Institute of Education
EPPI-Centre

# Review 2: Community engagement for health via coalitions, collaborations and partnerships

## A systematic review and meta-analysis

Ginny Brunton, Jenny Caird, Dylan Kneale, James Thomas, Michelle Richardson

EPPI-Centre
Social Science Research Unit
UCL Institute of Education
University College London

EPPI-Centre report • July 2015

REPORT

The authors of this report are:

G Brunton, J Caird, D Kneale, J Thomas, M Richardson (EPPI-Centre).

Funding

This is an independent report commissioned and funded by the National Institute for Health and Care Excellence (NICE). The views expressed are not necessarily those of NICE.

Conflicts of interest

There were no conflicts of interest in the writing of this report.

Contributions

The opinions expressed in this publication are not necessarily those of the EPPI-Centre or the funders. Responsibility for the views expressed remains solely with the authors.

This report should be cited as: Brunton G, Caird J, Kneale D, Thomas J, Richardson M (2015) Review 2: *Community engagement for health via coalitions, collaborations and partnerships: a systematic review and meta-analysis*. London: EPPI-Centre, Social Science Research Unit, UCL Institute of Education, University College London.

ISBN: 978-1-907345-78-4

## Contents

## Glossary and abbreviations

| | |
|---|---|
| ANOVA | Analysis of variance |
| ATOD | Alcohol, tobacco or drugs |
| Bidirectional communication | A method of communication which allows both engagees and engagers to express and receive the opinions of the other party. |
| Boolean minimisation | An algebraic formulisation which converts data into either ‘true’ or ‘false’ values; used in qualitative comparative analysis. |
| CERI | The original review (O’Mara-Eves et al. 2013) |
| CERUB | This current review |
| CI | Confidence interval |
| Coalition | An alliance of one or more groups of people with a common goal. |
| Collaboration | The act of working with others to achieve a common goal. |
| Collective decision making | A participatory process in which multiple individuals act collectively to make a decision. |
| Community-based participatory research (CBPR) | A partnership approach to research in which community members and researchers share expertise and decision making and contribute to all aspects of the research project. |
| Community engagement (CE) | The direct or indirect process of involving communities in decision making and/or in the planning, design, governance and delivery of services, using methods of consultation, collaboration and/or community control. |
| Conceptual framework | An analytical tool used to make conceptual distinctions and organise ideas. |
| Configuration | The term used to describe the combination of characteristics within a study during qualitative comparative analysis. |
| Consultation | A process through which information or advice is sought, but not necessarily acted upon. Decision making powers do not lie with those consulted. |

| | |
|---|---|
| Continuous outcome | Outcomes for participants are measured on a numerical scale and the results ordinarily summarised using the mean. |
| DH | Department of Health (UK) |
| Dichotomous (binary) outcome | an event which did or did not occur, e.g. death, pregnancy, disease state. |
| Framework synthesis | A structured approach to organising and analysing data in matrices or charts. |
| Fuzzy set | A dataset of studies which have had their outcomes standardised by their effect size for use in qualitative comparative analysis. |
| HEPA | Healthy eating/physical activity |
| Homogeneity | Of a uniform type. |
| Heterogeneity | Of a non-uniform type. |
| In-depth synthesis | A synthesis involving detailed scrutiny, as opposed to descriptive characterisation, of the available research. |
| IISP | Injury and infection screening |
| Logical remainders | An empty truth table row, indicating limited diversity of phenomena. |
| Mediator | A variable intervening in the causal pathway between two variables e.g., if A is significantly associated with C, and if A influences B and B influences C, then B is a mediating variable. |
| Meta-analysis | A statistical approach used to combine the results from multiple individual studies with improved power and greater precision in estimating effect size. |
| Meta-regression | A form of meta-analysis used to examine the impact of moderator variables on the study effect size via multiple regression analysis. |
| Moderator | A variable affecting the direction and/or strength of the association between a predictor and outcome variable. |
| Modifiable processes | A community engagement process capable of being changed or altered. |
| Necessary (of a condition) | *Must* be satisfied in order to obtain an outcome. If one condition is a necessary and sufficient condition of another, the former statement is true if and only if the latter is true. |

| | |
|---|---|
| NICE | National Institute for Health and Care Excellence |
| Odds ratio (OR) | A measure of association between an exposure and outcome - the odds of an outcome occurring given a particular exposure, compared to the odds of the outcome occurring without exposure. |
| OECD | Organisation for Economic Co-operation and Development |
| PHAC | Public Health Advisory Committee (UK) |
| Population churn | A measure of the turnover of individuals moving into or out of a group over a period of time. |
| Process evaluation | A study examining the development or implementation of an intervention or programme. |
| PROGRESS-Plus | This is used to denote markers of disadvantage: place of residence, race/ethnicity, occupation, gender, religion, education, socio-economic status, social capital and three further variables - age, disability and sexual orientation. |
| Qualitative comparative analysis (QCA) | A means of analysing the contribution of different conditions, or combinations of conditions, to an outcome. |
| RQ | Research question |
| Self-efficacy | Belief in ability to achieve a task, goal or outcome. |
| Sufficient (of a condition) | Its existence leads to the occurrence of a given outcome. If one condition is a necessary and sufficient condition of another, the former statement is true if and only if the latter is true. |
| Standardised mean difference | A measure of the effect size of an intervention - the difference in the means of between-study groups relative to the variability observed in the study under consideration. |
| Synthesis | The combination of separate elements to form a connected whole. In systematic reviews, it is a combination of the findings of individual studies in order to answer the review question. |

# Executive summary

## Background

Previous research suggests that interventions utilising community engagement show large beneficial effects, as well as considerable variation across populations, intervention types and outcomes (O'Mara-Eves et al. 2013). This variation makes it difficult to understand how community engagement works. Our previous research synthesis of community engagement did not suggest an ideal model; across different models of engagement, there is insufficient evidence that one particular model of community engagement (i.e. one combination of engagement across design, delivery and evaluation) is likely to be more effective for health outcomes than any other. Further, no clear model was identified that worked best across all contexts, populations and health issues (O'Mara-Eves et al. 2013). A synthesis of process evaluations from that review also suggested that implementation issues and consultative processes might influence the success of an intervention (O'Mara-Eves et al. 2013). This was corroborated by subsequent analyses of specific community engagement processes in a smaller subset of breastfeeding interventions, where we suggested that some processes, such as provider training, intervention feasibility and intensity may be more aligned with effective outcomes (Thomas et al. 2014). These findings are corroborated by the findings of Review 1 of this project, which suggested that both the extent and particular processes of community engagement may be linked to effects on people's health (Brunton et al. 2015).

## Aims and objectives of the review

The review addressed the following overarching research questions (RQ):

> *RQ1. How effective are community engagement approaches at improving health and wellbeing and reducing health inequalities?*
>
> *RQ2. Across disadvantaged groups, how effective are community engagement approaches at encouraging people to participate in activities to improve their health and wellbeing and realise their capabilities?*
>
> *RQ3. What processes and methods facilitate the realisation of community and individual capabilities and assets amongst disadvantaged groups?*
>
> *RQ4. Are there unintended consequences from adopting community engagement approaches?*
>
> *RQ5. What processes identified in the literature are more aligned with effective interventions, and which (if any) are more aligned with non-effective interventions?*

The aim of the in-depth synthesis undertaken in Review 2 is to examine and evaluate the processes and extent of community engagement across all stages of an intervention. This is done for the purpose of informing NICE PHAC members about the likely components and

processes of successful community engagement. Research questions for three separate syntheses were examined to address these aims:

1. Framework synthesis:
    a. What modifiable processes of community engagement are evaluated in the literature?
    b. Which modifiable processes of community engagement are associated with higher/lower extents of engagement?
    c. Are the processes of community engagement used differently with different populations?
2. Meta-analysis/modelling:
    a. Are potentially modifiable processes of community engagement associated with health outcome effects?
    b. What is the relationship between the extent of community engagement (high, moderate or low) and health outcome effects?
    c. Do direct comparisons of community engagement (i.e. studies that test a community engagement intervention versus the same intervention without community engagement) differ in health outcome effects from indirect comparisons (e.g. those that test community engagement versus usual care)?
    d. Do health outcome effects differ for:
        i. different age groups;
        ii. studies targeting men only versus those targeting women only;
        iii. studies specifically developed for low-income groups versus those that are not;
        iv. 'distal' (e.g. self-efficacy), 'intermediate' (e.g. health behaviour), or 'proximal' (clinical/physiological measure) outcomes?
3. Qualitative Comparative Analysis (QCA):
    a. Which *necessary* and *sufficient* intervention components are associated with effective interventions?

**Methods**

To be included in Review 2 of this project, studies had to:

- explicitly describe the use of coalitions, collaborations, or partnerships;
- provide data on processes of community engagement;
- provide data on health outcomes, i.e. self-efficacy, behavioural outcomes or clinical or physiological outcomes.

All outcome evaluations were assessed for risk of bias in Review 1 of this project (Brunton et al. 2015). Linked process evaluations were subjected to quality assessment using a tool developed from the original community engagement review (O'Mara-Eves et al. 2013).

*Framework synthesis*

Modifiable processes of community engagement described under the 'Actions' column of the conceptual framework developed in the previous review (e.g. administrative support or training support; see Appendix 1) were used as the 'framework' for the present

analysis. We 'populated' the framework with studies describing various processes, and then thematically compared and contrasted aspects of each process looking at differences in age groups, gender or socio-economic disadvantage using an adaptation of previously developed methods (Oliver et al. 2008; Ritchie and Spence 1994; Thomas et al. 2012).

*Meta-analysis*

Effect sizes were calculated to summarise the impact of community engagement interventions. We transformed odds ratios to standardised mean difference effect sizes using the methods described in Chinn (2000). Outcomes were classified into domains according to a conceptualisation of a causal pathway. The domains, in order of the theory of change, were self-efficacy, health behaviour change and finally clinical/ physiological consequences. Data were analysed using descriptive statistics, meta-analysis, homogeneity tests and meta-regression.

*Qualitative comparative analysis*

Qualitative comparative analysis (QCA) was employed to generate theory about *necessary* and *sufficient* components that are associated with effective interventions. Characteristics of studies examined in the meta-analysis were further analysed in the QCA, including health topic, extent of engagement, the population under study and effect sizes. Two researchers met to consider the findings for processes of community engagement, and discuss the interactions of these processes with the theories of community engagement in order to build a data table. QCA was undertaken in six steps (Rihoux and Ragin 2008). These were:

1. building the data table
2. constructing a 'truth' table
3. resolving contradictory configurations
4. Boolean minimisation
5. consideration of logical remainders
6. interpretation.

**Findings**

A total of 26 studies from Review 1 were included, as well as 38 studies from the original review of community engagement (n=64 studies) (O'Mara-Eves et al. 2013). Our findings represent the first attempt of which we are aware, to examine the processes of community engagement, rather than the processes of intervention implementation, i.e. those processes used to engage community members in the design, delivery or evaluation of interventions, rather than barriers or facilitators to intervention implementation. We found that evaluation of the processes of community engagement itself has not been undertaken routinely across studies. Understanding which processes have at least been reported, describing these processes, and looking at patterns across populations and in studies indicating different levels of engagement, helped us to understand which processes of community engagement, such as collective decision making, occur in studies which rated a higher extent of community engagement. It has also shed light on whether these processes, or other characteristics of the studies, such as populations under study or health topics, are associated with high community engagement and better outcomes.

*Framework synthesis*

Starting from the original community engagement review's conceptual framework that illustrated the modifiable processes of community engagement (i.e. how community members were engaged), we found that collective decision making, bidirectional communication and training support were most often reported.

Studies that were rated as having a high or moderate extent of engagement more often reported specific processes of community engagement, such as collective decision making, than did studies rated as having a low extent of engagement. This suggests that with more rigorous evaluation, such processes could be used in more intensive community engagement efforts.

Interventions targeting low-income groups more often reported specific processes such as collective decision making, bidirectional communication and training. This may indicate useful processes to be employed with these specific populations, although more rigorous evaluation is required.

Other modifiable processes were identified *de novo*, including skills around conflict resolution, negotiation and reflection, and planning collaborative meetings to suit community members' needs in terms of timing, location, and provision of transport and childcare. These could be incorporated into future evaluations of community engagement, to show specific ways to undertake it for health intervention design, delivery and evaluation. However, to establish their validity these and the other processes of community engagement require more robust evaluation.

*Meta-analyses*

A high extent of involvement was associated with higher effect sizes for interventions with behavioural outcomes measured longitudinally. Compared to interventions with a low extent of involvement, studies with a high extent had an effect size (*d*) that was larger (over half a standard deviation (0.673) larger). Extent of involvement was an important factor in measuring between-study variance, accounting for 52% of such variance. Analyses suggested that deploying multiple community engagement processes may be associated with statistically significantly larger effect sizes, but no one individual process could be attributed with increased effectiveness in the meta-analyses.

A meta-analysis of ten studies produced a pooled effect size of $d = 0.192$ (CI: 0.092-0.292), suggesting that interventions based on coalitions for community engagement made a small impact on clinical health outcomes on measures collected longitudinally. In studies measuring behavioural outcomes using cross-sectional methods, community engagement interventions appeared to produce a higher effect size in studies with children and young people, although this could be due to 'population churn' or to potential publication bias.

*Qualitative comparative analysis*

Qualitative comparative analysis indicated that four configurations aligned with effective interventions: the inclusion of lay delivery of interventions; the inclusion of lay delivery targeted to the general population; lay delivery of interventions focused on sexual health, organ donation or cancer prevention; and lay delivery interventions focused on infection or injury prevention targeted to the whole population. This suggests that lay delivery may

need to be present for coalitions to be successful, but on its own may not lead to effective outcomes.

One configuration was found more often in studies with lower effect sizes: coalitions employing a low extent of community engagement in design, delivery and evaluation. This suggests that interventions employing a higher extent of engagement tend to experience larger outcome effects.

**Discussion**

This review sought to address several research questions:

1. *How effective are community engagement approaches at improving health and wellbeing and reducing health inequalities?*

The findings from these analyses suggest that within projects that utilise coalitions, collaborations or partnerships with community members, higher behavioural outcome effect sizes are achieved through community members leading or collaborating on the design, delivery and evaluation of an intervention.

2. *Across disadvantaged groups, how effective are community engagement approaches at encouraging people to participate in activities to improve their health and wellbeing and realise their capabilities?*

While framework synthesis from Review 1 suggested that a high extent of community engagement was seen in low-income groups in particular, subsequent meta-analyses provided no firm evidence of differences between these or any other disadvantaged groups, due to methodological limitations of the studies.

3. *What processes and methods facilitate the realisation of community and individual capabilities and assets amongst disadvantaged groups?*

While evidence was located which evaluated the processes of intervention implementation, no studies provided evaluations of the processes of community engagement. We identified descriptions of several processes of community engagement, including bidirectional communication, collective decision making, training support for intervention provision for either community engagees or professionals, allowing adequate time for relationship development, negotiation/reflection/conflict resolution skills and arranging meetings to suit community members' needs. Future evaluation of these processes could provide a starting point for recommending good practices of community engagement.

4. *Are there unintended consequences from adopting community engagement approaches?*

No evidence was found which suggested unintended consequences from adopting community engagement approaches, although the findings from our QCA suggested that studies using less community engagement tended to show smaller effect sizes.

5. *What processes identified in the literature are aligned with effective interventions, and which (if any) are aligned with non-effective interventions?*

Findings from our QCA provided tentative evidence that a higher extent of community engagement was seen in studies with higher effect sizes; conversely, less community

engagement across design, delivery and evaluation tended to be found in studies with lower effect sizes.

We conclude that continued involvement of community members throughout the entire lifespan of a collaboration (i.e. through design, delivery and evaluation) leads to higher effects, and that community engagement addressing some health issues may show higher effects than in others. Further, specific modifiable processes of community engagement have been described in the literature. These suggest useful ways for researchers and service providers to work with community members in a collaborative way.

The analyses undertaken in Review 2 both confirm and further refine those found in Review 1. For example, the Review 2 findings confirm those of Review 1 that suggest that a high extent of community engagement across design, delivery and evaluation is associated with greater beneficial effects of health interventions, in comparison to either moderate or low extent of community engagement.

However, the more detailed analysis undertaken in the Review 2 meta-analysis suggests that health behaviour outcomes might be larger in studies focused specifically on injury or infection prevention or screening, rather than in the larger list of health topics suggested in Review 1. In addition, conclusions in Review 1 drawn from the examination of population subsets (i.e. women, men, children/young people, low-income populations) were not borne out in the more detailed analyses undertaken in Review 2. This is most likely due to the addition of studies from the original review of community engagement into the meta-analysis, which had two effects: to add power to the size of the sample for meta-analysis, thus making the findings more robust; and to reduce the amount of data to which the findings from the framework synthesis and QCA can be attributed, due to a lack of coding on community engagement processes from the original review's included studies.

**Conclusions**

Taken together, the findings across all three syntheses in this review suggest that community-led or community collaboration projects which design, deliver and evaluate health interventions are associated with larger behavioural outcomes. Where coalitions, collaborations and partnerships with community members include the use of bidirectional communication, collective decision making and community member or professional training support for intervention provision, a higher extent of community engagement across the project's design, delivery and evaluation was also found. Effective configurations of engagement within collaborations and coalitions generally include peer or lay delivery, and projects with a low extent of engagement were aligned with lower effect sizes.

**Evidence statements**

The evidence statements derived from these syntheses can be found in Chapter 7.

# 1. Background

## 1.1 Review context

Involving communities in decision making and in the planning, design, governance and delivery of services has become central to guidance and national strategy for promoting public health (Department of Health (DH) 2002, 2005, 2006a,b,c, 2010). The National Institute for Health and Care Excellence (NICE) plays a crucial role in providing guidance on best practice for community engagement. Since the publication of NICE Community Engagement guidance (National Institute for Health and Care Excellence 2008), there has been considerable activity with a view to understanding the nature of community engagement, its benefits, and challenges in its evaluation (for example, Phillips et al. 2014, Jamal et al. 2013, Sheridan and Tobi 2010, Sheridan et al. 2011).

Community engagement can take many forms, including volunteering, peer delivery, community coalitions, advocacy and social networks; and community members can be involved to varying degrees within a public health strategy, including leading, collaborating, consulting or being informed about the design, delivery or evaluation of an intervention (O'Mara-Eves et al. 2013).

Previous research suggests that interventions utilising community engagement show large beneficial effects, as well as considerable variation across populations, intervention types, and outcomes (O'Mara-Eves et al. 2013). This variation makes it difficult to understand how community engagement works.

A conceptual framework analysis in the same review identified several factors which influence community engagement, including: understanding motivations for seeking and participating in community engagement; conditions such as appropriateness and acceptability; actions, such as relationship building, communication techniques and other methods to engage communities; and the impacts for those who engage as well as the receiving community (O'Mara-Eves et al. 2013). This work identified some key issues in community engagement that merit further exploration. These include a consideration of the pathways through which an effective outcome can be achieved and a need for more research on the economic and implementation aspects of community engagement.

Some theoretical models of community engagement argue that involving the public in order to empower or enable them is crucial (Popay et al. 2007); others suggest that having community members' input into the design and/or delivery of an intervention improves its acceptability, thus making positive health outcomes more likely (Arblaster et al. 1996). Our previous research synthesis of community engagement did not suggest an ideal model: across different models of engagement, there is insufficient evidence that one particular model (i.e. one combination of engagement across design, delivery and evaluation) is likely to be more effective for health outcomes than any other.

Evidence suggested that peer-delivered interventions alone appeared to provide higher effect sizes in health outcomes than interventions with community members leading, collaborating or consulting on design (O'Mara-Eves et al. 2013). However, this did not

examine continued involvement of community members throughout the design, delivery and evaluation of an intervention. Further, no clear model of community engagement was identified that worked best across all contexts, populations and health issues (O'Mara-Eves et al. 2013).

A synthesis of process evaluations from that review also suggested that implementation issues and consultative processes might influence the success of an intervention (O'Mara-Eves et al. 2013). This was corroborated by subsequent analyses of specific community engagement processes in a smaller subset of breastfeeding interventions (Thomas et al. 2014). We suggested that some processes such as provider training, intervention feasibility and intensity may be more aligned with effective outcomes (Thomas et al. 2014).

### 1.2 Moving from Review 1 (map) to Review 2 (in-depth synthesis)

These findings appear to support those from the current synthesis of studies in Review 1 of this project, which suggested that both the modifiable processes and the extent of community engagement may be linked to effects on people's health (Brunton et al. 2015).

Consultations with our Advisory Group (listed in Review 1 report (Brunton et al. 2015)) highlighted the need to focus on the specific processes of community engagement (rather than motivations, mediators or conditions of community engagement), in order to inform PHAC members which components are contained within a successful community engagement initiative. A need was also identified to understand whether the process of community engagement varies throughout the life of a project, i.e. across design, delivery and evaluation of an intervention. Within this, a need was highlighted to examine differences in the processes of community engagement across age groups, health topics and type of disadvantage where possible.

### 1.3 Aims and objectives of the review

The aim of the in-depth synthesis undertaken in Review 2 was therefore to examine and evaluate the processes and extent of community engagement across all stages of a research project. This is done for the purpose of informing NICE PHAC members about the likely components and processes of successful community engagement.

### 1.4 Research questions

The review addressed the following research questions (RQ):

> *RQ1. How effective are community engagement approaches at improving health and wellbeing and reducing health inequalities?*
>
> *RQ2. Across disadvantaged groups, how effective are community engagement approaches at encouraging people to participate in activities to improve their health and wellbeing and realise their capabilities?*
>
> *RQ3. What processes and methods facilitate the realisation of community and individual capabilities and assets amongst disadvantaged groups?*
>
> *RQ4. Are there unintended consequences from adopting community engagement approaches?*

*RQ5. What processes identified in the literature are more aligned with effective interventions, and which (if any) are more aligned with non-effective interventions?*

To address these overarching research questions, sub-questions for three separate syntheses were examined:

1. Framework synthesis:
    a. What modifiable processes of community engagement are evaluated in the literature?
    b. Which modifiable processes of community engagement are associated with higher/lower extents of engagement?
    c. Are the processes of community engagement used differently with different populations?
2. Meta-analysis/modelling:
    a. Are potentially modifiable processes of community engagement associated with health outcome effects?
    b. What is the relationship between the extent of community engagement (high, moderate or low) and health outcome effects?
    c. Do direct comparisons of community engagement (i.e. studies that test a community engagement intervention versus the same intervention without community engagement) differ in health outcome effects from indirect comparisons (e.g. those that test community engagement versus usual care)?
    d. Do health outcome effects differ for:
        i. different age groups;
        ii. studies targeting men only versus those targeting women only;
        iii. studies specifically developed for low-income groups versus those that are not;
        iv. 'distal' (e.g. self-efficacy), 'intermediate' (e.g. health behaviour), or 'proximal' (clinical/physiological measure) outcomes?
3. Qualitative comparative analysis (QCA):
    a. Which *necessary* and *sufficient* intervention components are associated with effective interventions?

### 1.5 Operational definitions

A community is defined as a group of people either self-identified or identified by others, who share one or more common characteristics that can include geographical neighbourhood, health status, ethnicity, or shared interests, values, experience or traditions (Brenner et al. 2011). We have defined community engagement as a 'direct or indirect process of involving communities in decision making and/or in the planning, design, governance and delivery of services, using methods of consultation, collaboration, and/or community control' (O'Mara-Eves et al. 2013:p.6).

### 1.6 Identification of possible equality and other equity issues

Because of the large body of literature identified, this review has focused on community engagement involving disadvantaged communities. While the review provides a lot of

information on those experiencing health inequalities, it does not include information on non-disadvantaged communities.

### 1.7 Review team

The review team comprised researchers from the Evidence for Policy and Practice Information and Coordinating (EPPI-) Centre at the UCL Institute of Education. The team has a history of undertaking innovative systematic reviews that incorporate the public's views during review design, conduct or evaluation (i.e. advisory groups and peer review of reports). The EPPI-Centre team undertook a large-scale systematic review and meta-analysis examining the conceptual framework, processes, effectiveness and cost-effectiveness of community engagement strategies (O'Mara-Eves et al. 2013).

