addressed. I have tried to identify areas where this occurs. Greater clarity of reporting is required.  

   **Apologies for any inadvertent omissions.** We adhered to PRISMA quite strictly when writing the original review report, so are grateful to you for pointing out the areas where we needed more detail.

2. This paper is long and difficult to follow due to the complexity of the data presented. I did wonder whether it would be better to drop some of the data (indeed reference 10 suggests it might be available elsewhere) and provide a more focused paper referring to a smaller number of key messages, and perhaps where the data are more robust/conclusive. The ideas are important, but I did struggle to get to them.  

   **We have considered this, but are conscious of possible selection bias in the results we choose to retain or omit.** Instead, we have worked with the levels of headings and subheadings, and have added some additional ‘signposting’ at the beginning of the main results section, in the hope that this will ease the reader’s navigation.

3. **Discussion.** Whenever you do a meta-regression of this sort, combining data from a large number of diverse studies, I am generally uneasy at the appropriateness of combining such data statistically. That you can come out with a ‘d’ value doesn’t mean that you should, if you are essentially mixing ‘apples with pears’ i.e. ‘clinical’ heterogeneity is present. My sense when reading the descriptive data was that this may apply here. However, I acknowledge that the academic community is split over the appropriateness of such analytic approaches and there is no clear consensus. In the discussion the authors talk about the appropriateness of statistical techniques. Perhaps they could also reflect a little on the issue of clinical heterogeneity too here.  

   **We have added the following to the discussion, before we get into the discussion on statistical heterogeneity:**

   **Issues arising from the breadth of this review topic**

   This was a challenging review to undertake due to the breadth of research and perspectives it contains. As well as crossing multiple topic domains, there are also differing perspectives regarding the nature of community engagement and what should count as a community engagement intervention. Political issues loomed large, with some papers arguing for particular solutions from utilitarian and ethical positions. We navigated this uneven landscape by structuring our analysis according to the theories of change which underpin the interventions, thus transcending differences in both health topic and politics, and focusing on the intervention mechanisms which, in some situations, bring about a change in outcomes.

   While clinical and situational heterogeneity was inevitable and unavoidable, our conceptual framework afforded us homogeneity at the theoretical level, and any claims to generalizability must also be considered at this
Minor essential

Comment 1: Page 4 ‘problems with the designs of the primary studies (e.g. the time to follow-up in the mortality studies etc). Can you rephrase this sentence as it is unclear. Do you mean ‘primary prevention’ studies. The phrase ‘mortality studies’ doesn’t really mean anything – do you mean studies which included mortality rates as an outcome? Just a few more words here would help?

The sentence has been changed to read: “The authors found that no studies evaluated the effect of community engagement on outcomes directly, and that evaluations were often carried out too soon in the intervention lifecycle to demonstrate impact.”

Comment 2: Page 5 paragraph 2. You define the term ‘community’, but you don’t add any specific references against this term. Did you follow any specific definition or did you create your own? The same could be said for ‘engagee’, ‘participant’ and ‘deliverer’, although for these three terms I suspect you had to create your own definitions to suit your study aims.

We have added references for the definition of community; though you are correct – in that we needed to define many of these terms for ourselves, and so also reference our report.

Comment 3: Page 5 paragraph 3. The control conditions include ‘no or minimal community engagement’. How did you define minimal engagement and operationalise it?

Secondary aims. Might be helpful to give a few more examples of the potential effect modifiers here – it is a quite a long time before you describe this in any detail later in the paper.

The sentence now reads: “A secondary aim is to explore moderators of the intervention effect, including study characteristics (e.g., country in which the study was conducted), intervention characteristics (e.g., how community engagement was operationalised; and characteristics of the intervention providers), participant characteristics (e.g., age), and features of the evaluations (e.g., risk of bias).”

Comment 4: Use of acronyms (e.g. DARE, NIHR HTA, DoPHER, NHS EED, TRoPI) in the search strategy. For researchers outside the UK, many of these acronyms won’t mean a lot. You may need to spell some of the more UK-centric terms out more fully.

We have now spelled out the database names: “We searched the following sources without language restriction for systematic reviews of public health interventions: Cochrane CDSR and CCTR, Campbell Library, Database of Abstracts of Reviews of Effects, NIHR Health Technology Assessments programme website, Health Technology Assessments database, and the Database of promoting health effectiveness reviews (DoPHER).”