The team and their roles for the current review were as follows: Ginny Brunton, a Research Officer, acted as principal investigator; lead on framework synthesis coordinating analysis and writing of report; and project manager for the review. James Thomas, a Professor of Social Policy, was a co-investigator, leading on the qualitative comparative analysis. Jenny Caird, a Research Officer, was a co-investigator performing literature searches, screening and coding, and acted as lead analyst on the meta-analysis modelling. Dylan Kneale is a Research Officer; his role included coding, meta-analysis and modelling of studies. Michelle Richardson is a Research Officer; her role was data extraction, advisement on meta-analytic techniques and derivation of evidence statements. Claire Stansfield is an Information Specialist who has contributed to the review through the design, development and testing of the search strategy. Each team member has declared no conflict of interest.

# 2. Methodology

## 2.1 Evidence identification

An update of the recent systematic review (O'Mara-Eves et al. 2013) was undertaken, using innovative methods of locating and screening the literature. Full details of the searching, screening and quality assessment of included studies is described in Review 1 (Brunton et al. 2015).

A call for evidence to the project stakeholders was made by NICE during June and July 2014. Additional relevant evidence was added to the review process through this route. The authors of the included studies were contacted and asked for any additional data as required to undertake syntheses.

## 2.2 Evidence selection

To be included in the syntheses for Review 2, studies were considered from our original review of community engagement (O'Mara-Eves et al. 2013) and from the set of studies identified in Review 1. All references included in any synthesis are indicated with an asterisk in the References. To be included in the synthesis, studies had to:

- explicitly describe the use of coalitions, collaborations, or partnerships;
- provide data on processes of community engagement;
- provide data on health outcomes, i.e. self-efficacy, behavioural change or clinical or physiological outcomes.

## 2.3 Quality appraisal

All outcome evaluations were assessed for risk of bias in Review 1 (Brunton et al. 2015). Linked process evaluations were subjected to quality assessment using a tool developed from the original community engagement review (O'Mara-Eves et al. 2013).

## 2.4 Framework synthesis

The processes of community engagement described under the 'Actions' column of the conceptual framework developed in the previous review of community engagement were most likely to be modifiable (see Appendix 1). These were used as the 'framework' for the present analysis. Studies were coded with respect to whether there was evidence of the following modifiable processes taken from the original conceptual framework:

- bidirectional communication
- collective decision making
- training support (i.e. for community members to learn how to take part in the coalition/collaboration/partnership)
- administrative support (i.e. paid staff to organise meetings, take and circulate minutes etc.)
- sustainable funding processes
- frequency of coalition meetings
- duration of coalition meetings

- timing of coalition meetings
- adequacy of time to allow collaborative relationships to develop
- other modifiable processes not described above (to capture any newly emerging processes).

We extracted Yes/No data (or amounts stated by the authors) from all process evaluations for potentially modifiable processes of community engagement.

Consultation with NICE Stream 2 colleagues about emerging processes of community engagement in the literature identified a need to include provider and community engagee training as additional processes beyond those in the conceptual framework. Other processes were added as they emerged from the data, and all studies were reassessed for newly identified processes.

The resulting data extracted from the process evaluations underwent a framework synthesis, where we 'populated' the framework above with studies that describe each process, and then thematically compared and contrasted aspects of each process looking at differences in age groups, gender or socio-economic disadvantage, and looked for relationships with effect sizes, using an adaptation of previously developed methods (Oliver et al. 2008; Ritchie and Spencer 1994; Thomas et al. 2012).

Two members of the review team independently extracted data on processes of community engagement and made risk of bias assessments, then met to discuss and agree ratings. Emerging processes of community engagement, and themes derived from analysis across populations were discussed amongst team members. Data were analysed using EPPI-Reviewer 4 software and Excel.

### 2.5 Statistical moderator analysis

#### *2.5.1 Summary measures*

For the meta-analysis, effect sizes were calculated to summarise the impact of the interventions. Because many of the outcomes used different scales and different combinations of continuous and dichotomous data, we used the standardised mean difference (White and Thomas 2005) to compare and combine results of continuous measures, and odds ratios (ORs) for binary measures. We transformed the ORs to standardised mean difference effect sizes using the methods described in Chinn (2000). We adjusted the standard errors of cluster randomised trials that had a disproportionate weighting. Estimates of the intra-cluster correlation were imputed using estimates derived from other similar studies included in the review when not available in the report or obtainable via author contact.

The outcomes were classified into domains according to a conceptualisation of a causal pathway. The domains, in order of the theory of change, were self-efficacy, health behaviour change and finally clinical/ physiological consequences.

The 'extent' of community engagement was determined as follows:

For three aspects of the intervention (design, delivery and evaluation), the level of community engagement was rated as:

- leading or collaborating = 1
- consulted, informed or not involved = 0.

To determine the extent of community engagement, the level of engagement across all three aspects of an intervention was summed and the extent determined as follows:

- high = 3
- moderate = 2
- low = 1 or 0.

### *2.5.2 Meta-analysis*

The meta-analysis (quantitative synthesis) used various statistical methods to address our research questions, by testing whether any observed differences in the results of included studies might be associated with the type of community engagement employed. This is reported in Chapter 4. The methods used were descriptive statistics, meta-analysis (homogeneity tests), analysis of variance (ANOVA) and meta-regression (Thompson and Sharp 1999). For random-effects model analyses we followed the methods described in Lipsey and Wilson (2001).

Meta-regression models were fitted (where data permitted) using the *metareg* command in Stata v.12.1 (Statacorp, College Station, TX). A minimum of 10 studies was considered sufficient for undertaking meta-regression analyses, and for dichotomised constructs at least three studies were required in each category. For each potential moderator, we reported the pooled effect size and corresponding 95% confidence intervals (CIs), the proportion of between-cluster variability (Adjusted $R^2$) accounted for by the moderator variable and $I^2$ - the proportion of residual between-study variation due to heterogeneity (Borenstein 2009; Borenstein et al. 2011).

Specific aspects of the analysis, such as identification and treatment of outliers and skewed data, sensitivity analyses and assessment of publication bias, are reported in Chapter 4 alongside the relevant results to facilitate understanding of the findings.

## 2.6 Qualitative comparative analysis

Qualitative comparative analysis (QCA) is a method that aims to generate theory about the *necessary* and *sufficient* components that are associated with effective interventions. QCA allows for the identification of multiple pathways to an outcome and is especially suited to working with small numbers of studies. This makes it a suitable complement to meta-regression techniques. It can assist review users in choosing the components necessary to ensure success in a given situation (Thomas et al. 2014).

Two reviewers coded studies independently, using a previously developed tool. The reviewers met to discuss and agree coding, with disagreements resolved by a third reviewer. All data were entered into EPPI-Reviewer 4 software. QCA truth tables and configurations were calculated using STATA software. Two researchers met to consider the findings for processes of community engagement, and to discuss the interactions of these processes with the theories of community engagement in order to build a data table. QCA was undertaken in six steps (Rihoux and Ragin 2008). These were:

1. building the data table

2. constructing a 'truth' table

3. resolving contradictory configurations

4. Boolean minimisation

5. considering of logical remainders

6. interpretation.

*1. Building the data table*

Based on current understanding of theories of community engagement and the findings from the framework synthesis and meta-analysis/meta-regression, all the described processes of community engagement were considered. A series of processes of community engagement that provided the most robust evidence of the extent of community engagement were selected as the conditions to be tested. These were:

1. the presence of collective decision making;
2. adequate time for relationship development; and
3. cumulative processes of community engagement.

Studies were given a value of 1 if they met each condition or 0 if they did not. Studies that did not report on the process of engagement were given a value of 0.

The extent of community engagement in these studies indicates the impact of the process of community engagement. The original metric used in the framework synthesis was the rating of the extent of community engagement (a summing of whether community members led (score=4), collaborated (score=3), consulted (score=2), were informed (score=1) or were not involved (score=0) in each aspect of design, delivery and evaluation). Studies were rated out of a possible 12 points. These ratings were calibrated for use in the QCA analyses by converting them into a fuzzy set that allows for degrees of membership. Effect size estimates for constructing the fuzzy sets were calibrated as detailed in Table 2.1:

**Table 2.1:** Extent of community engagement: set membership determination

| **Extent of CE** | **Membership in 'Effective' set** | **Fuzzy set value** |
|---|---|---|
| 10-12 | In (full membership) | 1.00 |
| 7-9 | More in than out | 0.66 |
| 4-6 | More out than in | 0.33 |
| 0-3 | Out (non-membership) | 0 |

*2. Constructing a 'truth table'*

Studies were assigned to a configuration set (a 'truth table') depending on their combination of conditions. All possible configurations of the three conditions are illustrated in Table 2.2.

**Table 2.2:** Possible configurations and set labels of the three conditions

| Conditions | | | Configuration set label |
|---|---|---|---|
| **Collective Decision making (CDM)** | **Adequate Time for Relationship Development (ADEQTIME)** | **Cumulative Processes (CUMPROC)** | |
| 1 | 1 | 1 | CDM*ADEQTIME*CUMPROC |
| 1 | 1 | 1 | CDM*ADEQTIME*CUMPROC |
| 1 | 1 | 0 | CDM*ADEQTIME*~CUMPROC |
| 1 | 0 | 1 | CDM*~ADEQTIME*EXTCE |
| 1 | 0 | 1 | CDM*~ADEQTIME*CUMPROC |
| 1 | 0 | 0 | CDM*~ADEQTIME*~CUMPROC |
| 0 | 1 | 1 | ~CDM*ADEQTIME*CUMPROC |
| 0 | 1 | 1 | ~CDM*ADEQTIME*CUMPROC |
| 0 | 1 | 0 | ~CDM*ADEQTIME*~CUMPROC |
| 0 | 1 | 0 | ~CDM*ADEQTIME*~CUMPROC |
| 0 | 0 | 1 | ~CDM*~ADEQTIME*CUMPROC |
| 0 | 0 | 0 | ~CDM*~ADEQTIME*~CUMPROC |

* and ~ not

Truth tables were constructed and assessed for all ranges of engagement (e.g. from low to high extent of engagement). The resultant tables were checked for the spread of studies across the different configurations available and whether high, moderate and low extents of engagement were well covered or not.

*3. Resolving contradictory configurations*

Two reviewers assessed the dataset for any contradictory configurations (i.e. sets of studies in which identical configurations of conditions led to the different outcomes) and resolved (Rihoux and Ragin 200).

*4. Boolean minimisation*

The set was analysed using fsQCA software (Ragin et al. 2006). The primary metric for these analyses was a measure of *raw consistency*. Consistency is the proportion of all intervention studies with conditions of interest and the outcome of interest (Ragin 2006). We considered studies with a consistency value of >0.75 to be a valid combination, for two reasons: this was the suggested cut-off value (Ragin et al. 2006); and our set of studies was sufficiently heterogeneous in terms of outcomes to allow a consistency value toward the lower end of the scale. We also examined *coverage* as a metric. Coverage is the proportion of studies in the set of interest that have the condition of interest (Ragin 2006). This metric provides useful information about how often the conditions have occurred across all included studies.

*5. Consideration of logical remainders*

Two reviewers considered any remainders (i.e. configurations with no cases) in order to consider logical explanations in the light of the conceptual framework guiding the analysis.

*6. Interpretation*

Combinations of conditions or solutions were interpreted in the light of the studies on which they were based, the aims of this study and the original systematic review's research questions, including the conceptual framework which guided the review.

### 2.7 Formulation of evidence statements

Evidence statements for the findings from both Review 1 and the moderator and qualitative comparative analyses were derived, following the structure and process indicated in Section 5.5 of the NICE methods guidance (National Institute for Health and Care Excellence 2012). 'Evidence' referred to the sources of evidence (study type and references) and their quality in brief descriptive terms. In addition, each statement included summary information about the:

- **content** of the intervention, where applicable (for example, what, how, where?)
- **population(s)** and **setting(s)** (and country), where applicable
- **strength** of the evidence (reflecting the appropriateness of the study design to answer the question and the quality, quantity and consistency of the evidence)
- **outcome(s)**, the **direction** of effect (or correlation) and the **size** of effect (or correlation) (where applicable)
- **applicability** to the question, target population and setting.

The overall strength (quality, quantity and consistency) of the evidence was summarised (being clear about the sources and inclusion criteria) as:

- **No evidence**. For example: 'No evidence was found from English-language trials published since 1990…'.
- **Weak evidence.** For example, 'There was weak evidence from 1 (−) before-and-after study'.
- **Moderate evidence.** For example, 'There was moderate evidence from 2 (+) case-control studies'.
- **Strong evidence**. For example, 'There was strong evidence from 2 (++) and 1 (+) randomised controlled trials'.
- **Inconsistent evidence.** Where needed, further commentary was provided on the variability of findings in different studies. For example, when the results of (++) or (+) quality studies did not agree. In such cases, the review team qualified an evidence statement with an explanatory sentence or section giving more detail.

'Vote counting' (merely reporting on the number of studies yielding significant effects) is **not** an acceptable summary of the evidence.

Where appropriate, the **direction of effect** (impact) **or correlation** was summarised using one of the following terms:

- positive
- negative
- mixed
- none.

Where appropriate, the **size of effect** (impact) **or correlation** and, when possible, the degree of uncertainty involved, was reported using the scale applied in the relevant study. For example, an odds ratio (OR) or relative risk (RR) with confidence interval (CI), or a standardised effect size and its standard error, might be quoted. Where an estimate could not be explained, every effort was made to relate it to interpretable criteria or conventional public health measures. Where it was not possible to provide figures for each study, or where there were too many studies to make this feasible, the **size of effect** or **correlation** was summarised using the following standardised terms:

- small
- medium
- large.

In order to assist the PHAC in judging the extent to which the evidence reported in the reviews is applicable to the areas for which it is developing recommendations, we assessed each evidence statement to judge how similar the population(s), setting(s), intervention(s) and outcome(s) of the underpinning studies were to those outlined in the review question(s). The studies were assessed as a whole. Following this assessment, we categorised each evidence statement as:

- directly applicable
- partially applicable or
- not applicable

A statement detailing the category that the evidence statements fell into was included.

### 2.8 Quality assurance

Two reviewers independently conducted data extraction, and the final version was agreed upon to maintain accuracy. Where necessary, a third team member arbitrated in disagreements. Evidence tables were completed using templates based on those provided in NICE methods guidance (National Institute for Health and Care Excellence 2012). Records of the research identified by searches were uploaded to the specialist systematic review software, EPPI-Reviewer 4, for duplicate stripping and screening (Thomas et al. 2010). This software was used to record the bibliographic details of each study considered by the review, where studies were found and how, and the reasons for their inclusion or exclusion. EPPI-Reviewer 4 was also used to conduct and record the data extraction and quality appraisal stages for the included studies, using the required data fields and appropriate quality checklists detailed in the methods manual (National Institute for Health and Care Excellence 2012).

## 3. Findings: framework synthesis

Of the 28 studies included in Review 1, 26 were included in the Review 2 syntheses. Two studies were excluded: Bergstrom et al. (2013) did not describe community engagement with people from the identified population of need; and Wermert et al. (2012) because the population under study was not disadvantaged.

In the remaining 26 studies, data about the processes of community engagement were extracted using the conceptual framework previously developed (O'Mara-Eves et al. 2013). Due to the project timelines, this analysis was undertaken on the 26 new studies identified in this review only.

Descriptive frequencies were collected for each process of community engagement across the set of 26 studies. The number of studies reporting some evidence of each process of community engagement is provided in Table 3.1.

**Table 3.1:** Processes of community engagement (N=26 studies)

| Process | No. of studies |
|---|---|
| Bidirectional communication | 18 |
| Collective decision making | 18 |
| Training support | 18 |
| Adequacy of time to develop relationship | 5 |
| Administrative support | 1 |
| Other: Conflict resolution | 2 |
| Other: Arrange meetings to suit community members' needs | 2 |
| Other: External consultant to foster communication | 2 |
| Other: Negotiation/reflection skills | 1 |
| Other: Interagency working/communication | 1 |

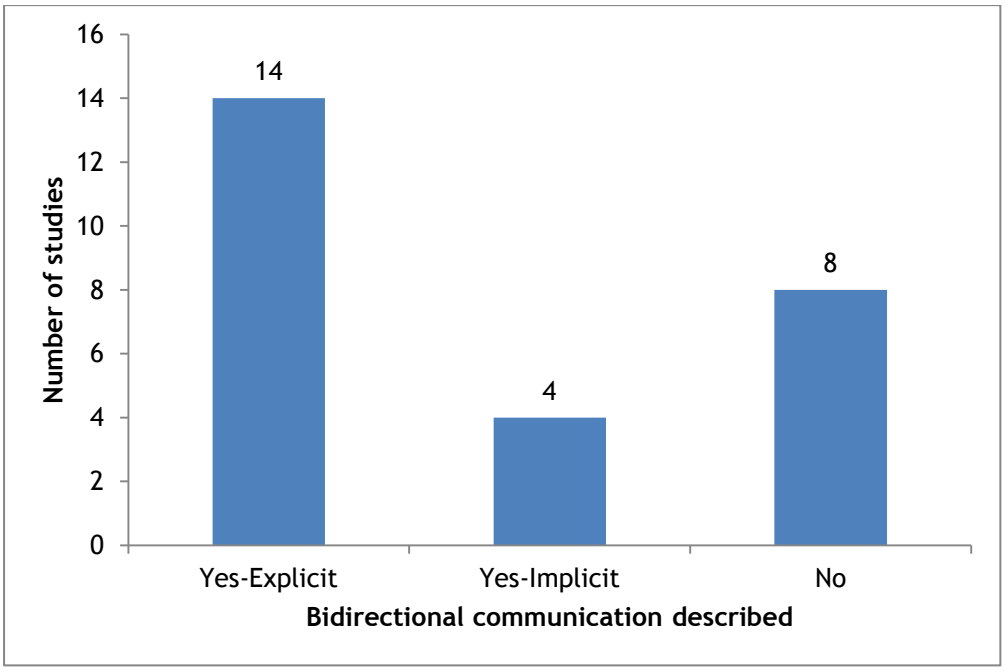
Of the processes originating from the previous review's conceptual framework, bidirectional communication, collective decision making and training support were the most frequently appearing processes described across the studies. Fewer studies provided evidence that allowed us to rate whether adequate time was given to develop collaborative relationships, or presented data concerning administrative support. Five new community engagement processes not identified in the original conceptual framework

were identified from a very low number of studies. These are listed as 'Other' in Table 3.1. The thematic analyses of each of these modifiable processes of community engagement are described below, in both overall terms and by age groups, gender and low-income groups where sufficient data permitted.

### 3.1 Bidirectional communication

To achieve good community engagement, it has been suggested that bidirectional communication should be sought between community members and other collaborative partners. Bidirectional communication can be thought of as the process by which partners listen and provide information to each other. Bidirectional communication such as this was described in 18 of the 26 studies (69%), as illustrated in Figure 3.1.

**Figure 3.1:** Bidirectional communication reported (n=26 studies)



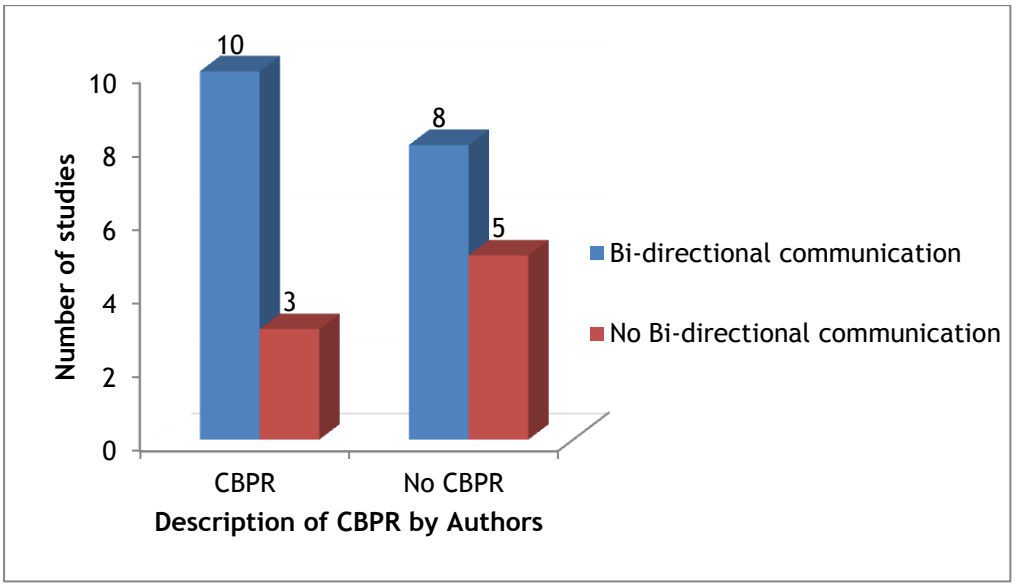

Most studies provided clear descriptions of bidirectional communication between partners: 14 of the 18 studies provided explicit descriptions of efforts, including descriptions such as 'fostering information exchange between community members and the research team' (Sussman et al. 2013: p.4). Four further studies (22%) described bidirectional communication implicitly, using phrases such as:

> The participation of Latinas from the same community in Healthy MOMs planning, design and implementation contributed to the cultural acceptability of its curriculum, activities and structure … it was guided by a steering committee of community resident women of childbearing age and representatives of community, academic and health-related organizations. (Kieffer et al. 2013: p.78).

#### *3.1.1 CBPR and bidirectional communication*

Bidirectional communication is considered a hallmark of CBPR (Wallerstein and Duran 2010). A total of 10 of the 18 studies that reported bidirectional communication also described the use of CBPR principles (56%), as shown in Figure 3.2.

**Figure 3.2:** Bidirectional communication and CBPR methods




However, three studies that claimed to base their study on CBPR principles did not also describe bidirectional communication. Further, eight studies that did not describe the use of CBPR methods did describe the use of bidirectional communication.

### *3.1.2 Extent of engagement and bidirectional communication*

A total of four studies were rated as having a high extent of community engagement, in that participants led or collaborated on design, delivery and evaluation. A high proportion of studies with high and moderate extents of engagement through design, delivery and evaluation, also provided evidence of bidirectional communication. All four of the 'high' and eight of the ten 'moderate' extent community engagement studies described bidirectional communication between collaboration partners. However, over half of the 11 studies rated as having a low extent of engagement also reported bidirectional communication (n=6, 55%).

### *3.1.3 Specific populations and bidirectional communication*

The use of bidirectional communication was proportionally more evident in studies that focused on men only and in those involving low-income groups. Only one of the five studies (20%) that focused exclusively on women reported bidirectional communication (Kneipp et al. 2011). Both of the studies targeting men only (100%) described this process of engagement (Andersen et al. 2013; Rhodes et al. 2011). Only two of the five studies (40%) directed towards children and young people reported the use of bidirectional communication (Berg et al. 2009; Bonell et al. 2010). Six of the eight studies (75%) targeting low-income groups reported the use of bidirectional communication (Berg et al. 2009; Cohen et al. 2013; Kneipp et al. 2011; Lassen et al. 2011; Martin et al. 2013; Phillips et al. 2014).