Also – did you really not search the major electronic search databases of PubMed and

Re our overall search strategy, we have amended the text to read: “Here we briefly
| Comment 5: Page 7 paragraph 1 after bullet list. | We have added the following clarification re the scope of the Marmot review: “which assembled evidence and advised the Department of Health, England on the development of a health inequalities strategy.” |
| What are the policy objective areas of the Marmot Review? This is a major eligibility criteria which contributed to excluding over half of the potentially eligible studies. Please could you describe them more fully to assist international readers. | |
| Comment 6: Page 7 – data extraction section. Please briefly describe how the process of data extraction was conducted and quality assured (e.g. were data double extracted by two independent reviewers?). | We have added the following “To ensure consistency in interpretation and to minimise error, data extraction was undertaken by researchers working independently in pairs, and then meeting to discuss and resolve any disagreements.” |
| Comment 7: Page 8. What are ‘community outcomes; or ‘engagee’ outcomes? You need to define your outcomes more fully – perhaps in a web-supplement if the lists/definitions of each type of potential outcome are lengthy. | We have added some examples to illustrate the community and engagee outcomes |
| Community outcomes (e.g. ‘local area improved in the last 3 years’) |
| Engagee outcomes (e.g. physical activity levels or health knowledge of the engagee) |
| Comment 8: Page 9: What are ‘outcome evaluations’ – do you mean a systematic review assessing effectiveness (irrespective of study design)? Can you reword please? | We have replaced with the wording “controlled trials”. |
| Comment 9: Page 11. The description of ‘what are funnel plots’ could be shortened. | The following text has been deleted “Funnel plots (scatter plots in which the effect size estimate from individual studies are plotted against a measure of study precision) are a common method for assessing the possibility of |
Notwithstanding this, the authors do not state explicitly how the funnel plots were used to ascertain the presence of bias. Was it based on visual inspection of the funnel plot? Please state explicitly.

<table>
<thead>
<tr>
<th>Comment 10: Page 12, paragraph 2. You refer to seven systematic reviews that failed to meet minimum methodological standards to be regarded as systematic reviews – but nowhere in your eligibility strategy do you define or cross-reference to what these are. At this point in the text, I am becoming confused as to whether your focus is on systematic reviews or trials (randomised or non-randomised). I think the problem is that the eligibility criteria and selection of studies aren’t that clear in the methods section. If this could be revised, then the results section will become clearer.</th>
</tr>
</thead>
<tbody>
<tr>
<td>We hope that the clarifications that we have made to the methods, and a considerably shorter account of our results will help to clarify the search methods used. The results section now reads:</td>
</tr>
<tr>
<td>Figure 1 describes the flow of literature through the review process. As outlined earlier, studies were identified for inclusion in the review by searches of databases of systematic reviews and databases of primary research. The flow chart below reflects this two-pronged approach.</td>
</tr>
<tr>
<td>We identified 943 records of potentially relevant systematic reviews, 81 of which were duplicate records. Of the 862 unique records, 622 were excluded during assessment of titles and abstracts. Full text copies of 240 systematic reviews were obtained and assessed for eligibility. Seven of these subsequently did not meet minimum methodological standards to be regarded as systematic reviews, and a further forty-two reviews did not include any relevant primary studies. The 7,506 primary studies from the remaining 191 systematic reviews were examined for relevance, an average of 39 studies per review, within a range of three to 547. This process identified 988 eligible studies, all of which were retrieved and re-assessed against our inclusion criteria on the basis of a full-text report.</td>
</tr>
<tr>
<td>We also searched TRoPHI and NHS EED databases for reports of primary studies directly, and came across other eligible studies (through recommendations from colleagues or</td>
</tr>
</tbody>
</table>
email alerts) before and while working on the review, resulting in 1,961 titles and abstracts to screen after duplicate checking. On the basis of their titles and abstracts, the full texts of 163 of these records were retrieved.

In total, this gave us 1,151 primary study reports to screen on full text, from which a total of 361 reports of 319 studies met our inclusion criteria. After mapping the characteristics of the 319 studies we had identified, we consulted our advisory group and narrowed the focus of the meta-analysis to those studies of high priority areas for the UK, as identified in the Marmot review (8). This is summarised in Figure 2.

Comment 11: Results section – description of studies included in meta-analysis sub-heading (including studies, participants, interventions sub-headings). This is a vitally important section, but the descriptive data are sparse across a range of core areas. There are also no cross-references to tables which describe the components of the selected studies to allow the reader to fill in the gaps. This section is brief to the point it is impossible to get a feel for the selected studies based on the data give. For example. When describing the ‘comparators’ used in intervention studies – the authors state that the largest groups. of evaluations employed usual care comparators (n=38, 29%), while 13 (10%) had a comparator that differed only from the intervention by not involving community members. However, 131 studies were selected for inclusion. What where the other comparator groups?

This is an area that we found challenging to write too, as it is difficult to describe so many studies from so many settings in detail without writing a great deal of text. We have expanded some sections, and the text now reads:

“The interventions were conducted over a range of health topics and settings. The most commonly-targeted health issue was substance abuse (n = 18, 13.7%), followed by cardiovascular disease (n = 14, 10.7%), breastfeeding (n = 13, 9.9%), obesity prevention / weight reduction (n = 13, 9.9%), smoking cessation (n = 12, 9.2%, public health / health promotion (n = 8, 6.1%) and antenatal care (n = 7, 5.3%).”

The sentence on the comparators now reads: “A variety of comparators were used in the intervention evaluations. The largest group of evaluations employed usual care comparators (n = 39, 30%); followed by inactive control (n = 31, 24%), alternative / placebo intervention (n = 28, 21%), waitlist / delayed treatment (n = 16, 12%), matched data from target population (n = 10, 7%), and other / unclear (n = 7, 5%).”

Comment 12: Presentation of results data. There are a substantial number of meta-analyses presented, and the results section is difficult to follow given the sheer volume and complexity of the data. From a technical perspective, I did wonder whether it was wise to prevent pooled estimates of effect (for example, page 15 – Maintenance of intervention effects on reported health)

Our approach to dealing with heterogeneity is outlined in the methods and discussed in the discussion section (‘issues in interpreting statistical findings’). In such a broad-ranging review, some residual heterogeneity might be expected.

With regards to the specific case of ‘maintenance’, we report the value for the sake
behaviour follow-up outcomes \(d=0.09\) when significant variation was observed on heterogeneity testing. This test is telling you that the pooling of data isn’t working well.

of transparency, but do say: “significant variation \(I^2 = 94.43\%\) suggests that the pooled estimate is not particularly meaningful”, so are not suggesting that readers should take the number as a recommendation for decision-making.

Comment 13: Discussion – practical significance sub-heading. This list is a really useful summary of the key findings. However, some of the recommendations lack comparators making it difficult to interpret some of the recommendations. For example “Interventions tended to be most effective in adult populations and less effective in general populations for health behaviour outcomes.”. What is the comparator to the ‘adult population’ of the ‘general population’. Please review this list, and ensure the comparators are clearly stated to help the reader.

Thank you – we have made the following revisions:

Interventions that engage community members in the delivery of the intervention are particularly effective (compared with interventions that empower the community or involve members in the design of the intervention).

... Interventions tended to be most effective in adult populations and less effective in general populations (i.e. those that included all age groups) for health behaviour outcomes.

Interventions tended to be most effective for health behaviour outcomes for participants classified as disadvantaged due to socioeconomic position (compared with those targeted to people on the basis of their ethnicity, place of residence, or being at / high risk). Interventions targeting participants on the basis of place of residence do not appear to be effective for health behaviour outcomes.

Reviewer 3 (Anne-Marie Bagnall)

(Discretionary Revisions from attached file)

<table>
<thead>
<tr>
<th>Abstract/ Methods: CCTR is now referred to as CENTRAL.</th>
<th>This is now corrected.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results/ Description of the studies included in the meta-analysis/ The participants: It is not strictly true that “most studies” were classified as being primarily targeted at or delivered to ethnic minority group, it would be more accurate to say “The largest group of studies” or something similar.</td>
<td>This is now corrected.</td>
</tr>
<tr>
<td>Additional analyses/ sensitivity analyses/ 2nd paragraph, 4th and 5th lines: “statistically</td>
<td>Thank you – these have been amended.</td>
</tr>
</tbody>
</table>
significant from...” should probably be
“statistically significantly different from...”

**Reviewer 4 (Malinee Laopaiboon - statistical reviewer)**

| The paper is interesting and presented clearly. Suggestions are only for the results presentation. For risk of bias of the 131 included trials, figure presentation is more attractive and easier to understand than the table. In addition, forest plots of effect sizes are more attractive than the tables that are presented. They were available in Cochrane RevMan free software. | Was there an analysis in particular that required a forest plot? We have included forest plots in separate files. We have produced figures which summarise the risk of bias assessment in a similar way to RevMan – we trust these are suitable. |