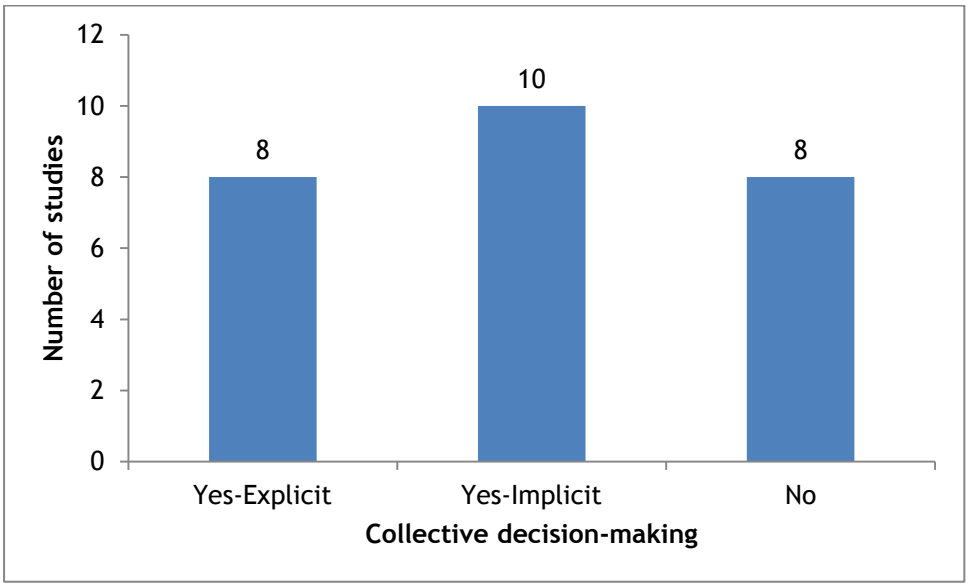
In summary, over two-thirds of the studies included in the analysis (18 of 26, 69%) reported bidirectional communication as a process of community engagement; 14 of these 18 studies described it explicitly. Bidirectional communication was reported in a high proportion of studies that were based on the principles of CBPR, and of those that were

not. Similarly, while a high proportion of studies with high and moderate extent of engagement through the design, delivery and evaluation utilised bidirectional communication, half of those rated as having a low extent of community engagement also provided evidence of this process. The use of bidirectional communication was proportionally more evident in studies that focused on men only and in those involving low-income groups.

### 3.2 Collective decision making

Collective decision making has been identified as an important process of community engagement (Wallerstein et al. 2008) Out of the 26 studies included in this analysis, just over two-thirds (n=18 studies, 69%) provided some evidence of community members' involvement with other partners in collective decision making. However, this was not consistently described across the studies. This is illustrated in Figure 3.3.

**Figure 3.3:** Collective decision making




In less than half of these 18 studies (n=8, 44%), 'shared' or 'collective' decision making was explicitly described (Kneipp et al. 2011; Kong et al. 2013; Parikh et al. 2010; Phillips et al. 2014; Plescia et al. 2008; Rhodes et al. 2011; Segal et al. 2011; Zoellner et al. 2013). The remaining ten studies (56%) were judged by our review team to have implicit evidence of collective decision-making. This was evidenced by descriptions such as:

> The worksites themselves were responsible for initiating and implementing activities to achieve a high level of local project ownership. (Lassen et al. 2007: p.728)

> The study design, staffing plans and recruitment, retention, intervention, and evaluation methods and materials were developed by a community-based steering committee in accordance with community-based participatory research (CBPR) principles. (Kieffer et al. 2014: p.526)

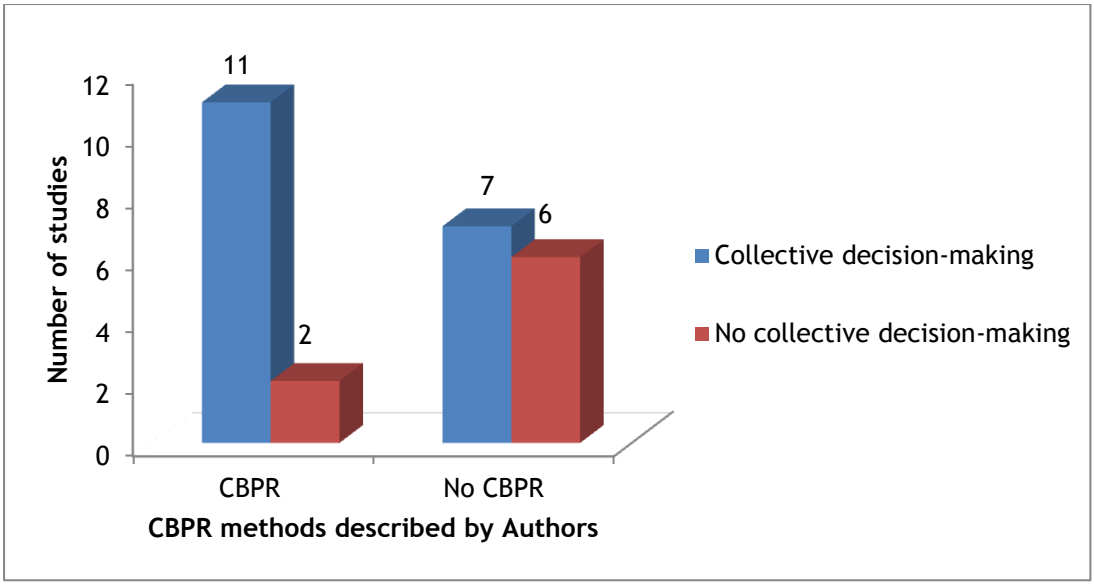
While neither quote explicitly identifies collective decision making, it can be inferred from the authors' descriptions of the project processes. Thus collective decision making

was apparent in a large proportion of the included studies, but was not always specifically referenced as a process of community engagement.

### *3.2.1 Community-based participatory research (CBPR) and collective decision making*

Despite reports from authors such as the quote above that CBPR principles were utilised, evidence of important processes enacting those principles did not appear to be consistently demonstrated within these studies. Thirteen of the 26 studies (50%) described explicit use of CBPR methods. Of the 18 studies that described collective decision making methods, 11 (61%) also described using CBPR methods. These are illustrated in Figure 3.4.

**Figure 3.4:** Collective decision making and CBPR methods




Collective decision making within the CBPR studies was not consistently evident. Only five of the thirteen studies (38%) reporting CBPR as the study's theoretical underpinning described explicitly the use of collective decision making. We were able to infer from descriptions in a further six studies using CBPR methods that collective decision making probably occurred. However, two studies that described using CBPR methods did not provide any evidence of collective decision making (Harper et al. 2009; Wright et al. 2013). Conversely, of the 13 studies that did not describe using CBPR methods, two explicitly described collective decision making and another five studies inferred collective decision making. These were evidenced by statements such as:

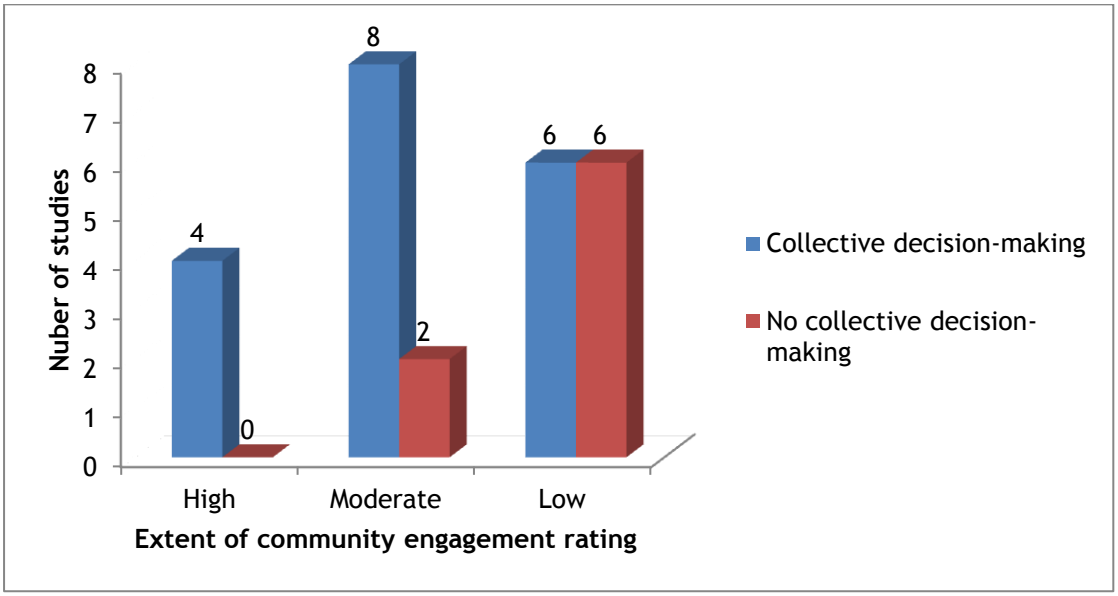
> Toward the conclusion of the six CAC sessions in the first year of the study, the group reached consensus on three sections to feature on the DVD: (1) adolescent motivation for change, (2) strategies targeting energy balance and nutritional quality, and (3) physical aerobic dance and strength/resistance training instructional segments. The CAC reviewed each of the elements to be included in the DVD as well as offered stylistic suggestions relating to background music and video editing techniques to appeal to adolescents. (Kong et al. 2013)

This suggests that interventions using CBPR methods did not always use collective decision making as an explicit process, and that some studies using other (often undescribed) philosophical underpinnings did use collective decision making.

### *3.2.2 Extent of engagement and collective decision making*

Collective decision making appeared to be used consistently in studies that were also rated as having high or moderate extents of community engagement across project design, delivery and evaluation. These are shown in Figure 3.5.

**Figure 3.5:** Collective decision making and extent of community engagement




Of the fourteen studies having either a high or moderate extent rating in this domain, only two (Andrews et al. 2012; Harper et al. 2009) did not describe collective decision making. Interestingly, a total of six of the 12 studies (50%) rated as having a low extent of community engagement described collective decision making.

### *3.2.3 Specific populations and collective decision making*

Collective decision making did not appear to be a process used more often in studies targeted exclusively to women or men compared to studies targeting mixed sex. For example, only one study out of five focused specifically on women (Kneipp et al. 2011) and only one of two studies targeting men (Rhodes et al. 2011) reported collective decision making as a process of community engagement. However it was used as a process more often with children/young people: three of five studies (60%) directed specifically toward children/young people described collective decision making (Berg et al. 2009; Bonell et al. 2010; Hoelscher et al. 2010). This finding might be due to the substantive areas: screening and physical activity for women/men (Hoelscher et al. 2010) versus 'whole school' approaches for children/young people (Berg et al. 2009; Bonell et al. 2010). Collective decision making was also used more often where authors identified low-income populations as the group of interest. Of the eight studies targeting low-income populations, six studies (75%) described collective decision making processes (Berg et al. 2009; Cohen et al. 2013; Kneipp et al. 2011; Lassen et al. 2011; Martin et al. 2013; Phillips et al. 2014).

In summary, collective decision making was identified as a process in over two-thirds of the studies included in this analysis; however authors did not often describe it explicitly as a process of community engagement. As a hallmark of good practice in CBPR, one might

expect to find collective decision making referred to consistently in studies that also reported CBPR principles. However only 60% of studies using CBPR method also reported collective decision making processes. Studies rated as having a high or moderate extent of community engagement across design, delivery and evaluation reported collective decision making, but also over half of the studies with a low extent of community engagement. Studies targeting women or men only did not tend to use collective decision making, but studies focused on children and young people did. Collective decision making was used most often in studies targeting low-income populations.

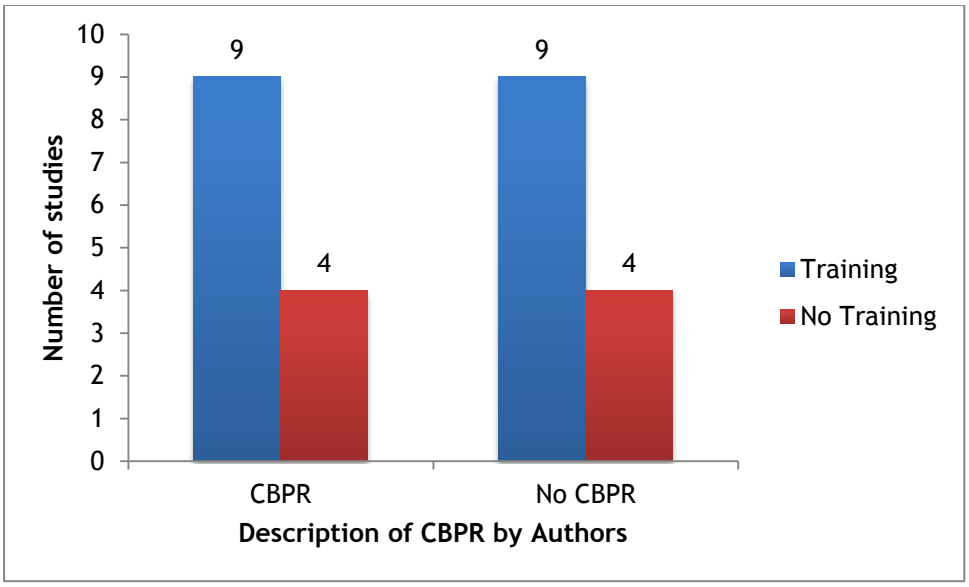
### 3.3 Training support

A total of 18 of the 26 included studies (69%) provided evidence of training support. In all but one of these studies, training was for the purposes of intervention provision rather than to function as participating members of collaborative groups with researchers and/or service providers. Only one study, by Bonell et al. (2010), offered teachers training in communication skills and class management techniques in addition to intervention provision training. The recipients of training across these studies were most often peers (n=9) and staff members (n=7). Three studies trained community engagees, all in intervention provision. Cohen et al. (2013) offered training to park advisory board members on making parks more useable through 'outreach, the importance of visibility and excellent customer service, and how to use special events to promote routine activities and programs' (Derose et al. 2014: p.15). Kneipp et al. (2011) trained community members hired as research staff for their respective roles and intervention fidelity. And Phillips et al. (2014) trained neighbourhood volunteers to support other residents to participate in the intervention.

#### *3.3.1 CBPR and training support*

Training support occurred as often in studies which did report CBPR methods as in those that did not: nine studies reported the use of CBPR and provided training support; and an equal number provided training without CBPR methods. These are shown in Figure 3.6. This suggests that training support is not a process of engagement related to the use of CBPR methods.

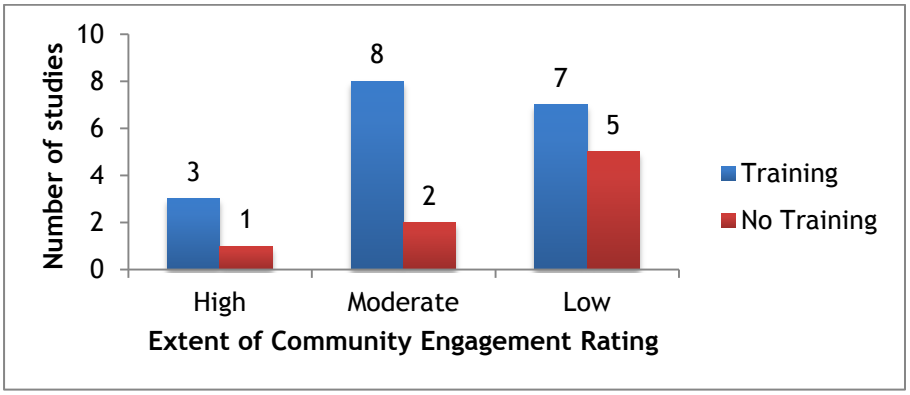
**Figure 3.6:** Training support and CBPR




*3.3.2 Extent of community engagement and training support*

Studies rated as having a high or moderate extent of engagement were slightly more likely to also offer training support, as shown in Figure 3.7. Three of the four studies that were rated as having a high extent of community engagement across design, delivery and evaluation provided training support (Berg et al. 2009; Cohen et al. 2013; Islam et al. 2013), as did eight of the ten moderate extent of engagement studies and six of the eleven low extent of engagement studies.

**Figure 3.7:** Training support and extent of community engagement




*3.3.3 Specific populations and training support*

Across the studies focused on specific populations, training for intervention provision was provided. Studies that provided training support occurred slightly more often when targeted to women than to men. Three of the five studies (60%) that focused specifically on women utilised training support (Eades et al. 2012; Kneipp et al. 2011; Russell et al. 2010), and one (50%) of the studies that focused on men (Rhodes et al. 2011). A similar pattern was noted in studies focused on children and young people. Three of these five studies provided training (Berg et al. 2009; Bonell et al. 2010; Hoelscher et al. 2010). All three provided training for intervention provision only, but Bonell et al. also reported

training for teachers in communication skills and class management techniques. Studies working specifically with low-income populations appeared more likely to provide training support: six of these eight studies (Berg et al. 2009; Cohen et al. 2013; Kneipp et al. 2011; Martin et al. 2013; Phillips et al. 2014; Russell et al. 2010). Three of these studies provided intervention provision training to community engagees (Cohen et al. 2013; Kneipp et al. 2011; Phillips et al. 2014).

In summary, studies of training support as a process of community engagement were evident in over two-thirds of the included studies. These tended to focus on staff or peer training to provide an intervention offered to peers, staff and less often to community engagees as advisory board members or researchers. Interventions with low-income populations were more likely to contain a component of training, and the three studies which offered training to community engagees focused on low-income populations.

### 3.4 Adequacy of time to develop a collaborative relationship

In order to understand whether there was adequate time for community members and other partners to develop a collaborative relationship, we extracted information reported on the frequency, duration and length of time partners met. Only 11 of the 26 studies included in the analysis (42%) provided information about the meetings between community members and other collaboration partners. The method of reporting varied widely with respect to how often coalitions met, and for how long, and over what period of time. This is shown in Table 3.2.

**Table 3.2:** Frequency, duration, timing of meetings and adequacy rating

| Study | Meeting Frequency | Meeting Length | Meeting Duration Period | Intervention Duration | Adequate Time to Develop Relationships |
|---|---|---|---|---|---|
| Berg (2009) | Daily | 4 hours | 7 weeks | 52 weeks | Yes |
| Bonell (2010) | Ten times | -- | -- | 36 weeks | Yes |
| Cohen (2013) | Four times | -- | -- | -- | Yes* |
| Rhodes (2011) | Weekly | -- | -- | 36 weeks | Yes* |
| Zoellner (2013) | Six times | -- | 27 weeks | 27 weeks | Yes* |
| Hoelscher (2010) | Twice | -- | 36 weeks | 104 weeks | No* |
| Kong (2013) | Monthly | -- | 27 weeks | 104 weeks | No |
| Lassen (2011) | Once | -- | -- | -- | No |
| Russell (2010) | -- | -- | -- | 78 weeks | No |
| Woods (2013) | Four times | 4 hours | -- | 68 weeks | No* |
| Wright (2013) | Four times | -- | 52 weeks | 104 weeks | No* |

'--' : Not reported

Based on the completeness of the descriptions of the meetings provided by authors, which are summarised in Table 3.2, we judged five of the studies included in analysis (19%) to have provided evidence of adequate time for collaborative relationships to develop.

### *3.4.1 CBPR and relationship development*

Three of the five studies that demonstrated adequate time for relationship development (60%) also reported the explicit use of CBPR methods (Cohen et al. 2013; Rhodes et al. 2011; Zoellner et al. 2013), compared to three of the six studies (50%) that did not demonstrate adequate time to develop relationships. This suggests that studies that allowed for adequate time to develop collaborative relationships were slightly more likely to be based on CBPR methods than those that did not allow for relationship development.

### *3.4.2 Extent of community engagement and relationship development*

The extent of engagement did not appear to be related to the development of collaborative relationships. Two of these five studies were rated as having a high extent of community engagement across design, delivery and evaluation (Berg et al. 2009; Cohen et al. 2013). Two studies were rated as having a moderate extent (Bonell et al. 2010; Rhodes

et al. 2011) and one was considered to have a low extent (Zoellner et al. 2013). This distribution was similar for studies which did not provide adequate time for relationship development or did not report it.

### *3.4.3 Specific populations and relationship development*

No studies directed toward women alone were judged to have adequate time for relationships to develop, and one of the two studies targeting men specifically (Rhodes et al. 2011). Two of the five studies focused on children and young people were judged to have provided adequate time for relationship development (Berg et al. 2009; Bonell et al. 2010), and just two of the eight studies involving low-income groups (Berg et al. 2009; Cohen et al. 2013).

In summary, less than half of the included studies (n=11) provided enough data to determine whether adequate time was given to allow for the development of working relationships between community members and other collaborative partners. Of those, only five studies did allow adequate time for this important process of engagement. A slightly higher proportion of these five studies described the use of CBPR methods, suggesting that CPBR methods emphasise time for adequate relationship development. No pattern in relationship development could be discerned for studies rating as having high, moderate or low extents of community engagement. Two of five studies allowing adequate time for relationship development focused on children and young people, and two studies focused on low-income populations.

## 3.5 Administrative support

Only one study described the presence of administrative support. Dzewaltowski et al. (2010) undertook a CBPR project to improve the capacity of after-school staff to increase physical activity and improve dietary intake in children to prevent childhood obesity. In this study, the authors reported provision of 5% contribution towards a Family and Consumer Science County agent to conduct local community development work, attend school wellness council meetings and work with food services. In addition, they reported the provision of a salary for a half-time Cooperative Extension assistant, who co-ordinated the staff training and delivery of the intervention.

## 3.6 Other processes

Five other processes were identified which had not originated from the conceptual framework. These included: conflict resolution; arranging meetings to suit community members' needs; use of external consultants; negotiation/reflection skills development; and interagency communication. These processes are described narratively below. However, they have not been further synthesised as the number of studies reporting each process are too small to permit meaningful comparisons.

### *3.6.1 Conflict resolution*

The framework synthesis identified two studies reporting on the process of conflict resolution between collaboration partners (Berg et al. 2009; Plescia et al. 2008). Berg and colleagues (2009) used peer research as a way to increase feelings of collective empowerment and reduce or delay the onset of risky behaviour around drugs and sex amongst young people in a US city. In this study, young people were supported in

undertaking different research projects with the aim of helping to understand the needs and identify potential solutions for social issues in their community. Plescia and colleagues (2008) developed a lay health advisor programme in the US to target and improve three behavioural risk factors for heart disease and diabetes. Both studies appeared to describe the role of conflict resolution in community engagement delivery differently: Plescia et al. (2008) described conflict as something to be mediated and largely resolved by a third party. However, Berg et al. (2009) described cognitive conflict and negotiating conflicts of option choices as steps toward 'changing peer culture' and 'supporting pro-prevention norms' among young people.

#### *3.6.2 Arranging meetings to suit community members' needs*

Two studies described arranging collaborative meetings in ways to suit the needs of community members. These included meetings at trusted settings, providing childcare and transport (Kieffer et al. 2013) and scheduling the meeting times to suit community members' needs (Kong et al. 2013).

#### *3.6.3 Use of external consultants*

Two studies reported the use of an external consultant or facilitator to foster communication between coalition partners. Plescia et al. (2008) made use of an external consultant to resolve conflicts and provide suggestions for collaborative communications. A workshop facilitator was employed to ensure that all group members participated in information sharing and decision making in the study by Zoellner et al. (2013).

#### *3.6.4 Negotiation/reflection skills*

The study by Berg et al. (2009) reported the development of 'reflection skills' by community members over the course of the study. In this study, community engagement was the intervention under study; thus community members learned negotiation and reflection skills as part of the intervention.

#### *3.6.5 Interagency working/communication*

One study reported fostering the process of interagency working and communication between county-level community services agency and the local school level, where the intervention was implemented (Dzewaltowski 2010).

### 3.7 Tests of convergence

#### *3.7.1 Bidirectional communication and collective decision making*

It could be argued that bidirectional communication and collective decision making are closely related concepts, and may measure the same concept. We assessed the congruence of these two processes of community engagement by comparing the ratings for both processes across all studies. In all but four studies (n=22, 85%) the two processes were congruent. Two of these studies described bidirectional communication but not collective decision making, evidenced by data including:

> In close collaboration … we developed… (Andersen et al. 2013: p.102)

> we consulted with key opinion leaders to develop and review the materials… (Andrews et al. 2012: p.162)

Two other studies reported collective decision making but in the absence of data supporting bidirectional communication. This was illustrated, for example, by:

> design was jointly discussed and developed in conjunction with our Hmong community collaborators (Chen et al. 2013: p.786)

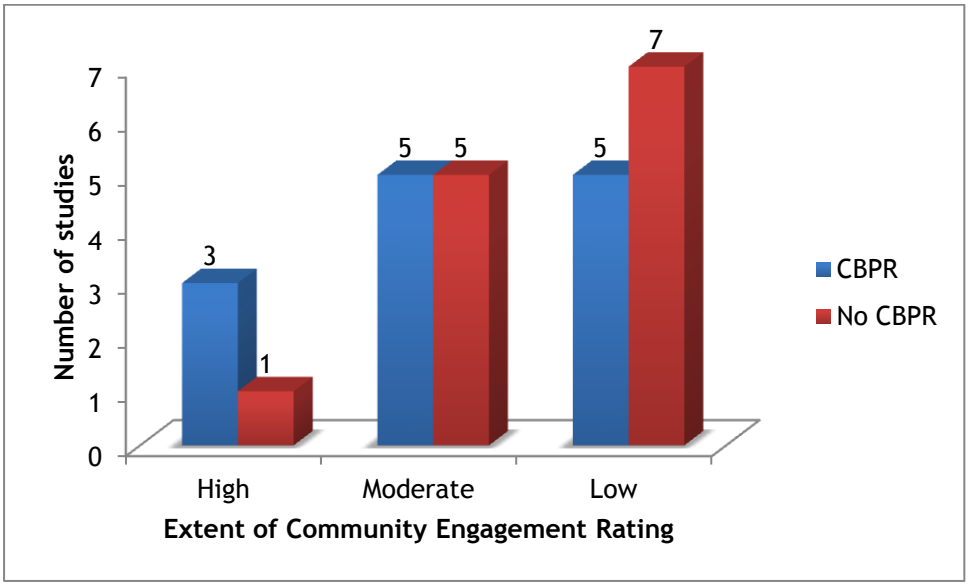
> CATCH Community Action Teams were asked to … select an activity each semester from a CATCH Community Café menu … we elected to include the community partners in various decisions (Hoelscher et al. 2010: p.43)

These findings suggest that while bidirectional communication and collective decision making appear to be convergent concepts, bidirectional communication does not necessarily lead to collective decision making (or the latter is just assumed and not reported), and that collective decision making may take place in an environment that does not necessarily report or privilege bidirectional communication.

### *3.7.2 CBPR and extent of community engagement*

Similarly, it could be argued that community-based participatory research and ratings of the extent of community engagement were measuring the same concepts. To test this idea, we examined the convergence of CBPR and the extent of community engagement across the 26 included studies. These findings are illustrated in Figure 3.8.

**Figure 3.8:** Convergence of CBPR and extent of community engagement




If CBPR and the extent of community engagement ratings were measuring the same concept, one could expect the proportions of those that are and are not CBPR, for each rating of extent, to be the same. The figure illustrates that while the two concepts overlap, they are probably measuring slightly different concepts.

# 4. Findings: meta-analysis and modelling

## 4.1 Chapter summary

- In most cases, studies were too heterogeneous to produce a summary effect size statistic for self-efficacy, behavioural or clinical outcomes. Further analyses were conducted to examine factors moderating the variability of effect between studies.
- A meta-analysis of five studies examined behavioural *change* outcomes (measured immediately post-intervention; all five aimed to improve healthy eating and physical activity), indicated that community engagement coalitions had no effect upon health behaviour d = 0.005, 95%CI: 0.064 to 0.007, Q = 2.68, $I^2$ = 0).
- When we focused on those studies that collected clinical measurements we found that there was little between-study variance (Q=5.17, df=9). Our pooled effect size for this group of studies stood at 0.192 (CI: 0.092-0.292), suggesting that interventions based on coalitions for community engagement made a small impact on clinical health outcomes for measures collected longitudinally.
- Analyses suggested that deploying multiple community engagement processes may be associated with statistically significantly larger effect sizes, but no one individual process could be attributed with increased effectiveness in the meta-analyses.
- A high extent of involvement was associated with higher effect sizes for interventions with behavioural outcomes measured longitudinally. Compared to interventions with a low extent of involvement, studies with a high extent of involvement had an effect size (*d*) that was larger (over half a standard deviation (0.673) larger). Extent of involvement was an important factor in measuring between-study variance accounting for 52%.
- There was insufficient evidence to determine whether direct comparisons of community engagement (i.e. community engagement intervention versus the same intervention without community engagement) differed in health outcome effects from indirect comparisons of community engagement.
- Studies using a longitudinal design and that focused on the screening for and prevention of injury or infection were associated with higher effect sizes for behavioural outcomes - compared to studies focusing on healthy eating or physical activity, these studies had higher average effect sizes in the region of 0.479 d; those with a focus on alcohol, tobacco or drugs had a lower average effect size than studies with an alternative focus such as injury prevention, healthy eating and physical activity or sexual health.
- It was not possible, using the evidence available, to determine whether or not the effect of community engagement interventions, in terms of producing variable health outcome effects for different age groups, genders or low-income groups, was a substantive finding, rather than a spurious finding attributable to other potential study-level confounders.
- There was insufficient data with which to test the correlation between self-efficacy outcomes and health behaviour, and the correlation between health behaviour and clinical or physiological outcomes.

## 4.2 About this chapter

The purpose of this chapter is to consider the effectiveness of interventions that incorporate community engagement (specifically those employing coalitions, collaborations or partnerships), compared with controlled conditions in which no or minimal community engagement is evident. Once the overall effectiveness of such interventions is ascertained, moderators of this effect will be explored.

As discussed in the Methods section (Chapter 2), moderators tested include community engagement characteristics (such as processes adopted for achieving community engagement) participant characteristics (such as age, gender or socio-economic status) and features of the evaluations (such as risk of bias). We updated and re-ran analyses from the original review using information from the studies located in the update. In addition, we attempted to address the following research questions:

1. Are potentially modifiable processes of community engagement associated with health outcome effects?
2. What is the relationship between the extent of community engagement (high, moderate or low) and health outcome effects?
3. Do direct comparisons of community engagement (i.e. studies that test a community engagement intervention versus the same intervention without community engagement) differ in health outcome effects from indirect comparisons of community engagement (e.g. those that test community engagement versus usual care)?
4. Do health outcome effects differ for:
    a. different age groups;
    b. studies targeting men only versus those targeting women only;
    c. studies specifically developed for low-income groups versus those that are not;
    d. 'distal' (e.g. self-efficacy), 'intermediate' (e.g. health behaviour) or 'proximal' (clinical/physiological measure) outcomes?

Before we synthesise the evidence to answer the RQs, we give a very brief overview of the theoretical framework underpinning the analyses, including the proposed causal pathway. This is followed by descriptive information about the participants, interventions, comparators and evaluation characteristics, outcomes and effect size estimates included in the meta-analyses. Thereafter we examine the heterogeneity of effect across studies, present syntheses addressing the above research questions and finally, report on sensitivity analyses, including an examination of publication bias and the impact of risk of bias upon intervention effect size estimates.

## 4.3 Theoretical framework

A conceptual framework describing a range of dimensions to explore and categorise differences between community engagement approaches was developed by O'Mara et al. (2013; see Chapter 8, page 112).

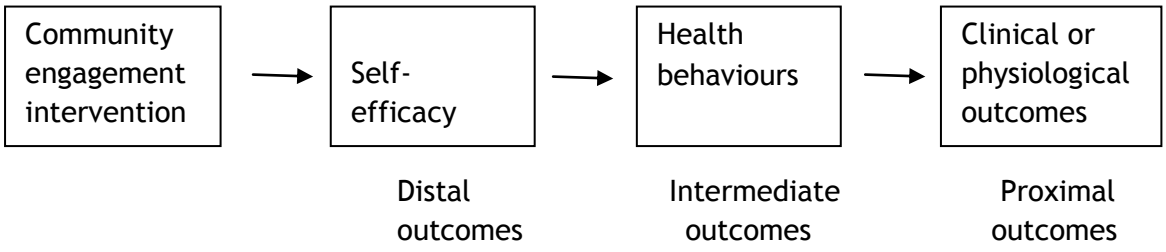
We extracted data from each of the primary studies on aspects of these dimensions to enable us to assess whether differences between the community engagement approaches

were associated with differing levels of intervention impact upon health consequences (clinical/physiological outcomes), health behaviours, and self-efficacy outcomes:

- community identified health need
- collaboration in intervention design
- consultation in intervention design
- lay delivery
- extent of community engagement (calculated as described later in this chapter)
- socio-demographic moderators
- modifiable processes of community engagement including: collective decision making; training support; administrative support; adequate time for relationship development.

Across studies, a variety of outcomes were used to assess impact. Following the theoretical development in the original review (O'Mara-Eves et al. 2013), a causal chain was assumed (see Figure 4.1) in which self-efficacy (i.e. self-esteem and belief in ability to change behaviour) needed to be changed in order for health behaviours (i.e. actions that people do, such as smoking, healthy eating and physical activity) to subsequently have an impact upon physiological or clinical outcomes (such as blood pressure and body mass index).

**Figure 4.1:** Proposed causal pathway from intervention to clinical/physiological outcomes.




We extracted available data on intervention effectiveness from the additional studies included in this update for the following outcomes:

- self-efficacy in relation to health behaviours;
- health behaviours – outcomes extracted were alcohol, tobacco and drug use, donor registration, healthy eating (e.g. fruit and vegetable consumption, dietary fat or sugar intake), physical activity, service use and sexual health behaviours (e.g. condom use);
- physiological or clinical consequences – outcomes extracted were body mass index, depression, mental health (Warwick-Edinburgh mental wellbeing scale), serum insulin levels, weight.

## 4.4 Description of the studies included in the meta-analysis

### *4.4.1 Included studies (N = 58)*

In this current update, *Community Engagement for Health Via Coalitions, Collaboration and Partnerships* (CERUB) we identified 26 studies concerning coalitions, collaborations or

partnerships for community engagement interventions intended to improve health. We were able to extract effect sizes from 20 of these 26 studies.

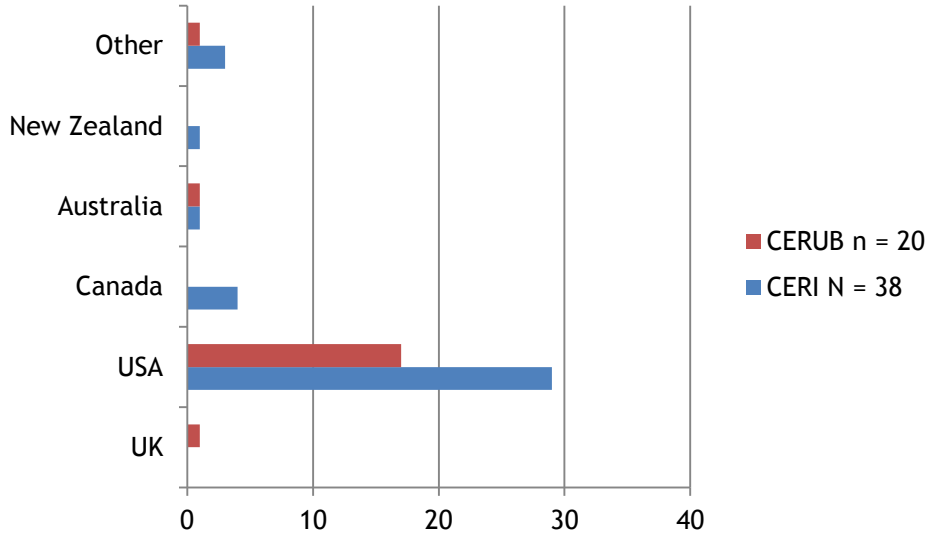
The findings of these studies were synthesised with the findings from a sub-set of studies (those which employed coalitions, collaborations or partnerships for community engagement interventions) from the original review by O'Mara et al. (2013) *Community engagement to reduce inequalities in health* (CERI) (n = 38).

Thus, a total of 58 studies were available for meta-analysis: 38 studies from the original CERI review, and 20 studies from the CERUB update review.

### *4.4.2 Country of origin*

Of the 58 studies included in the meta-analysis, one was conducted in the UK (Phillips et al. 2014), 46 (79.3%) in the USA, four (6.9%) in Canada, two (3.4%) in Australia, one in New Zealand, and four (6.9%) in other OECD countries (see Figure 4.2).

**Figure 4.2:** Country in which studies included in the meta-analysis (n = 58) were conducted




### *4.4.3 Publication date*

In terms of publication date, 5 (8.6%) of studies were published 1992-1994, 14 (24.1%) 1995-1999, 11 (18.9%) 2000-2004, 9 (15.5%) 2005-2009 and 19 (32.8%) 2010-2014. With the exception of Harper et al. (2009) all the new studies from the update of the review which were included within the meta-analyses (n = 19) were published between 2010 and 2014 (see Table 4.1).

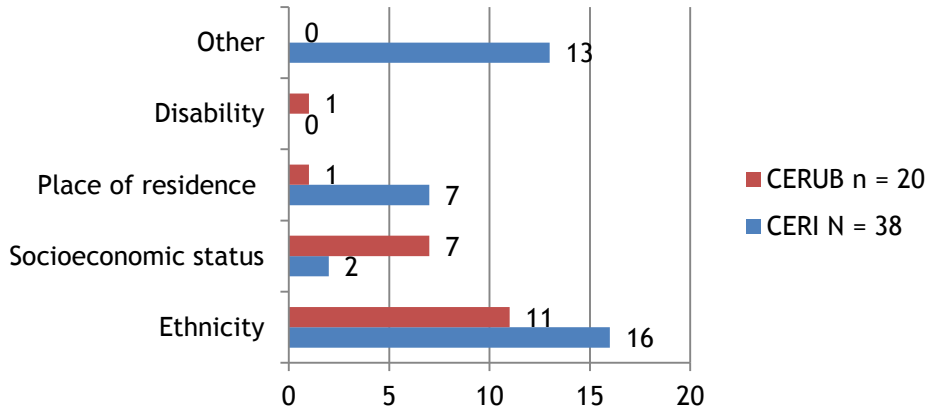
**Table 4.1:** Publication date for the studies in the review

| Publication date | Number of studies |
|---|---|
| 1992-1994 | 5 |
| 1995-1999 | 14 |
| 2000-2004 | 11 |
| 2005-2009 | 9 |
| 2010-2014 | 19 |

### *4.4.4 Primary marker of disadvantage*

In terms of the main characteristic of the study population that distinguished them as being disadvantaged, 27 (46.6%) studies were classified as being primarily targeted at, or delivered to, ethnic minority groups, followed by 9 studies (15.5%) aimed at people considered to be of low socio-economic position. Many studies had multiple PROGRESS-Plus categorisations;[1] the majority of these represented a combination of ethnic minority group status and low-income and/or inner-city status. Most of the ethnic minority participants were classified as African American or Hispanic/Latino. This is illustrated in Figure 4.3.

**Figure 4.3:** Primary marker of disadvantage of populations in studies included in the meta-analysis (n = 58)

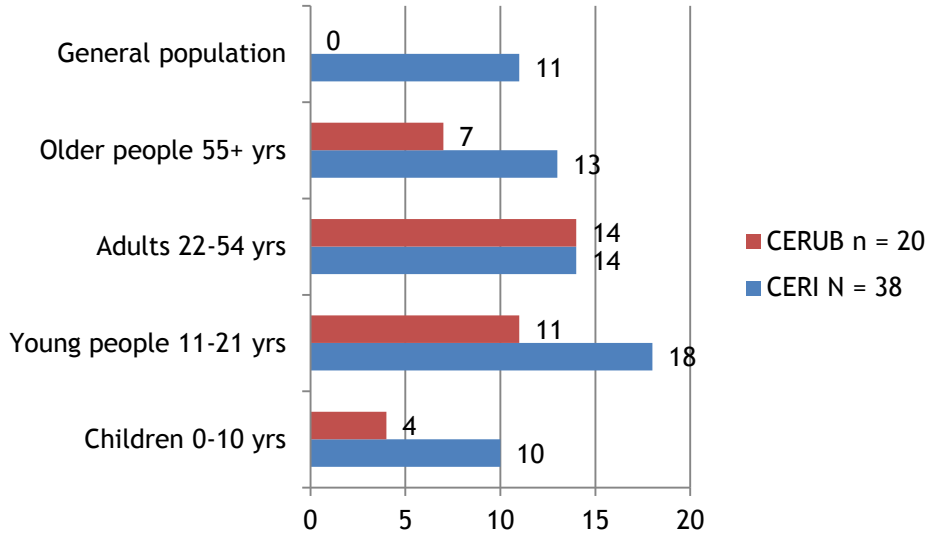



### *4.4.5 Age*

The studies did not focus predominantly upon one age group, such as young people or older people. Often study populations included a variety of age groups, although in the analysis we examined those that were categorised as focusing on children and young people only (see Figure 4.4).

[1] See the Glossary for a definition of PROGRESS-Plus.

**Figure 4.4:** Age groups of study populations in studies included in the meta-analysis (n = 58)

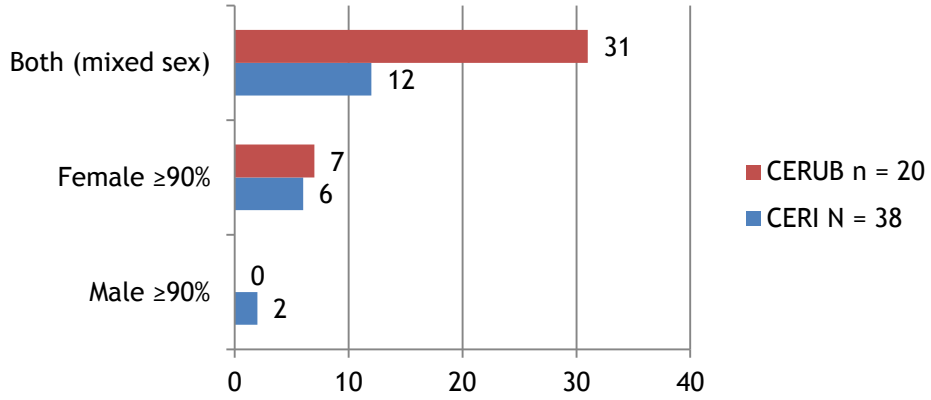



Note: Studies could examine more than one age group.

*4.4.6 Gender*

Of 58 studies included in the meta-analyses, only 2 (3.4%) focused solely on males (Andersen et al. 2013, Rhodes et al. 2011). Thirteen (22.4%) studies focused upon a predominantly female population. The majority of studies (n = 43) examined a mixed sex population (see Figure 4.5).

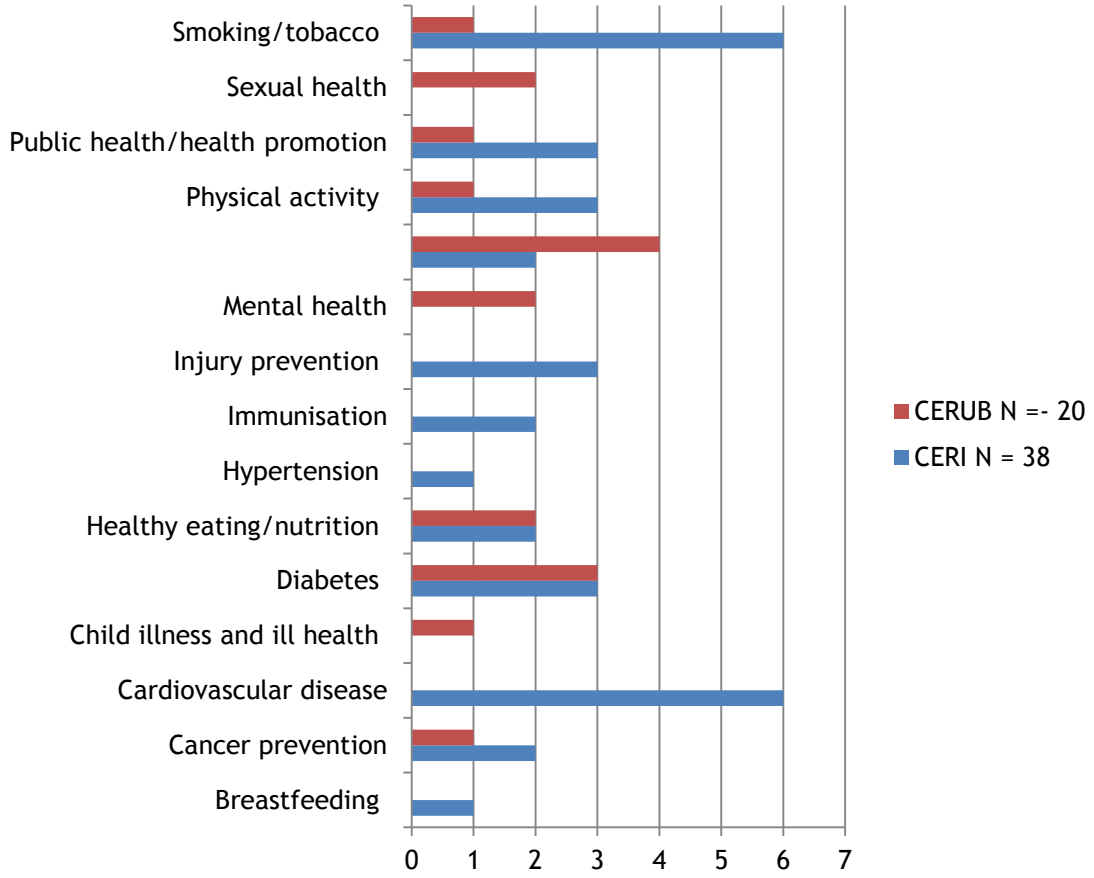
**Figure 4.5:** Gender of populations in studies included in the meta-analysis (n = 58)




*4.4.7 Health topic focus*

The interventions were conducted over a range of health topics (see Figure 4.6). The most commonly targeted health issues were smoking (n = 7, 12.1%), cardiovascular disease (n =6, 10.9%), diabetes (n = 6, 10.9%), and obesity or weight management (n =6, 10.9%). In our later analyses, we group several of these categories together.

**Figure 4.6:** Primary health issues targeted by the interventions in studies included in the meta-analysis (n = 58)




### 4.5 Effect size estimates across different outcome domains

Studies could contribute to more than one effect size estimate under the following conditions:

- when there were both immediate post-test and follow-up measures (to test the persistence of effects over time)
- when there were outcomes from different points in the causal pathway i.e. self-efficacy outcomes, health behaviour outcomes or clinical/physiological outcomes (see Figure 4.1 and Table 4.2)
- when standardised mean differences were recorded using means, or alternatively, mean changes (means and mean changes are not equivalent, often having substantially different standard errors, and therefore have been considered separately in this analysis).

Due to the conceptual and statistical heterogeneity in outcome domains, we did not view combining different domain types, measurement points, or data types as appropriate. As such, the meta-analyses in this chapter constitute a number of stratified analyses.

**Table 4.2:** Number of studies contributing to separate meta-analyses by outcome domain, measurement point and mean change outcome type.

| | Outcome type | Measurement point | Change data | Number of studies included | CERUB |
|---|---|---|---|---|---|
| 1 | Behavioural | Post-test | No | 49 | 16 |
| 2 | Behavioural | Post-test | Yes | 5 | 3 |
| 3 | Behavioural | Follow-up | No | 6 | 2 |
| 4 | Clinical | Post-test | No | 19 | 7 |
| 5 | Clinical | Post-test | Yes | 7 | 6 |
| 6 | Clinical | Follow-up | No | 1 | 1 |
| 7 | Self-efficacy | Post-test | No | 10 | 4 |
| 8 | Self-efficacy | Post-test | Yes | 4 | 2 |
| 9 | Self-efficacy | Follow-up | No | 2 | 2 |

Adjusted effect size estimates were used in preference to unadjusted when the choice was available. Where more than one outcome was reported per domain, the primary outcome was selected for that domain, with reference to the aims and objectives of the intervention. Effect sizes were extracted from each study by two reviewers working independently, who then met to resolve any discrepancies. When necessary, we contacted authors for missing data to enable effect sizes to be calculated. The intra-cluster correlation coefficient was used to adjust for the impact of clustering upon effect size in cluster randomised controlled trials. Estimates of the intra-cluster correlation were imputed using estimates derived from other similar studies included in the review when not available in the report or obtainable via author contact.

Multiple effect size estimates were calculated for some studies, a total of 103 across 58 studies. Of the 58 studies, 94 were calculated from post-test measurements and 9 from follow-up measurements. It was not possible to examine persistence of intervention effects across outcome domains due to the small number of studies reporting final follow-up outcomes. The majority of the remainder of this section refers only to the post-test effect size estimates (i.e. not follow-up and not representing mean or other change data).

For post-test effect size estimate,[2] 31 studies contributed an effect size estimate to only one domain: 27 studies had a behaviour outcome only and 4 studies a clinical outcome only. Only three studies (Islam et al. 1013; Resnicow et al. 1992; Zoellner et al. 2013)

[22] Which did not represent change data

provided an effect size estimate for all three domains (i.e. self-efficacy, behaviour and clinical outcomes).

Effect size estimates based on continuous data are represented by Cohen's 'd', whereas log-odds ratio (LOR) was used to represent effect size estimates calculated from binary data. To enable the synthesis of the effect size outcome across both binary and continuous measures, the LOR effect size estimates were converted to 'd' effect size estimates using the procedures described by Chinn (2000).

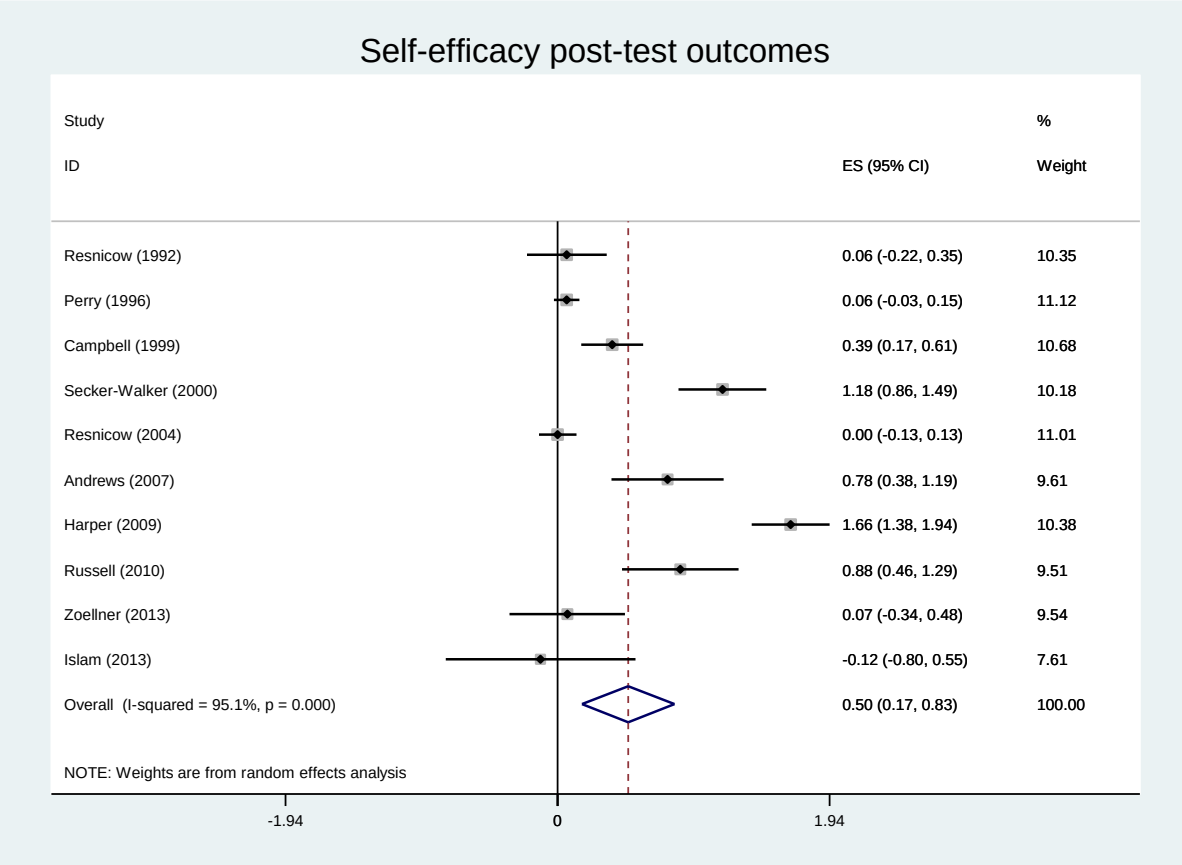
For self-efficacy outcomes, effect size estimates ranged from *d* = -0.122 to d = 1.663 and did not require Winsorising. For health behaviour outcomes, effect size estimates ranged from *d* = -0.098 to *d* = 1.756. The range of values after outliers was Winsorised to two standard deviations above or below the mean. Two post-test effect sizes were Winsorised (Macaulay et al. 1997; Schwarz et al. 1993). Both were more than two standard deviations above the mean. For clinical outcomes, effect size estimates ranged from *d* = -0.034 to *d* = 0.519. The range of values after outliers was Winsorised to two standard deviations above or below the mean. One post-test effect size (more than two standard deviations above the mean) was Winsorised, that of Kieffer et al. (2013).

Data were reverse coded when necessary such that a positive value indicated that the intervention favoured the intervention or treatment group. A positive d indicates that participants in the treatment group, on average, scored higher than those in the control group. An effect size estimate of d = 1 means that participants in the treatment group scored, on average, one standard deviation higher than participants in the control group on the outcome measured.

### 4.6 Synthesis

Figures 4.7, 4.8 and 4.9 show the effect size estimates, confidence intervals and relative weight for each intervention by outcome type (i.e. self-efficacy, behaviour and consequences or clinical outcomes). In the following sections we report upon the heterogeneity of the studies in each synthesis, after which the results of the quantitative syntheses are organised according to the research questions proposed. Finally, we report on sensitivity analyses to examine the effect of different analytical strategies upon our results.

**Figure 4.7:** Forest plot of effect size estimates and standard errors of studies reporting post-test self-efficacy outcomes (n = 10)




As can be seen from Figure 4.7, although largely positive, indicating a favourable impact of the intervention, the effect sizes for the 10 studies examining post-test self-efficacy outcomes were highly variable and widely dispersed: Q = 182.1, $I^2$ = 95.1% (See Table 4.3). The studies were too heterogeneous to produce a summary effect size statistic although further analyses were conducted to examine factors moderating the variability of effect between studies.

Figure 4.8 displays the effect sizes for the 49 studies examining post-test health behaviour outcomes. Although largely positive, indicating a favourable impact of the intervention, the effect sizes were highly variable and widely dispersed across this group of studies: Q = 297.8, $I^2$ = 83.9% (see Table 4.3). The studies were too heterogeneous to produce a summary effect size statistic. Nevertheless, further analyses were conducted to examine factors moderating the variability of effect between studies.

**Figure 4.8:** Forest plot of effect size estimates and standard errors of studies reporting post-test health behaviour outcomes (n = 49)

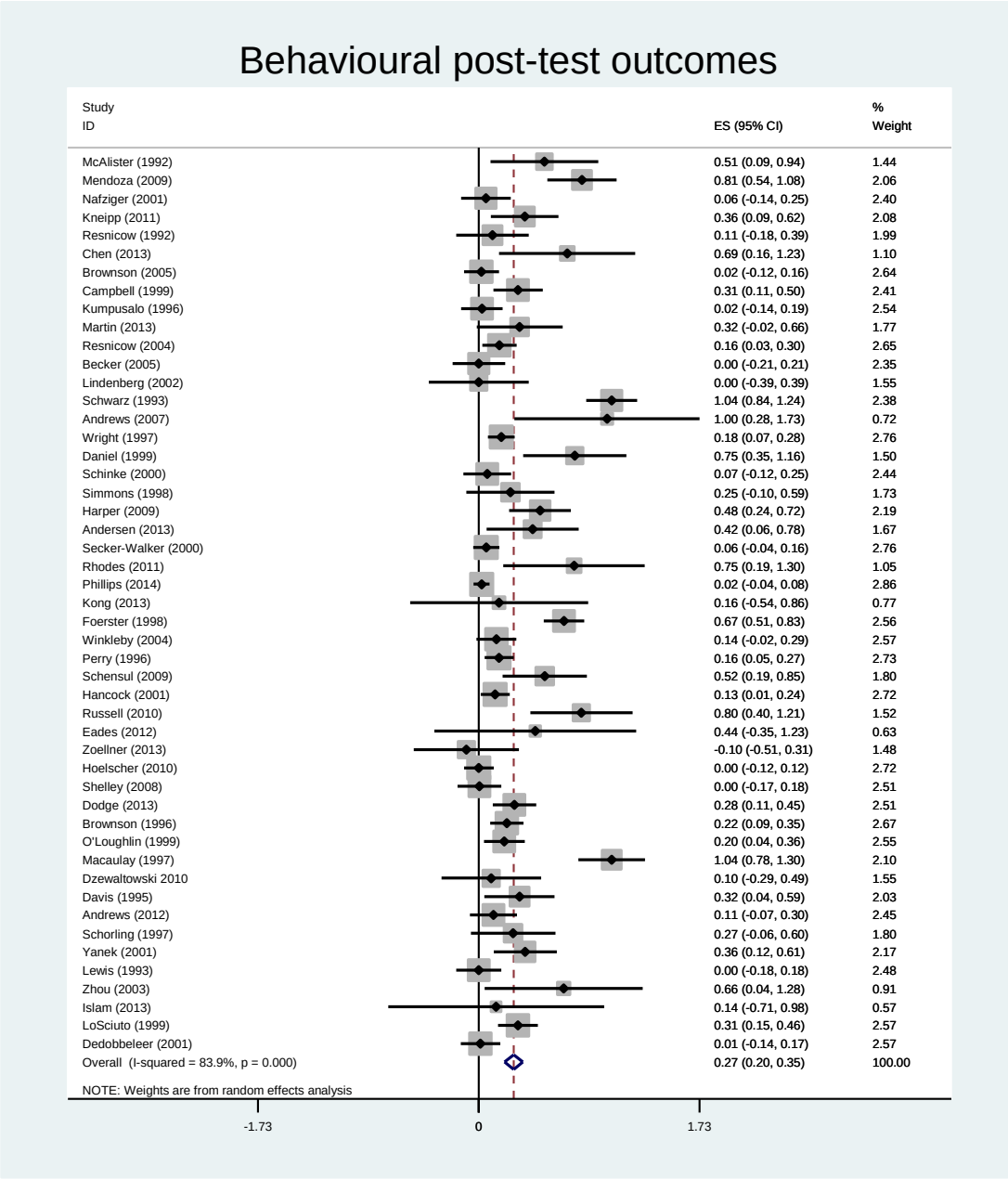
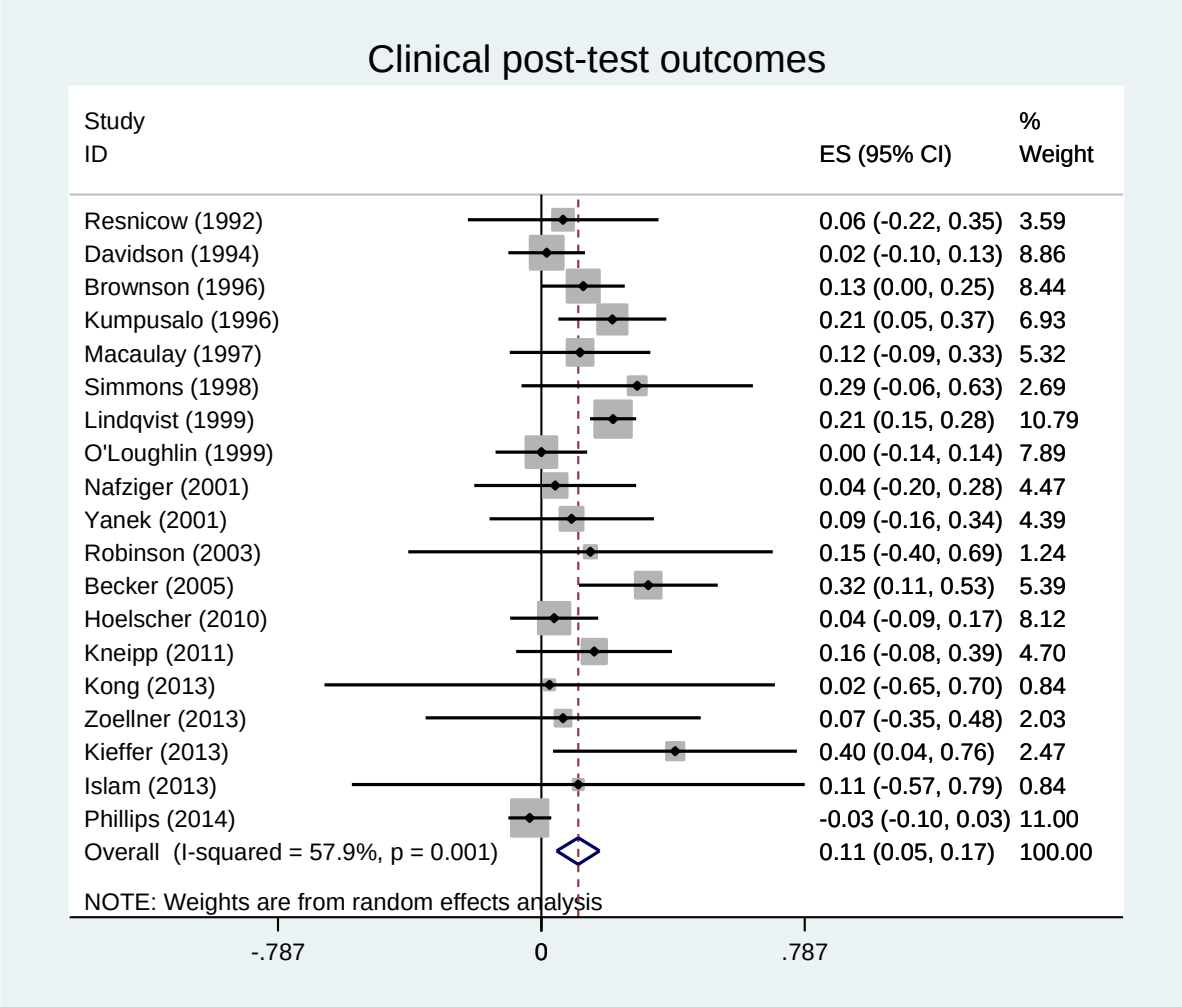



Figure 4.9 displays the effect sizes for the 19 studies examining post-test clinical or physiological health consequences outcomes. As for self-efficacy and behaviour outcomes, the effect sizes were largely positive, indicating a favourable impact of coalitions for community engagement interventions upon outcomes. However, the effect sizes were too heterogeneous to produce a summary effect size statistic: $Q = 42.8$, $I^2 = 57.9\%$. Notwithstanding, further analyses were conducted to examine factors moderating the variability of effect between studies.

Two factors are relevant to determining whether it makes sense to combine outcomes: between-group heterogeneity and the direction of each sub-group's pooled effect size estimate. Whereas the pooled effect size estimate for each outcome domain (measured at

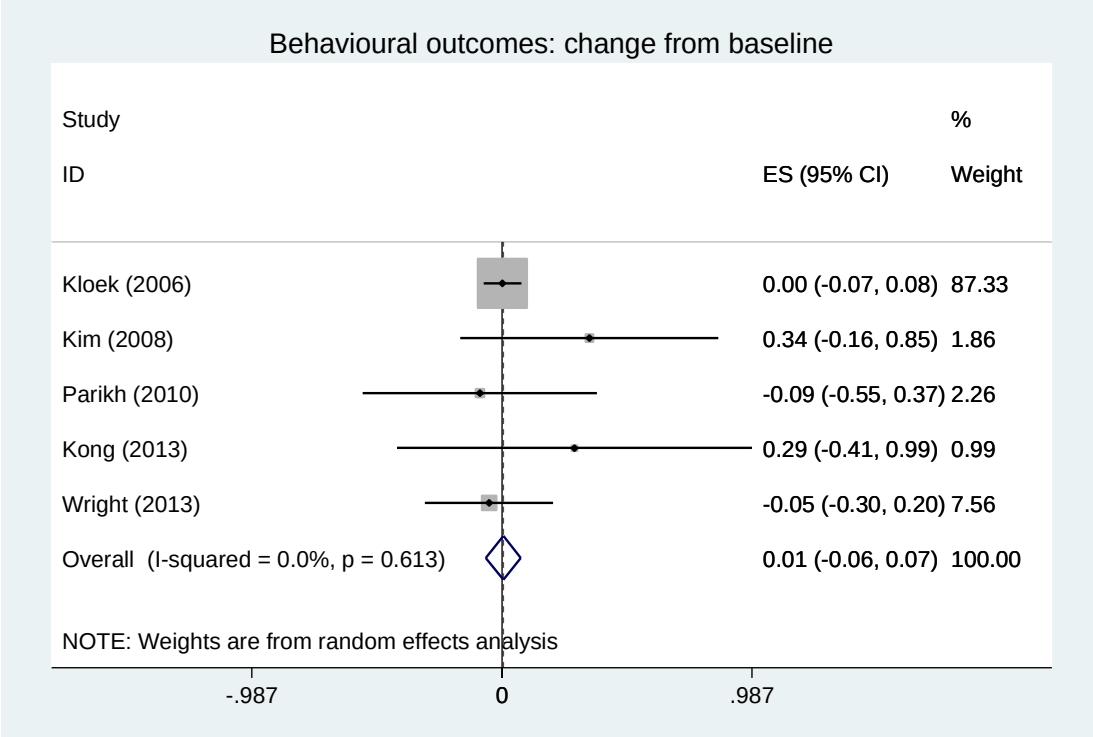
post-test and not directly measuring changes) was statistically significant from 0 in a positive direction, the variation in the magnitude of effects of these results suggested that in most cases, on the basis of statistical differences, it was not appropriate to pool the effect sizes for self-efficacy, behavioural and clinical outcomes in the analyses. However, such a large degree of heterogeneity can be explored further to examine which study level factors may be most associated with this, and these analyses form the basis of much of the remainder of the chapter.

**Figure 4.9:** Forest plot of effect size estimates and standard errors of studies reporting post-test clinical health consequences outcomes (n = 19)




A meta-analysis of five behaviour *change* studies examining community engagement interventions (all intended to improve healthy eating and physical activity behaviours; see Figure 4.10) had sufficiently homogenous studies to produce an overall pooled summary estimate of effect: d = 0.005, 95%CI -0.064 to 0.0075, Q = 2.68, $I^2$ = 0).

**Figure 4.10:** Forest plot of effect size estimates and standard errors of studies reporting change in health behaviour outcomes (n = 5)




The pooled effect size estimate indicated that community engagement coalitions for the purposes of improving healthy eating and physical activity had no effect upon health behaviour outcomes. It should be noted that one study involving adults aged 18-65 years living in deprived neighbourhoods in Eindhoven, The Netherlands, contributed a substantial proportion towards the weighted estimate (87%).

### 4.7 Exploring heterogeneity

Because the interventions covered a broad range of health topics and health outcomes which we have combined in analyses under the domains of 'self-efficacy outcomes', 'health behaviour outcomes' and 'clinical outcomes', we tested the variability of effect size estimates across studies within separate analyses (See Table 4.3). We also split the behavioural outcomes into two different groups based on study design – those that followed individuals longitudinally and those that implemented a different study design, such as repeated cross-sectional measures. This stratification reflects later sensitivity analyses around potential publication bias, but also often reflects a distinction in dose response and the number of individuals from which information is collected (see below, where we explore heterogeneity on the basis of age); there were an insufficient number of studies to conduct the same stratification across other studies.

**Table 4.3:** Heterogeneity across studies within separate analyses

| Group | Outcome type | Measure-ment point | Change data | Number of studies | Pooled effect estimate | Confidence Interval of effect | $\tau^2$ | Q-statistic | $I^2$ (%) |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Behavioural [S] | Post-test | No | 49 | 0.274*** | 0.198-0.336 | 0.045 | 297.63 | 83.88 |
| 1a | Behavioural: Longitudinal Measures of Individuals [S] | Post-test | No | 34 | 0.303*** | 0.207-0.382 | 0.049 | 162.57 | 79.17 |
| 1b | Behavioural: Cross-sectional Measures of Individuals [S] | Post-test | No | 15 | 0.210*** | 0.102-0.317 | 0.034 | 103.32 | 86.45 |
| 2 | Behavioural [S] | Post-test | Yes | 5 | 0.005 | -0.064-0.075 | [SS] | 2.68 | 0.00 |
| 3 | Behavioural [S] | Follow-up | No | 6 | 0.539** | 0.162-0.916 | [SS] | 82.63 | 93.95 |
| 4 | Clinical [S] | Post-test | No | 19 | 0.110** | 0.046-0.174 | 0.009 | 42.78 | 57.93 |
| 5 | Clinical [S] | Post-test | Yes | 7 | 0.285* | 0.023-0.548 | 0.069 | 14.58 | 58.84 |
| 6 | Clinical | Follow-up | No | 1 | 0.166 | -0.218-0.550 | [SS] | - | # |
| 7 | Self-efficacy [S] | Post-test | No | 10 | 0.504** | 0.175-0.833 | 0.251 | 182.13 | 95.06 |
| 8 | Self-efficacy [S] | Post-test | Yes | 4 | 0.768 | -0.640-2.175 | [SS] | 458.96 | 99.35 |
| 9 | Self-efficacy [S] | Follow-up | No | 2 | 0.784* | 0.151-1.417 | [SS] | 6.52 | 84.67 |

[S] Random effects model estimated; [SS] $\tau^2$ is not presented as estimates are based on a small number of studies (for further information see (Borenstein et al., 2011); ****p≤0.001; **p≤0.01; *p≤0.05.* # No $I^2$ is presented as there is no between-study heterogeneity for one study

With the exception of a group of five studies examining post-test behaviour change outcomes (group 2 in Table 4.3), within-group heterogeneity statistics indicate that the

groups are statistically significantly different from each other, although in this latter group, the evidence suggested that the pooled effect size did not differ from zero (see the forest plot in Figure 4.10). Where $I^2$ exceeds 50%, it is inadvisable to combine effect size estimates in an overall summary effect size (Borenstein et al. 2011). Where variation in the magnitude of effects is large, we do not focus on the average effect size statistic, but upon explaining the variability between the observed effect size estimates – this can be done through sub-group analyses or through meta-regression. All analyses of heterogeneity are based on groups identified a-priori in our research questions.

## 4.8 Addressing the research questions

### *4.8.1 Research Question 1: Are potentially modifiable processes of community engagement associated with health outcome effects?*

To address research question 1, we considered two sets of information that measured processes of community engagement. As set out in the methods section of this chapter we:

1. initially examined the extent to which modifiable processes of the intervention explain between-trial heterogeneity of effect size in univariate random effects models prior to
2. modelling significant covariates simultaneously in multivariate random effects regression models to examine the way in which study-level covariates (i.e. processes of community engagement) can help to explain heterogeneity, also taking into consideration the impact of other study-level covariates where possible in later analyses (e.g. socio-demographics).

We did this by running analyses separately for those permutations where we had a sufficient number of studies: these were post-test measures for behavioural, clinical and self-efficacy outcomes (that did not represent change measures). In addition, we examined three indices and different concepts around community engagement:

1. community engagement processes identified in Brunton et al. (2015; Review 1) that we examined as individual processes and as a composite score
2. community engagement processes originally identified in O'Mara et al. (2013) that we examined as individual processes and as a composite score
3. a composite score around the extent of community engagement

#### ***4.8.1.1 Community engagement processes, including communication, training support and collective decision making***

We began by examining new processes identified within the review update - whether studies reported (sufficient levels of):

a. collective decision making
b. bidirectional communication
c. training of coalition members
d. administrative support for coalition members
e. time for relationships to develop

These processes were only available for new studies identified since the original O'Mara and colleagues (2013) study was undertaken, i.e. data extraction for these specific processes was not undertaken for the studies appearing in the previous review, but only for those in the current update.

**Table 4.4:** Number of studies by number of community engagement processes identified (CERUB only)

| | | Collective decision making | Bidirectional communication | Training of coalition members | Supported with administration | Adequate time for relationship |
|---|---|---|---|---|---|---|
| Behavioural outcomes* | Studies not reporting | 7 | 7 | 4 | 15 | 14 |
| | Studies reporting | 9 | 9 | 12 | 1 | 2 |
| Clinical outcomes | Studies not reporting | 0 | 1 | 1 | 7 | 6 |
| | Studies reporting | 7 | 6 | 6 | 0 | 1 |
| Self-efficacy outcomes | Studies not reporting | 2 | 2 | 2 | 4 | 3 |
| | Studies reporting | 2 | 2 | 2 | 0 | 1 |

*Because of the small number of studies, we do not disaggregate behavioural outcomes here by study design, although as earlier, all outcomes are post-test measures

As there were too few studies to look at clinical or self-efficacy outcomes (See Table 4.4), we focused our efforts on behavioural outcomes, and in particular, collective decision making, bidirectional decision making and training of coalition members, where there were sufficient distributions in our categories. When we tested these processes among the updated (CERUB) studies, we were not able to identify any individual processes that were significantly associated with effect size.

We also created a score from these processes to examine the impact of each additional community engagement process on effect size. Potentially, studies could employ up to five processes, although the highest recorded in our sample was four, which was found in one study only (Rhodes et al., 2011). We modelled this continuously as well as treating it as a categorical variable (to help identify non-linear effects). Neither model suggested that additional processes of engagement made an impact on effect size for behavioural outcomes. We present the pooled effect size for these studies in Table 4.5. These analyses were conducted solely for the new studies identified in our CERUB review update - they do

not reflect pooled estimates of studies also identified in the original CERI review by O'Mara-Eves et al. (2013). It should be acknowledged therefore that while most of the effect sizes for these behavioural outcomes suggest that they differ from zero, and there is comparatively little within-group heterogeneity ($I^2$ under 55% within each group), the low number of studies available reduces the reliability of these pooled effect sizes.

**Table 4.5:** Pooled effect size by number of community engagement processes identified (CERUB only)

| **Community Engagement Processes** | **Number of studies[$]** | **Pooled effect estimate (Cohen's *d*)** | **Confidence interval of effect estimate** |
|---|---|---|---|
| None | 2 | 0.361*** | 0.168-0.553 |
| One | 3 | 0.578*** | 0.320-0.836 |
| Two | 4 | 0.119 | -0.064-0.301 |
| Three/Four | 7 | 0.204* | 0.000-0.409 |

[$]Because of the small number of studies, we do not disaggregate behavioural outcomes here by study design.

***p≤0.001; *p≤0.05.
Note: The results of a random-effects model are shown. Heterogeneity statistics for meta-analysis: Q = 54.21, p < 0.001; $I^2$ =72.3%

#### *4.8.1.2 Community engagement processes, including lay delivery, how health needs were identified and involvement in design*

To explore processes of community engagement further, we looked at the processes identified in the original O'Mara-Eves et al. (2013) review (referred to as CERI processes here); these included:

a. whether the health need was identified by the community
b. whether the community collaborated in the design
c. whether the community was consulted on the design
d. whether the intervention was delivered by lay members of the community.

As with the analyses above, we combined these into a score reflecting the number of domains reported, the distribution of which is shown by domain type in Table 4.6. As was the case for the previous score, studies could potentially record up to three processes, which 11 individual studies did (one from CERUB and ten from CERI). There was little variation for self-efficacy outcomes – all studies reported two processes, and therefore variation in self-efficacy outcomes could not be explored further. In modelling the number of CERI-identified processes, we employed a continuous variable, so that the *B* coefficient reflected the impact of each additional community engagement process on the effect size.

**Table 4.6:** Number of studies by number of community engagement processes identified (CERUB and CERI)

| Number of processes identified | Behavioural | | | Clinical | Self-efficacy |
|---|---|---|---|---|---|
| | Overall | Longitudinal measures of individuals | Other study design | | |
| None | 1 | 1 | 0 | 0 | 0 |
| One | 13 | 12 | 1 | 5 | 0 |
| Two | 27 | 17 | 10 | 9 | 10 |
| Three | 8 | 4 | 4 | 5 | 0 |

Our results suggested that for each additional process of community engagement reported in a study, there was an increase in effect size for behavioural outcomes (see Table 4.7). Statistical significance for this effect was restricted to those studies that measured behavioural outcomes using a longitudinal design to track the outcomes of individual beneficiaries (as opposed to repeated cross-sectional measurements or other study designs). Each additional process of community engagement reported corresponded to a 0.15 increase in average effect size. While this can only be considered to be a moderate impact, the results suggest that studies with higher numbers of identified engagement processes may also have the largest impact in terms of effect size. Modelling each additional process of community engagement helped to explain 21.2%[3] of between-study variance for these behavioural outcomes, although the results also showed that a substantial degree of heterogeneity remained.

[3] Based on adjusted R-squared

**Table 4.7:** Output for regression modelling the impact of number of community engagement processes identified† on effect size (CERUB and CERI)

| | **All behavioural outcomes** | **Behavioural outcomes: longitudinal measures of individuals** | **Behavioural outcomes: cross-sectional measures of individuals** | **Clinical outcomes** |
|---|---|---|---|---|
| Additional community engagement process | 0.0929 [-0.0179,0.204] | 0.157* [0.0370,0.278] | 0.0380 [-0.272,0.348] | -0.0114 [-0.0943,0.0715] |
| Constant | 0.115 [-0.0982,0.329] | 0.0535 [-0.158,0.264] | 0.153 [-0.536,0.842] | 0.132 [-0.0493,0.313] |
| N | 49 | 34 | 15 | 19 |
| Q | 297.3 | 144.3 | 99.74 | 42.57 |
| $\tau^2$ | 0.0624 | 0.0430 | 0.0942 | 0.00766 |
| $I^2$ | 0.842 | 0.778 | 0.870 | 0.601 |

95% confidence intervals in brackets

* $p \leq 0.05$

†Based on CERI framework

In Table 4.8, we also present the pooled effect size for the number of community engagement processes for behavioural outcomes from longitudinal studies.

**Table 4.8:** Pooled effect size for behavioural outcomes measured longitudinally, by number of community engagement processes identified (CERI and CERUB studies; CERI framework)

| **Community engagement processes** | **Number of studies** | **Pooled effect estimate (Cohen's *d*)** | **Confidence interval of effect estimate** |
|---|---|---|---|
| None | 1 | 0.066 | -0.121-0.254 |
| One | 12 | 0.170** | 0.069-0.271 |
| Two | 18 | 0.339*** | 0.225-0.453 |
| Three | 4 | 0.539* | 0.011-1.067 |

***p≤0.001; **p≤0.01; *p≤0.05.

Note: The results of a random-effects model are shown. Heterogeneity statistics for the meta-analysis: Q = 162.57, p < 0.001; $I^2$ =79.7%

We also attempted to unpack community engagement processes (based on the CERI framework) through examining the impact of individual processes by looking at these individually in univariate analyses (see Table 4.8). These analyses were intended to reveal whether one particular community engagement process was driving the results above for behavioural outcomes (in studies that collected longitudinal measurements). However, these analyses did not reveal one particular community engagement process that was driving a trend towards a stronger effect size. We did observe that among this set of studies, those that reported that the community either collaborated or were consulted in the design also had larger effect sizes for behavioural outcomes measured longitudinally, although this was not significant at the 5 per cent level (*B*=0.171, p=0.11).

Therefore, while these analyses suggest that additional community engagement processes may be associated with statistically significantly larger effect sizes, no one individual process can be attributed with this effect in the meta-analyses. This issue is unpacked further in later qualitative analyses.

### *4.8.2 Research Question 2: What is the relationship between the extent of community engagement (high, moderate or low) and health outcome effects?*

While the analyses above focused on specific processes of community engagement, we were also interested in exploring the overall rating of the extent of engagement given to each study. This rating – high, medium or low – captures a greater diversity of processes and is based on whether coalitions were either leading or collaborating in the design, delivery or evaluation of interventions. This gives a different distribution of scores compared with the earlier composite scores that we examined.

'Extent' of community engagement was determined as follows:

For three aspects of the intervention (design, delivery and evaluation), the level of community engagement was rated as follows:

- Leading or collaborating = 1
- Consulted, informed or not involved = 0

To determine the extent of community engagement, the level of engagement across all three aspects of an intervention was summed and the extent determined as follows:

- High = 3
- Moderate = 2
- Low = 1 or 0

**Table 4.9:** Number of studies by number of community engagement processes identified and extent of engagement scores

| **Community engagement processes** | **CERUB studies** | **Extent of engagement[$]** | **CERUB and CERI studies** |
|---|---|---|---|
| *CERUB Processes[†]* | | | |
| None | 2 | Low | 20 |
| One | 3 | Medium | 28 |
| Two | 4 | High | 10 |
| Three or more [φ] | 7 | | |
| *CERI Processes[‡]* | | | |
| None | | | 1 |
| One | | | 15 |
| Two | | | 31 |
| Three | | | 11 |

†CERUB processes: collective decision making; bidirectional communication; training support; administrative support; adequate time for relationship development.
‡CERI processes: health need identified by community; consultation or collaboration in design of intervention; lay delivery of intervention.
φ Maximum number of community engagement processes employed in any one study = 4.
$ Based on design, delivery and evaluation of intervention

As with the analyses described above, we modelled the extent of engagement in random effects regression models, although we included two dummy variables representing 'high' and 'medium' extent of engagement and examined the association with effect size relative to low extent of engagement. We did this in five models representing overall behavioural outcomes as well as our two sub-sets based on study design (longitudinal or repeated cross-sectional follow-up), as well as for clinical and self-efficacy outcomes. The results are presented in Table 4.10, and demonstrate that a higher extent of engagement was associated with higher effect sizes for behavioural outcomes for studies with longitudinal follow-up of individuals.

The size of the coefficient ($B$=0.673; $p$<0.01) suggests that a high extent of engagement is associated with substantially higher average effect sizes for these behavioural outcomes (using low extent of engagement as a reference category); however, it is worth noting that this high engagement group is composed of only four studies for this outcome domain

type.[4] Nevertheless, modelling the impact of high, medium and low extents of engagement helped to explain 51.8%[5] of between-study variance for these behavioural outcomes. As with our previous analyses, these add to the evidence that studies with higher levels of community engagement and community engagement processes are also more likely to have, on average, higher effect sizes, although we were unable to identify specific processes that contributed to this result. It should be noted that, with the exception of models for self-efficacy, the $\tau^2$ reflecting between-study variance is relatively modest in all cases (see Table 4.3). We have also pictorially depicted the difference in effect size for behavioural outcomes measured longitudinally in the forest plot in Figure 4.11.

[4] When we explored the results further, we also found that the coefficient for the studies deemed to have a high extent of engagement suggested significantly higher effect sizes compared to those studies with a medium extent of engagement. Again the caveat to this result is the small number of studies with high extent of engagement. Combining studies with high and medium extents of engagement and comparing effect sizes with studies with a low extent of engagement suggested that a difference persisted in effect size which was significant at the 10% level (p=0.09), although this combined group had a higher $I^2$ value than among the groups of studies that were disaggregated by three categories of extent of engagement.

[5] Based on adjusted R-squared

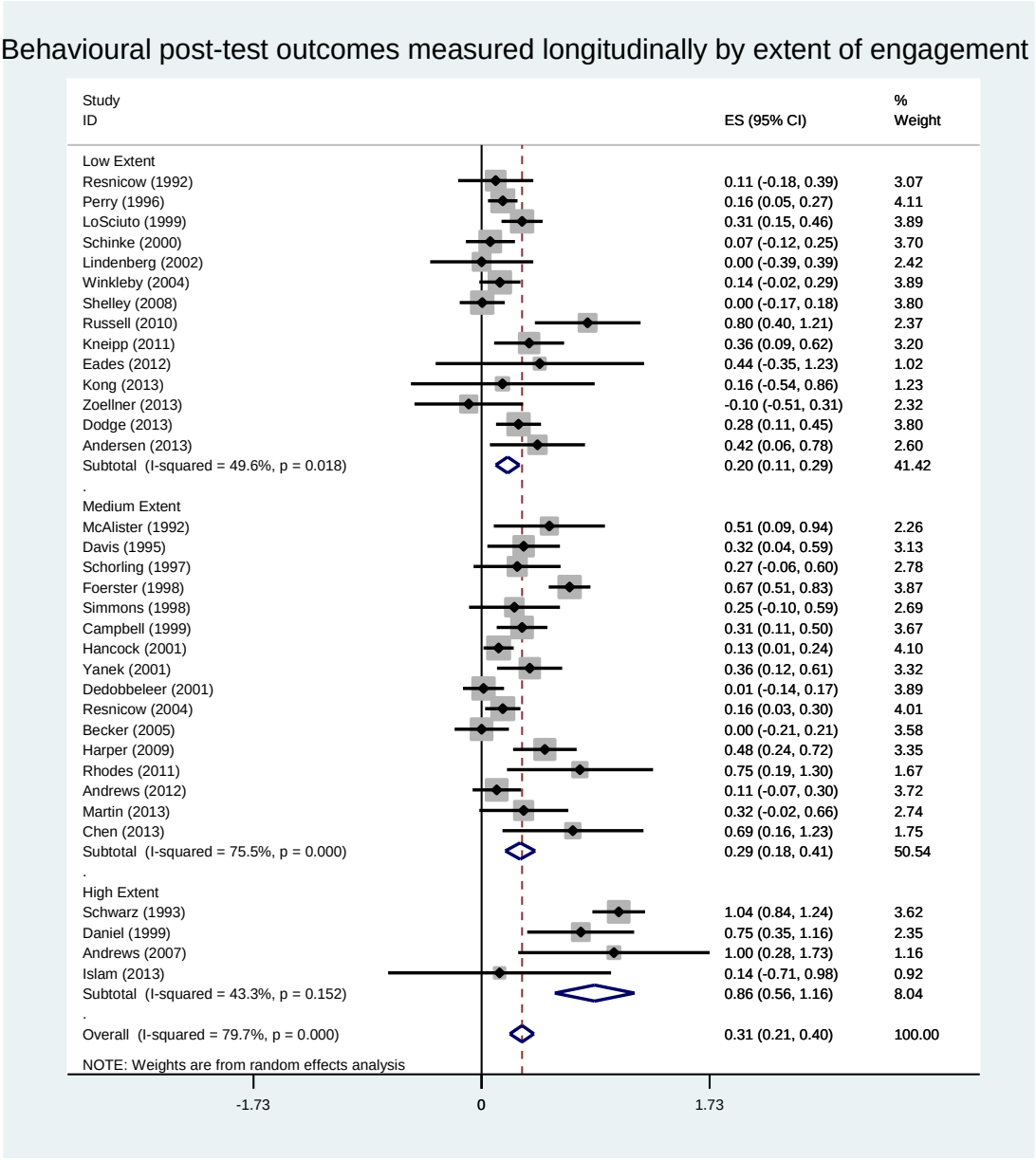
**Table 4.10:** Output for regression modelling the impact of number of community engagement processes identified† on effect size (CERUB and CERI)

| | **All behavioural outcomes** | **Behavioural outcomes: longitudinal measures of individuals** | **Behavioural outcomes: cross-sectional measures of individuals** | **Clinical outcomes** | **Self-efficacy outcomes** |
|---|---|---|---|---|---|
| *Extent of engagement (baseline: Low extent)* | | | | | |
| Medium | 0.151<br>[-0.0306,0.333] | 0.0845<br>[-0.0763,0.245] | 0.309<br>[-0.107,0.726] | 0.0410<br>[-0.104,0.186] | 0.535<br>[-0.478,1.549] |
| High | 0.217<br>[-0.0399,0.474] | 0.673***<br>[0.365,0.982] | 0.0560<br>[-0.411,0.523] | -0.0343<br>[-0.353,0.285] | 0.113<br>[-1.193,1.419] |
| Constant | 0.175*<br>[0.0375,0.313] | 0.202**<br>[0.0830,0.322] | 0.0733<br>[-0.256,0.403] | 0.0865<br>[-0.0284,0.201] | 0.259<br>[-0.463,0.982] |
| N | 49 | 34 | 15 | 19 | 10 |
| Q | 291.8 | 92.22 | 101.3 | 42.40 | 155.5 |
| $\tau^2$ | 0.0626 | 0.0263 | 0.0794 | 0.00806 | 0.346 |
| $I^2$ | 0.842 | 0.664 | 0.882 | 0.623 | 0.955 |

95% confidence intervals in brackets; * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

† Based on CERI framework

**Figure 4.11:** Behavioural post-test outcomes measured longitudinally by extent of engagement




*Research Question 3: Do direct comparisons of community engagement differ in health outcome effects from indirect comparisons of community engagement?*

In order to examine whether there was a difference between those studies that explicitly tested community engagement and those that did not, we required studies for which the only difference between each group or arm in the study was the presence or absence of community engagement, i.e. treatment and control groups received the same intervention but with the intervention receiving an additional community engagement element.

Only 2 of 20 studies included in the update meta-analysis provided direct comparisons of community engagement (Hoelscher et al. 2010; Segal et al. 2011). Of the 13 studies examining direct comparisons of community engagement in the

original review (O'Mara-Eves et al. 2013), only one involved community engagement with coalitions, collaborations or partnerships (Elder et al. 1993), and it was not possible to extract an effect size estimate from this study. Therefore, insufficient studies were available in order to address this question.

### *4.8.3 Research Questions 4a, 4b and 4c: Do community engagement interventions produce different health outcome effects for different age groups, genders or low-income groups?*

In our next set of analyses, we explored the extent to which the average effect size differed by a number of study-level characteristics. These included studies with a different health focus/topic, studies that focused on particular age or gender groups, and studies that focused on low-income groups.

#### ***4.8.3.1 Age***

We first looked at studies that explored different age groups, and found that around a third (19) of our 58 included studies focused exclusively upon children or young people (aged 18 or under). Other studies may have included children and young people, although in combination with other age groups. In general, studies that had an exclusive focus on children and young people did not have higher average effect sizes than other studies. One exception was studies with a focus on children that also collected information cross-sectionally on behavioural outcomes; these had higher average effect sizes than those that did not share this focus. This result was statistically significant, and was associated with a small-to-medium positive association with effect size ($B$=0.377; $p$<0.05; N=15), which also explained approximately 30% of between-study variance.[6]

Looking more closely within this group (studies with behavioural outcomes not measured longitudinally), and comparing the 10 studies that did focus on children and young people in this group with the 5 studies that did not, many of those that focused on children and young people tended to be in schools (four out of five studies). In this case, schools represented something of a closed population and were subject to less population churn than other settings, and even among those studies that used a cross-sectional design, the nature of the population and the intervention meant that most of the subjects assessed at post-test would have received the full intervention.

In contrast, other studies in this group were predicated on a variety of study designs, and for some, a limitation of their methods was the inability to ensure that those assessed received the same 'dose' of the intervention. For example, in Phillips and colleagues' (2014), a study that took place among 20 neighbourhoods in London, 40% of respondents assessed at post-test had not received a 'full dose' of the intervention, having moved into the study areas after the intervention had started. While the result described could suggest that studies that focus on children and young people are associated with higher effect sizes, we are unable to

[6] Based on adjusted R-squared

discount the possibility that this may be due to a 'dose-response' effect rather than an age effect per se.

#### *4.8.3.2 Gender*

We initially intended to explore whether studies which were targeted on males or females or which did not specify a focus, made an impact on average effect sizes. We found that among our 58 included studies, just two had a focus on males, while 13 focused on females and 43 did not specify a focus on either gender; we therefore grouped the two studies focused on males (both of which had behavioural outcomes only) with those without a focus on either gender. When we explored the impact of focusing on females only in explaining average effect size, we found that this variable explained very little of the between-study variance, with the exception of studies that measured self-efficacy outcomes. Here, there was evidence that studies that focused on females were more likely to report higher average effect sizes than those that did not have a gender-based focus (*B*=0.822; *p*=0.02; Table 4.11). Gender explained a large amount of between-study variance (46.9%), which in itself was substantially higher for self-efficacy outcomes than other domain types (see the $\tau^2$ values in Table 4.11). While these results could indicate that community engagement interventions are particularly effective in raising the self-efficacy of females; other factors relating to study quality may also be driving these results (see later sensitivity analyses).

**Table 4.11:** Output for regression modelling the impact of gender focus of intervention on effect size (CERUB and CERI)

| | **All behavioural outcomes** | **Behavioural outcomes: longitudinal measures of individuals** | **Behavioural outcomes: cross-sectional measures of individuals** | **Clinical outcomes** | **Self-efficacy outcomes** |
|---|---|---|---|---|---|
| Female focus (vs no gender focus or male) | 0.0317<br>[-0.179,0.242] | 0.0800<br>[-0.155,0.315] | -0.137<br>[-0.657,0.383] | 0.0649<br>[-0.123,0.252] | 0.822*<br>[0.146,1.497] |
| Constant | 0.275***<br>[0.180,0.371] | 0.289***<br>[0.176,0.401] | 0.254*<br>[0.0541,0.455] | 0.0990**<br>[0.0290,0.169] | 0.0142<br>[-0.508,0.536] |
| N | 49 | 34 | 15 | 19 | 10 |
| Q | 297.0 | 157.6 | 104.9 | 41.71 | 68.41 |
| $\tau^2$ | 0.0662 | 0.0555 | 0.0927 | 0.00699 | 0.177 |
| $I^2$ | 0.842 | 0.797 | 0.876 | 0.592 | 0.883 |

95% confidence intervals in brackets

* p ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001

#### *4.8.3.3 Lower socio-economic status*

We examined whether studies reported a focus on people from lower socio-economic status groups and their consequent effect sizes in random effects models. Lower socio-economic status was author-defined. Study authors reported this population characteristic in a variety of different ways, for example 'deprived' (Kloek et al., 2006), 'low-income' (O'Loughlin et al., 1999) and simply 'low socioeconomic status' (Dzewaltowski et al., 2010). When we examined the extent to which a focus on low socio-economic status was associated with average effect size, we found no association for self-efficacy or behavioural outcomes. We did find evidence that studies which included a focus on low-income groups and measured clinical outcomes had lower average effect sizes (*B*=-0.161; *p*=0.002; N=19), although we are sceptical of the validity of this finding given the low $\tau^2$ values for clinical outcomes (in pooled estimates, and on the basis of later sensitivity analyses (see the following section). In addition, care should be taken in the interpretation of the result, which does not suggest that studies focusing on lower socio-economic status groups had lower effect sizes because of this focus necessarily – other factors, including the study quality, may also be instrumental in driving this trend.

#### *4.8.3.4 Health topic*

From a long list of 13(+) categories (see Brunton et al. 2015), we grouped the health topic that was the focus of the intervention into four categories: 'healthy eating/physical activity' (HEPA, 32 studies), 'alcohol, tobacco, or drugs' (ATOD, 13 studies), service use and mental health (5 studies) and 'injury and infection screening and prevention' (IISP, 8 studies). Across our outcome domains, we examined the way in which the health topic was associated with effect size, finding that those studies that aimed to reduce the risk of injury or infection had larger effect sizes for behavioural outcomes relative to studies that focused on healthy eating or physical activity (Table 4.12). In the case of our model for behavioural outcomes for studies that followed individuals longitudinally, the average effect size for studies that focused on injury and infection screening and prevention was 0.479 higher relative to studies that focused on healthy eating or physical activity. Five studies included in this group examined issues such as Hepatitis-B screening, HIV, STI or pregnancy prevention (2 studies), mammography screening or having a smoke detector in the home. In this model, the health topic that was the focus of the intervention accounted for 57.8% of between-study variance. We do not present the results for studies reporting behavioural outcomes that did not collect longitudinal information, as there was just one study in each category besides the baseline (HEPA).

**Table 4.12:** Output for regression modelling the impact of health focus of intervention on effect size (CERUB and CERI)

| | **All behavioural outcomes** | **Behavioural Outcomes: longitudinal measures of individuals** |
|---|---|---|
| *Health topic: baseline HEPA* | | |
| ATOD | -0.0751<br>[-0.244,0.0936] | -0.126<br>[-0.293,0.0409] |
| Service use | 0.0587<br>[-0.202,0.319] | -0.0481<br>[-0.296,0.200] |
| IISP | 0.513***<br>[0.256,0.771] | 0.479***<br>[0.235,0.724] |
| Constant | 0.241***<br>[0.144,0.338] | 0.291***<br>[0.172,0.409] |
| N | 49 | 34 |
| Q | 203.8 | 79.45 |
| $\tau^2$ | 0.0426 | 0.0230 |
| $I^2$ | 0.779 | 0.622 |

95% confidence intervals in brackets

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$. HEPA: healthy eating/physical activity; ATOD: alcohol, tobacco, or drugs; IISP: injury and infection screening and prevention

We replicated the model above and found a similar association for self-efficacy outcomes, although we do not present the results, as just two studies in the model had a focus on injury and infection screening and prevention. Therefore, due to sample size, we collapsed some of the categories below and for both clinical and self-efficacy models, we ran models that included a coefficient for a HEPA focus (vs other health focus). In the model for clinical outcomes, we found that a HEPA focus was not associated with effect size. However, for the self-efficacy outcomes model, the result suggested that a HEPA focus was associated with a lower effect size (*B*=-0.804; *p*=0.02; N=10); however, as the comparison category was heterogeneous in terms of focus and was small (3 studies), we do not discuss this result further.

In exploring study-level characteristics therefore, our strongest evidence in explaining between-study variation comes from the studies' health focus. Here the evidence suggests that studies that use a longitudinal design and that focus on the reduction of risk from injury or infection are associated with higher effect sizes for behavioural outcomes. Compared to studies focusing on healthy eating or physical activity, these studies have higher average effect sizes for behavioural outcomes in the region of 0.479 (which would be considered a medium effect in itself). As with much of the analyses here, the modest number of studies in the analyses (34 in total, 5 with a focus on injury and infection screening and prevention) is a caveat.

Nevertheless, health focus accounted for almost 60% of between-study variance for behavioural outcomes measured longitudinally.

### *4.8.4 Research Question 4d: Do community engagement interventions produce different effects for 'distal' (e.g. self-efficacy) outcomes, 'intermediate' (e.g. health behaviour) outcomes, or 'proximal' (clinical/physiological) outcomes?*

#### ***4.8.4.1 The correlation between effect size in one domain and another***

We considered the extent to which outcomes across different domains varied within studies and the extent to which we could model causal pathways towards improved health outcomes. However, we were unable to test the direct and indirect pathways to improving health outcomes because of a lack of data for different outcome types from the same studies. Modelling this causal pathway was also limited because we did not have longitudinal data to test the proposed causal ordering. We ran correlation analyses to test whether there were any relations between outcome domains, although we found little evidence of any correlation, and none of the correlation analyses we undertook[7] came close to achieving statistically significant results; these are therefore not presented here.

#### ***4.8.4.2 The stability of effect size across time***

The majority of our data in this chapter derive from post-test measures of differences, and we were able to extract follow-up measures only from a limited number of studies. Here we have compared post-test and follow-up measures from five studies collecting information on health behaviour outcomes. These studies were Andersen et al. (2013), Andrews et al. (2007 and 2012), Dedobbeleer and Desjardins (2001), Harper et al. (2009) and Winkleby et al. (2004), and the pooled effect size is displayed below with measures of heterogeneity (see Table 4.13). For each measurement point, the value of $I^2$ is high, which suggests considerable heterogeneity, undermining the reliability of the pooled estimates (although the direction of effect is consistent across studies). Nevertheless, it is noteworthy that the follow-up pooled effect size is larger than the post-test effect size, which could suggest that the impacts of coalition-based health interventions actually amplify across behavioural outcomes over time rather than attenuate. However, a larger body of evidence (with lower levels of between-study variance) is needed before such a conclusion can be drawn with confidence. For three of the studies, the effect size was larger; in another, there was a small decrease in effect size, while in a fifth study similar effect sizes were recorded. In one study, the increasing effect size was attributed to the ongoing support that the participants in the intervention group received, including 'more support for being physically active, stronger beliefs that positive outcomes will follow participation in physical activity, and the perception that they had more control over being physically active when faced with barriers' (Andersen 2013, p116). In the second study that recorded higher effect size at follow-up, in this case higher numbers recording having sex with a condom, this was attributed to more girls being sexually active at

[7] Weighted by an average of the standard error for the pairwise correlation.

post-test, so that the content of the intervention could be put into practice (safe sex) (Harper et al. 2009). In the third study, the higher effect size is not discussed, although attrition may have had an impact on the results (Dedobbeleer and Desjardins 2001). Given that there is no consistent pattern or explanation for higher effect sizes at follow-up and for the way in which community engagement may help to bring about such an effect, and the findings in Table 4.13 are based on very small samples, it would be difficult for such a finding to be incorporated into further guidance without examining a larger pool of studies.

**Table 4.13:** Pooled effect size and measures of heterogeneity for behavioural outcomes among studies with both post-test and follow-up data (CERUB and CERI)

| | Outcome type | Measurement point | Number of studies included | Pooled effect estimate (Cohen's *d*) | Confidence Interval of effect estimate | $\tau^2$ | Q-statistic | $I^2$ (%) |
|---|---|---|---|---|---|---|---|---|
| 1 | Behav-ioural | Post-test | 5 | 0.296* | 0.066-0.527 | 0.0463 | 17.25 | 76.8 |
| 2 | | Follow-up | 5 | 0.661*** | 0.172-1.151 | 0.275 | 74.2 | 94.6 |

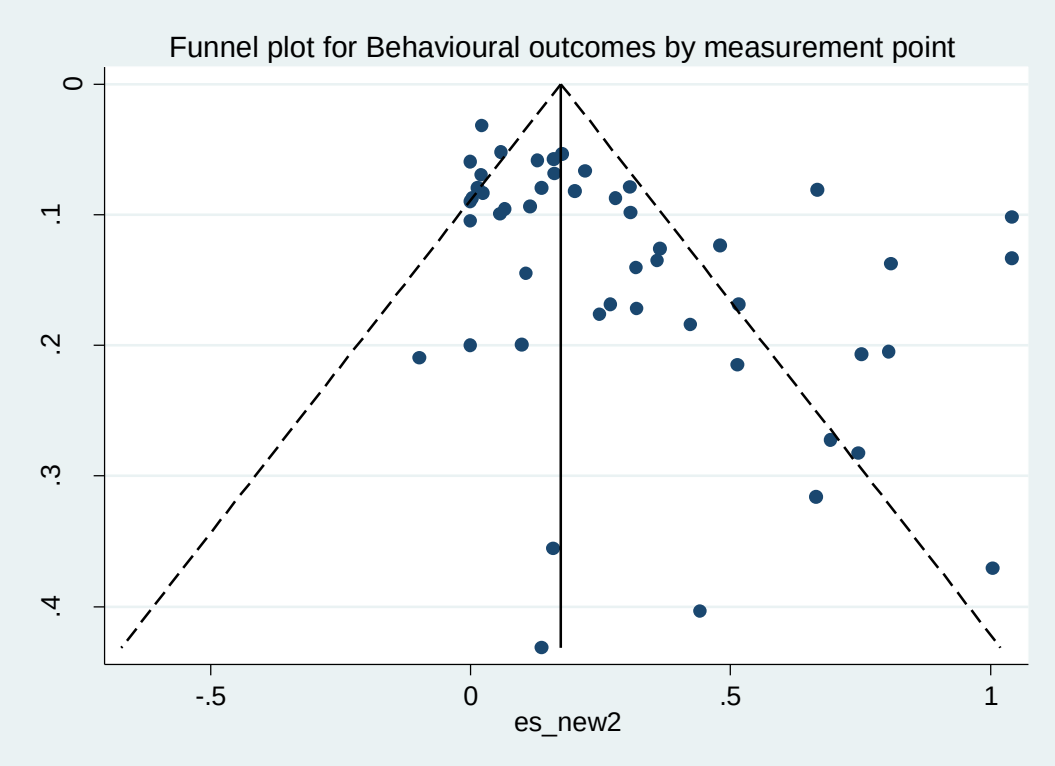
* $p \leq 0.05$, *** $p \leq 0.001$.

## 4.9 Sensitivity analyses

### *4.9.1 Risk of bias: consideration of potential publication bias*

We looked for potential publication bias across the nine comparisons, excluding those using funnel plots and Egger's test (Harbord et al., 2009). We looked across all three domains, and found little evidence for clinical outcomes or for self-efficacy outcomes (p>0.05 for small-study effects). However, we identified potential evidence of publication bias when looking at behavioural outcomes at post-test across the 49 studies; the estimated bias coefficient measured 2.370 (p <0.01). This is presented in the funnel plot in Figure 4.12.

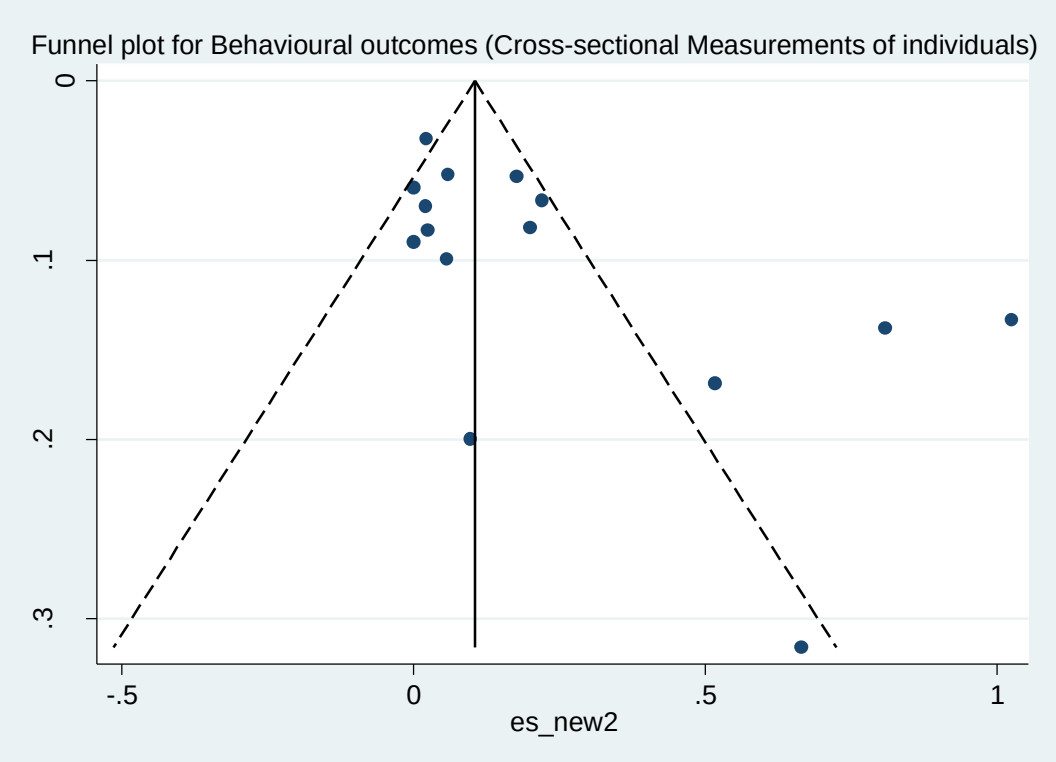
**Figure 4.12:** Funnel plot of effect size estimates and standard errors of studies reporting post-test health behaviour outcomes n = 49.




Part of this effect may be accounted for by the large number of studies that use large cross-sectional surveys to collect information on intervention impacts (which can collect data from a large sample and thereby reduce the standard errors of estimates). We have therefore stratified our analyses further by study type and examined those studies that follow individuals separately from studies that follow clusters or neighbourhoods or otherwise use cross-sectional methods (or a combination of methods) as illustrated in Figure 4.14. Thus we have found that evidence for publication bias is eliminated for the larger group of studies that followed individuals, which included 34 studies (the estimated bias coefficient measured 1.459 ($p>0.05$); for the smaller group of 15 studies that did not follow individuals longitudinally (see Figure 4.13), or that use a combination of methods, evidence of publication bias remained (the estimated bias coefficient measured 3.329 with a standard error of 1.220, giving a p value of 0.02). While there are strategies available that attempt to account for publication bias, each has its own limitations, and therefore these are not employed here. If this finding is due to publication bias, this evidence suggests that pooled effect sizes for behavioural outcomes measured cross-sectionally may be inflated and that studies that find null or negative effects are underrepresented in our sample of studies. However, in such a heterogeneous data set, an asymmetric funnel plot may be caused by numerous other factors, including small study sample size or multiple phenomena under study. A cautious interpretation of this is that this potential source of publication bias does therefore represent a caveat to our results (but this should not be overstated) (O'Mara-Eves et al. 2013).

**Figure 4.13:** Funnel plot of effect size estimates and standard errors of studies reporting post-test health behaviour outcomes measured longitudinally (n = 34)


**Figure 4.14:** Funnel plot of effect size estimates and standard errors of studies reporting post-test health behaviour outcomes measured cross-sectionally (n = 15)



### *4.9.2 Risk of bias: continuous and binary measures of effect size and study quality*

We conducted two further sets of sensitivity analyses to test whether (i) Cohen's *d* effect size estimates based on binary data were statistically similar to *d* effect size estimates based on continuous data; and (ii) to measure the between group variance of studies deemed as methodologically 'sound' compared to those deemed 'not sound'. Those studies deemed to be methodologically sound were those that did not have selection bias (of subjects included) or attrition bias, or reported their results selectively. We restricted these sensitivity analyses to post-test outcome measures. Looking first at studies with behavioural outcomes, we found no evidence of statistically significant differences in effect size by whether effect sizes were originally based on binary or continuous measures or whether studies were methodologically sound or not.

For the 19 studies with clinical outcomes, we found that there was significant between group heterogeneity on the basis of methodological quality. Those studies deemed methodologically sound tended to have higher effect sizes than those that were not sound. This potentially compromises the validity of our earlier findings exploring heterogeneity (although $I^2$ for clinical outcomes was relatively low and consequently we found little, besides socio-economic status, that was associated with effect size). In our final (consolidated) models presented in the next section, we have therefore applied a similar stratification as for behavioural studies above, and have focused on the ten studies that collected information longitudinally from study subjects. Among these studies, the pooled effect size stood at 0.192, which was statistically significantly different from zero, with a confidence interval of 0.092-0.292. Among these 10 studies with clinical outcomes, there was little between-study variance and $I^2$ stood at zero; therefore we have presented a constant-only model in our final set of results.

Among the 10 studies with self-efficacy outcomes, we also found that there was significant between group heterogeneity on the basis of methodological quality ($p<0.05$). Studies that were deemed sound had higher effect sizes. The number of studies reporting these outcomes was too small to apply a further stratification. Therefore, for our results for self-efficacy we have presented the results, although with the caveats that methodological quality and original measurement scales may present systematic bias in the effect size estimates.

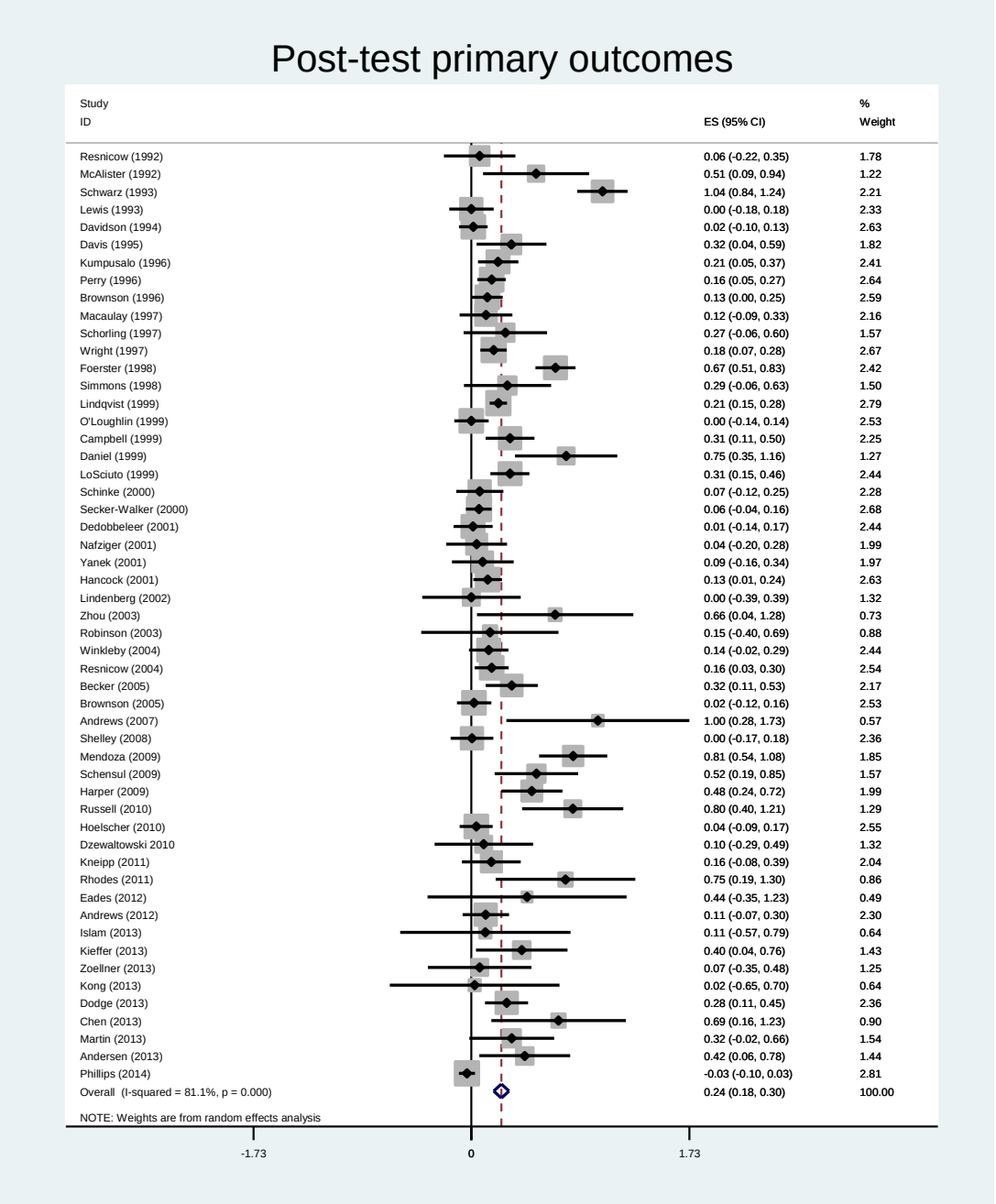
## 4.10 Final analyses: primary effect sizes and consolidating the evidence

### *4.10.1 Primary effect size*

In our final analyses, we examined the results for the primary outcome in studies. As studies often failed to identify a single primary outcome, we ranked the outcomes by importance, starting with clinical, followed by behavioural, followed by self-efficacy outcomes. This was under the assumption that most studies would aim to ultimately make a difference to clinical outcomes in their theory of change. In total, we had information for 53 studies (other studies only had change data or data collected at follow-up); of these, the primary outcomes for 19 studies were clinical and for 34 studies, behavioural. The overall effect size for these analyses stood at a modest 0.237 (CI:0.176-0.297), although this did significantly differ from

zero, and there was substantial heterogeneity ($I^2$=81%). These results are shown in the forest plot in Figure 4.15.

**Figure 4.15:** Post-test 'primary' outcomes




We have presented results that explore this variance further in Table 4.14. As with earlier analyses around individual domains, we first looked at the extent to which individual factors helped to explain between-study variance. Few factors appeared to explain effect size differences statistically significantly, with a regression model for extent of engagement only achieving borderline statistical significance and, similarly, a model with health topic only suggesting that this factor significantly explained differences in effect size. These study-level factors were brought together into a multivariate regression model, and the results mirrored those

observed earlier for our behavioural outcomes (measured longitudinally), where studies that focused on screening and prevention of injury or infection (IISP) and, to a lesser degree, studies that involved community members to a high extent were those that tended to have higher effect sizes. In both cases, the impact on effect size was modest – relative to HEPA focused studies, a focus on IISP was associated with an average effect size that was 0.304 ($p<0.01$) higher, while a high extent of community engagement in designing, delivering and evaluating the intervention was associated with an average effect size that was 0.191 higher. However, a number of factors should be highlighted. Firstly, the $I^2$ was high in the null model for primary effect size (without covariates), standing at 81% with a $\tau^2$ value of 0.0464; in model three (in Table 4.14) which includes covariates for both extent of engagement and health topic, the value of $\tau^2$ reduced only by 11.1%. This indicates that the vast majority of the between-study variance for primary outcomes remained unexplained, and model fit statistics actually suggested that model 2 (Table 4.14) provided the best fit (adjusted $R^2$=11.4%).

**Table 4.14:** Output for regression modelling the impact of covariates† on primary effect size (CERUB and CERI)

| | **Model 1** | **Model 2** | **Model 3** |
|---|---|---|---|
| *Extent of engagement (baseline: Low extent)* | | | |
| Medium | 0.0959<br>[-0.0571,0.249] | | 0.0868<br>[-0.0655,0.239] |
| High | 0.212[BS]<br>[-0.0130,0.437] | | 0.191 [BS]<br>[-0.0329,0.415] |
| *Health topic: baseline HEPA* | | | |
| ATOD | | -0.0423<br>[-0.206,0.122] | -0.00247<br>[-0.173,0.168] |
| Service use | | 0.0429<br>[-0.210,0.295] | 0.0624<br>[-0.194,0.319] |
| IISP | | 0.301**<br>[0.0952,0.507] | 0.304**<br>[0.0971,0.510] |
| Constant | 0.167**<br>[0.0501,0.283] | 0.208***<br>[0.115,0.300] | 0.127<br>[-0.0154,0.269] |
| N | 53 | 53 | 53 |
| Q | 268.2 | 252.4 | 244.4 |
| $\tau^2$ | 0.0455 | 0.0410 | 0.0413 |
| $I^2$ | 0.814 | 0.806 | 0.808 |
| df_Q | 50 | 49 | 47 |

95% confidence intervals in brackets

** $p \leq 0.01$, *** $p \leq 0.001$, [BS] Borderline statistically significant ($p<0.10$)

† These are covariates that have been found to be significant in univariate random effects regression models.

### *4.10.2 Consolidating the evidence across outcome domains*

In Table 4.15, we present the evidence from those factors that were found to be significant in our univariate random effects models that we subsequently entered into multivariate random-effects meta-regression models for each domain type. For the model for self-efficacy post-test outcomes and the model for clinical post-test outcomes, these models remain the same as those presented earlier, given that we identified only one study-level factor that was associated with effect size (gender and socio-economic status respectively), and we also present a null model (no covariates) for a subset of clinical outcomes (see earlier sensitivity analyses).

For behavioural outcomes, we found that among those studies with a longitudinal design, both higher extent of community engagement (relative to lower engagement) and having a focus on reducing the risk of injury or infection (relative to a focus on healthy eating or physical activity) were associated with significantly higher effect sizes. These effects remained after we entered these covariates into the model simultaneously, so that health focus and the extent of engagement remained (independently) associated with effect size after controlling for one another. In the case of health focus, relative to studies that had a focus on healthy eating or physical activity, studies focusing on screening and prevention of injury or infection were associated with a 0.358 ($p<0.01$) increase in average effect size; studies that reported a high extent of engagement (relative to low) were associated with a 0.452 ($p<0.01$) increase in average effect size. Differences of this magnitude in effect size observed between groups would ordinarily be considered small-to-medium (Cohen 1992). This random effects model accounted for 69% of between-study variance, while the percentage of residual variance reduced to a moderate-to-substantial level (from substantial-to-considerable) (Higgins and Green 2011).

**Table 4.15:** Output for multivariate regression models measuring the impact of covariates[†] on effect size across different outcome domains (CERUB and CERI)

| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** |
|---|---|---|---|---|---|---|
| | **All behavioural outcomes** | **Behavioural outcomes: longitudinal measures of individuals** | **Behavioural outcomes: cross-sectional measures of individuals** | **Clinical outcomes** | **Clinical outcomes longitudinal measures of individuals** | **Self-efficacy outcomes** |
| *Health topic: baseline HEPA* | | | | | | |
| ATOD | -0.0409<br>[-0.218,0.136] | -0.105<br>[-0.271,0.0615] | | | | |
| Service use | 0.0849<br>[-0.181,0.351] | -0.0150<br>[-0.250,0.220] | | | | |
| IISP | 0.498***<br>[0.237,0.758] | 0.358**<br>[0.117,0.599] | | | | |
| *Extent of engagement (baseline: Low extent)* | | | | | | |
| Medium | 0.112<br>[-0.0516,0.276] | 0.0205<br>[-0.133,0.174] | | | | |
| High | 0.133<br>[-0.102,0.367] | 0.452**<br>[0.144,0.761] | | | | |

| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** |
|---|---|---|---|---|---|---|
| | **All behavioural outcomes** | **Behavioural outcomes: longitudinal measures of individuals** | **Behavioural outcomes: cross-sectional measures of individuals** | **Clinical outcomes** | **Clinical outcomes longitudinal measures of individuals** | **Self-efficacy outcomes** |
| Focus on children and young people | | | 0.380*<br>[0.0384,0.722] | | | |
| Focus on low-income groups | | | | -0.161**<br>[-0.265,-0.0579] | | |
| Focus on females | | | | | | 0.822*<br>[0.146,1.497] |
| Constant | 0.157*<br>[0.00766,0.307] | 0.249**<br>[0.0851,0.413] | 0.119<br>[-0.0610,0.300] | 0.153***<br>[0.0932,0.213] | 0.192**<br>[0.0764,0.307] | 0.0142<br>[-0.508,0.536] |
| N | 49 | 34 | 15 | 19 | 10 | 10 |
| Q | 203.1 | 61.73 | 93.33 | 16.73 | 5.169 | 68.41 |
| $\tau^2$ | 0.0440 | 0.0169 | 0.0599 | 0.00205 | 0 | 0.177 |
| $I^2$ | 0.788 | 0.546 | 0.861 | 0 | 0 | 0.883 |
| df_Q | 43 | 28 | 13 | 17 | 9 | 8 |

† These are covariates that have been found to be significant in univariate random effects regression models.

From model 2 in Table 4.15, for behavioural outcomes measured longitudinally, we present the predicted pooled effect size across the factors that are found to be statistically significantly associated with effect size in Table 4.16. These predicted values are based on a random effects model that accounts for the impact of health topic and extent of engagement simultaneously. The results show that the pooled predicted effect size (d) for studies focusing on IISP would be over 0.5 higher than for studies focusing on ATOD. Similarly, the predicted average effect size (d) for studies that have a high level of engagement is over 0.5 higher compared to studies with a low extent of engagement for behavioural outcomes.

**Table 4.16:** Pooled predicted effect size for behavioural outcomes measured longitudinally based on Model 2 (see Table 4.15)

| | **Pooled effect estimate (Cohen's *d*)** | **Confidence Interval of effect estimate** |
|---|---|---|
| *Extent of engagement* | | |
| Low | 0.207*** | 0.140-0.275 |
| Medium | 0.306*** | 0.221-0.391 |
| High | 0.767*** | 0.577-0.958 |
| *Health topic* | | |
| HEPA | 0.306*** | 0.225-0.387 |
| ATOD | 0.174*** | 0.105-0.243 |
| IISP | 0.700*** | 0.544-0.857 |
| Service use | 0.243** | 0.096-0.390 |

***p≤0.001; **p≤0.01; *p≤0.05.

Note: The results of a random-effects model are shown. Heterogeneity statistics for the meta-analysis: Q = 165.57, p < 0.001; $I^2$ =79.5%

# 5. Findings: qualitative comparative analysis

Qualitative comparative analysis (QCA) was originally developed by Charles Ragin in the field of political science, but its guiding principles address many of the issues faced by reviewers synthesising complex interventions (Thomas et al. 2014), and we therefore undertook an exploratory QCA in this review. The main purpose of a QCA is to uncover 'configurations' of conditions which are associated with effectiveness; this is of course similar to the aims of the meta-regression, and the two analyses complement one another, rather than presenting radically different conclusions. One of the main differences in QCA is its ability to cope with multiple causal pathways. For example, it is possible for a condition such as 'intervention deliverers were trained' to be associated with effectiveness in some situations without it necessarily being associated with a lack of effect in others; such multiple causal pathways are often difficult to model in a meta-regression. 'Variables' within a QCA are known as 'conditions', though in the text that follows, the term 'variable' is also retained. All analyses were conducted in Stata using the 'fuzzy' command.

A QCA follows an iterative process, which falls into six main phases, outlined below:

1. building the data table,
2. constructing a 'truth table',
3. checking the quality of the truth table,
4. resolving contradictory configurations,
5. Boolean minimisation,
6. interpretation (results).

### 5.1 Building the data table

In order to complement the meta-analysis, the same studies that appeared in the meta-analysis were analysed using QCA; and in order to have the largest possible dataset, the outcome used was behaviour, measured at post-test. This sampling strategy is somewhat at variance with the 'purposive' strategy described in the QCA literature, but is coherent with systematic review methodology (Thomas et al. 2014).

Since QCA is a set-based mode of analysis, the outcomes, originally expressed as standardised mean differences, were 'calibrated', to form fuzzy sets. The outcome set was defined as those studies with 'large effect sizes', where a large effect was defined as being above 0.5 standard deviations. This led to the following breakdown, giving the analysis reasonable numbers of studies both inside and outside the outcome set:

- full membership (scoring 1.0; N = 12) in the set of large effect sizes: SMD > 0.5
- partial membership (scoring 0.66; N = 3) in the above set: SMD > 0.4 < 0.5
- partial non-membership (scoring 0.33; N = 10) in the above set: SMD > 0.2 < 0.4
- full non-membership (scoring 0.0; N = 26) in the above set: SMD < 0.2

## 5.2 Constructing a truth table

A truth table is a rearrangement of studies by the conditions they contain; and the configurations of these conditions are looked at across studies. The truth table thus contains configurations of conditions as the main unit of analysis, rather than individual studies. There were a large number of possible configurations, covering aspects of engagement, outcome domains and study quality. Many variables (and configurations) were evenly distributed across the outcome set, meaning that they were not able to 'explain' different results. Analysis was conducted in a stepwise fashion, with variables added into the analysis and the emerging pattern of configurations examined. When a condition did not increase the ability of our model to explain differences between outcomes, it was discarded. For this reason, the modifiable processes initially selected for analysis (i.e. presence of collective decision making, adequate time for relationship development and cumulative processes of community engagement) were discarded in favour of other study characteristics that better explained the differences between outcomes.

### *5.2.1 Checking the quality of the truth table*

The analysis proceeded in a stepwise fashion, drawing on the results of the meta-analysis. (i.e. beginning with 'extent_low', the analysis added and removed variables in order to identify configurations of conditions which best explained the variation in effect size observed.) We found that the dataset was extremely heterogeneous, and the resulting truth table was not entirely satisfactory. It may be that additional data extraction of subject- and process-specific details across the studies might yield more insights, but this was not practicable in the time available.

### *5.2.2 Resolving contradictory configurations*

The truth table that resulted is free from direct contradictions, though it does not explain the entirety of the diversity in the dataset.

## 5.3 Boolean minimisation

After applying the Stata 'reduce' criterion, the solution was reduced to its minimum logical expression.

## 5.4 Results

The meta-analysis had found that the extent of engagement was positively related to effect size, with studies with more engagement having larger effects. We therefore began with the 'extent' variables, finding that in the QCA, this pattern was most apparent in the studies with low extent of engagement (i.e. that low engagement was associated with low membership in the set of highly effective studies, but that high engagement did not necessarily guarantee good results). Four studies were in the 'highly effective' set, but four were either entirely (N = 3) or partially (N = 1) outside it. No clear pattern explains this.

Additional variables were added and retained if they helped to explain patterns in the dataset. As with the statistical analysis, we did not stick rigidly to a $p<0.05$ threshold if some explanation seemed to be possible.

In summary, this qualitative comparative analysis suggested several combinations of conditions that tended to occur in more effective interventions. Firstly, coalitions that included lay delivery of interventions were found more often in interventions with larger effect sizes. This suggests that lay delivery used within a coalition may enhance an intervention's effectiveness.

Further, coalitions including lay delivery of interventions that were targeted to all ages of the (disadvantaged) population tended to have larger effect sizes, suggesting that lay delivery provided across all age groups (as opposed to those targeting specifically children, for example) may improve effect sizes.

When looking at interventions addressing infection or injury prevention, those coalitions that targeted the entire (disadvantaged) population and incorporated lay delivery were more aligned with effective interventions than other health topics. But additionally, coalitions that focused on sexual health or cancer prevention and also included lay delivery as part of the intervention strategy were aligned with more effective interventions than other topics (i.e. physical activity, healthy eating, mental health). This suggests that for these health topics in particular, a combination of coalitions and lay delivery may improve outcomes.

Conversely, coalitions employing a low extent of engagement across design, delivery and evaluation tended to have lower effect sizes.

# 6. Discussion

Our previous review examined and extended theories on the mechanisms of community engagement. We suggested that a number of community engagement mechanisms positively influenced outcomes, including: theories of change for patient/consumer involvement to influence service development; theories of change for peer-delivered interventions; and theories of empowerment for community development to reduce health inequalities (O'Mara-Eves et al. 2013). Our previous review also examined the effectiveness of process implementation in interventions that contained community engagement (O'Mara-Eves et al. 2013). The included studies provided weak evidence (due to the quality of intervention process evaluations) that successful partnerships and efforts to build relationships between partners appeared to influence programme outcomes. Further, good relationships between community members and professionals providing an intervention were important to programme implementation. Finally, good project management and specific, adequate ongoing training and support for community members impacted on implementation (O'Mara-Eves et al. 2013). The present work has aimed to extend that knowledge, to understand how to undertake good community engagement.

The findings from Review 1 of this review suggested that two characteristics might influence outcomes: the extent of community engagement throughout the lifespan of a project's design, delivery and evaluation; and the processes of that engagement.

To explore these characteristics further, we undertook three sequential syntheses:

- framework synthesis to understand which modifiable processes of community engagement were associated with higher and lower extent of engagement and differences between populations;
- meta-analytic modelling to understand which modifiable processes of engagement or other characteristics of studies were associated with larger and smaller effect sizes; and
- qualitative comparative analysis to understand which processes of engagement or other study characteristics were associated with effective (and non-effective) interventions.

## 6.1 Main findings

### *6.1.1 Framework synthesis*

The framework synthesis examined the processes of community engagement from data in the studies identified in Review 1 (n=26). This was done in order to understand which processes of community engagement led to successful outcomes. The main finding arising from this synthesis was that while processes of intervention implementation were evaluated in the published literature, processes of community engagement were only described and not rigorously evaluated by authors. This helps us to understand what processes have been used, but limits our ability to make recommendations concerning good practice for community engagement in health intervention design, delivery and evaluation.

Looking for modifiable processes of community engagement identified in our previous review's conceptual framework, we identified that when authors described how community members were engaged, the strategies most often described included collective decision making, bidirectional communication and training support. These processes have been described elsewhere as key components of collaborative research effort in community-based participatory research (CBPR) methods (Wallerstein et al. 2008; Wallerstein and Duran 2010). However, in studies where authors identified the application of CBPR principles, they did not always demonstrate these key modifiable processes. Collective decision making, bidirectional communication and training support were more consistently reported in studies that were rated as being of high or moderate extent of community engagement – in effect, community members led or collaborated on the design, delivery and evaluation of intervention studies. Tests of congruence of the concepts of CBPR and extent of engagement demonstrated that while there was some overlap, studies that self-identified as CBPR were not always rated as having high or moderate extent of community engagement, suggesting that these concepts are slightly different in nature.

Very little evidence was found of researchers allowing adequate time for collaborative relationships to develop between community members and researchers, service providers or other coalition members, despite this being described as an important part of relationship building (Jones et al. 2008; Wallerstein and Duran 2010).

Studies that were rated as having a high or moderate extent of engagement more often reported specific processes of community engagement, such as collective decision making, than did studies rated as having a low extent of engagement. Interventions targeting low-income groups more often reported specific processes, such as collective decision making, bidirectional communication, and training.

Other modifiable processes were identified *de novo*, including skills around conflict resolution, negotiation and reflection, and planning collaborative meetings to suit community members needs in terms of timing, location and provision of transport and childcare. These could be incorporated into future evaluations of community engagement, to show those wishing to implement such strategies some specific ways to undertake community engagement for health intervention design, delivery and evaluation. However to establish their validity, these and the other a priori processes of community engagement require more robust evaluation.

### *6.1.2 Meta-analysis/meta-regression*

Using studies included in our original review and identified in Review 1, we used meta-analysis and meta-regression techniques to synthesise the evidence extracted from 58 individual studies that collected 103 pieces of outcome information, using effect size as our variable of primary interest.[8] This effect size (Cohen's *d*) represents the difference (in standard deviations) between the mean values of the control group and the intervention group. For those interventions that measured the impact of the intervention

[8] However, most of the information was derived from the 53 studies that included information collected post-test that was not measured in units of 'change'.

dichotomously, we transformed these data on to the same scale. As many of the outcomes measured differed conceptually, we grouped these according to whether they represented clinical health measures, health behaviour measures or health self-efficacy measures. Other stratifications were also applied, reflecting the study design and the type of measures collected, and most of the analyses explored differences in post-test measures (not follow-up and not representing change data).

One of the most consistent trends in these data was in terms of their heterogeneity – our studies exhibited a great deal of diversity in effect size, which made interpreting their pooled effect sizes difficult without applying several caveats. However, this heterogeneity also provided an opportunity to explore some of the factors that could underlie this between-study variance. There were some exceptions to this pattern of heterogeneity. Outcomes measuring *change* in health behaviours, derived from five studies, exhibited little variation, although the combined effect size for these studies was close to zero, suggesting that the intervention had virtually no advantage over control conditions. These studies were not distinctive in terms of their health topic, study population characteristics or community engagement processes. Our effect size for the 19 studies measuring (post-test) clinical outcomes also exhibited less heterogeneity compared to other measures for outcome domains. A focus on low socio-economic groups appeared to initially explain some between-study heterogeneity that was detected. However, we were later not able to rule out the influence of the methodological quality of the study in driving differences in effect size. When we focused on those studies that collected measurements longitudinally (in a similar stratification to that which we had applied to studies measuring health behaviour outcomes), we found that there was little between-study variance ($Q=5.17$, $df=9$). Our pooled effect size for this group of studies stood at 0.192 (CI: 0.092-0.292), suggesting that interventions based on coalitions for community engagement made a small impact on clinical health outcomes, for longitudinally collected measures.

We explored the impact of coalitions in health interventions using measures of self-efficacy, where there was a substantial degree of heterogeneity between the ten studies we examined. A focus on gender appeared to explain part of this between-study variance, with studies focusing on females appearing to have higher effect sizes. However, we were unable to ascertain the extent to which this was driven by differences in the methodological quality of the studies or the measurement of their outcomes. As there were comparatively few studies, we did not apply further stratification. While the pooled effect size stood at 0.504, and all studies reported positive effects for self-efficacy (measured post-test), the substantial variation among studies undermines the precision of this estimate. Therefore, while self-efficacy is likely to improve as a result of coalition-based engagement processes implemented on health interventions, the extent of this improvement is difficult to quantify. Such variation may be expected, given the nature of self-efficacy as an outcome, with self-efficacy being more proximal to the intervention, potentially more transient or malleable as an outcome than behaviour or clinical outcomes (Gist and Mitchell 1992), and therefore being more sensitive to study measurement and study design differences that were not accounted for in these analyses.

Our strongest results appeared to derive from examining behavioural outcomes. Here, we made a distinction between studies that implemented a longitudinal design and those that implemented a repeated cross-sectional or other form of design that did not track

progress among individuals. This was to account for conceptual differences in the studies as well as on the basis of later sensitivity analyses. Looking first at those studies that did not implement a longitudinal design (n=15), we found that there was a substantial level of between-study variation in effect size, and that study-level age effects appeared to be a significant explanatory factor. Those studies that focused on children and young people appeared to have substantially larger effect sizes. While we posited that some of this might be due to the location of the intervention, with studies based in schools having less population churn and higher levels of full intervention 'dosage' than among those taking place in community or neighbourhoods, there was also evidence of publication bias among this group of studies, which may undermine our conclusions.

In contrast, the most robust evidence was derived from studies that measured health behaviour outcomes longitudinally, where studies that had higher levels of community engagement had higher effect sizes – compared to studies with a low effect size, the predicted pooled value from a multivariate random effects model was 0.56 higher. This represents a difference equivalent to a medium effect size, and suggests that a coalition alone is not sufficient to generate a medium or high effect size, but that the community needs to be highly involved (leading or collaborating) in designing, delivering and evaluating the intervention. We were unable to identify specific processes driving this result in these analyses, although this is explored further in our qualitative synthesis. We also observed that studies focused on injury and infection screening and prevention (IISP) had higher average effect sizes than studies with other foci, although the largest gulf was between studies with a focus on ATOD (alcohol, tobacco or drugs) and IISP, where there was a difference of 0.53 in predicted pooled effect size. Such a difference may be reflective of the nature of behaviour change between the different health foci. Behaviour change for ATOD-focused studies may deal with a change in practice as well as breaking chemical cycles of addiction. For studies focused on IISP, there were clear and relatively short-term advantages that were communicated to intervention beneficiaries of the benefits of taking measures to avoid, in this case, HIV/STI infection, injury or death from fire, or avoidance of cancer.

### *6.1.3 Qualitative comparative analysis*

The qualitative comparative analysis was undertaken to test whether specific combinations of conditions were seen in more (and less) effective studies. We found several combinations of conditions that led to successful (i.e. larger effect size) and less successful (i.e. smaller effect size) interventions. QCA indicated four configurations aligned with (i.e. related to) effective interventions: including lay-person delivery of the interventions; the inclusion of lay people in delivery of the interventions targeted across age groups; lay-person delivery of interventions focused on IISP, sexual health, organ donation or cancer prevention interventions; and lay-person delivery of interventions focused on infection or injury prevention targeted across an entire population. These findings suggest that intervention delivery by lay people may need to be present for coalitions to be successful, but intervention delivery by lay people alone may not lead to successful outcomes.

One configuration of conditions tended to occur in studies which showed lower effect sizes: those where coalitions employed a low extent of community engagement in design, delivery and evaluation.

## 6.2 Addressing the review's research questions

This project sought to further examine theories of community engagement that clarified our understanding of which interventions work, with which groups, and under what circumstances. To meet those aims, this review sought to address several research questions:

*1. How effective are community engagement approaches at improving health and wellbeing and reducing health inequalities?*

The findings from these analyses suggest that within projects that utilise coalitions, collaborations or partnerships with community members, higher behavioural outcome effect sizes are achieved through community members leading or collaborating on the design, delivery and evaluation of an intervention. Further, behavioural outcomes appear to be larger for interventions focused on infection or injury prevention in comparison to other health issues such as healthy eating, physical activity or mental health.

*2. Across disadvantaged groups, how effective are community engagement approaches at encouraging people to participate in activities to improve their health and wellbeing and realise their capabilities?*

While framework synthesis suggested that a high extent of community engagement was seen in low-income groups in particular, subsequent meta-analyses provided no firm evidence of differences between these or any other disadvantaged groups, due to methodological limitations of the studies.

*3. What processes and methods facilitate the realisation of community and individual capabilities and assets amongst disadvantaged groups?*

While evidence evaluating the processes of intervention implementation was located, no studies provided evaluations of the processes of community engagement. We identified descriptions of several processes of community engagement, including bidirectional communication, collective decision making, training support for intervention provision, allowing adequate time for relationship development; negotiation/reflection/conflict resolution skills and arranging meetings to suit community members' needs. Future evaluation of these processes could provide a starting point for making practical suggestions about how to undertake community engagement.

*4. Are there unintended consequences from adopting community engagement approaches?*

No evidence was found which suggested unintended consequences from adopting community engagement approaches. However, findings from our QCA suggested that a low extent of community engagement was aligned with lower effect sizes.

*5. What processes identified in the literature are more aligned with effective interventions, and which (if any) are more aligned with non-effective interventions?*

Finally, findings from our QCA supported those of the meta-analyses by providing tentative evidence that more community engagement was aligned with higher effect sizes; conversely, less community engagement across design, delivery and evaluation was aligned with lower effect sizes.

The findings from the syntheses presented here represent the first attempt, to our knowledge, to examine the processes of community engagement per se, rather than the processes of intervention implementation. These syntheses add detail to our previous theoretical development around the processes of community engagement. We found that evaluation of the processes of community engagement itself has not been undertaken routinely across studies. It was important to learn what processes had at least been reported, describe these processes, and look at patterns across populations and in studies indicating different levels of engagement. This helped us to understand whether some processes of community engagement, such as collective decision making, occurred more often in studies that also rated a higher extent community engagement. It has also shed light on whether these processes, or other characteristics of the studies such as populations under study or health topics, are more related to high community engagement and subsequent effective outcomes.

### 6.3 Setting the findings within current policy and research

Previous NICE guidance on community engagement presented overarching principles to evaluate community engagement across policy and initiative development and for working in partnership with community members (National Institute for Health and Care Excellence 2008). The guidance also outlined approaches to community engagement that included building on mutual trust and respect, identifying organisational change, agreeing levels of engagement and specific initiatives to engage community groups (National Institute for Health and Care Excellence 2008). The findings from this review add to that guidance in suggesting that continued involvement of community members throughout the entire lifespan of a collaboration (i.e. through design, delivery and evaluation) leads to greater effects. It also contributes by suggesting that community engagement in addressing some health issues may show greater effects than in others. Further, specific modifiable processes of community engagement (e.g. shared decision making) have been described in the literature which suggests *how* to build mutual trust and respect. These include bidirectional communication, collective decision making, and training support (for either engagers or engages). The findings from the synthesis of UK process evaluations undertaken for Stream 2 may further illuminate the processes.

The findings from the framework synthesis have added to previous findings concerning processes of community engagement based on our previous review's conceptual framework. Both reviews highlight the paucity of evidence on the underlying processes of community engagement and the lack of detail included in many studies, limiting exploration of how the specific processes that take place may contribute to variations in effect size. However, both reviews were consistent in highlighting the importance of consultation and collaboration as processes of community engagement. The current review also identified bidirectional communication, collective decision making, scheduling of meetings to meet community members' needs, adequate time for relationship development, and skills development in conflict resolution, negotiation, and reflection, as potentially useful processes that have been undertaken when developing coalitions.

Community based participatory research (CBPR) is a term often used in the studies included in the current review, although it was used as a shorthand term to describe a variety of different processes and often lacked a description. To ensure consistency in the implementation of CBPR, more guidance may be needed as to what CBPR should constitute

and the key activities that should take place in order for CBPR to be appropriately labelled.

Few studies included in either review considered the sustainability of the coalitions developed, and many did not see the coalition itself as a mechanism by which the intervention might be sustained and developed in the future.

In the current review, analyses from the framework synthesis and QCA suggested that studies focused on low-income groups recorded the use of specific processes of community engagement and had higher levels of extent of engagement, but the meta-analysis did not support this finding. Given that meta-analyses were undertaken on a smaller set of studies, some of which were not included in the framework synthesis or QCA, this is perhaps not surprising. Although timelines did not allow, further data extraction and analysis of the original review's included studies would have clarified this finding. Low-income groups are subject to poorer health outcomes than higher-income groups. Interventions with a high level of community engagement may be effective for the former group, so this approach has the potential to reduce health inequalities if targeted at such groups. And it should be noted that the meta-analyses showed that a higher level of community engagement was also associated with higher effect sizes, this finding highlights the potential of this approach in addressing income-based health inequalities in the future.

This review adds to the evidence that interventions that focus on self-efficacy outcomes may have the highest effect sizes, followed by studies focusing on health behaviours and then clinical outcomes. However, given the especially high levels of heterogeneity in effect size among studies measuring self-efficacy in both our previous review and this current one, we can say with greater confidence that studies focusing on health behaviours are likely to have higher effect sizes than studies focused on clinical outcomes. This is probably not a reflection around the quality of interventions of studies with a clinical focus compared to a behavioural focus, but more a reflection of the added complexity in altering clinical outcomes, which in the majority of studies included in both reviews, is implicitly dependent first on a change in behaviour.

The current review's analyses present evidence that the association between community engagement and effect size is additive, and the greater the number of processes implemented, the higher the average effect size. This was most prominent in the analyses of studies that collected health behaviour outcomes.

Unlike the analyses undertaken in our previous review, this review presented analysis that suggests that specific health topics (injury and infection prevention and screening) have an association with effect size. In our previous review, the between-study variance in effect size based on health topic was not significant, although in the present review, we found that studies focusing on the reduction of risk of injury or infection had higher effect sizes compared to studies with a focus on other health topics. Studies that focused on alcohol, tobacco or drugs had lower effect sizes, and we interpreted this as being reflective of the need to both change practice/behaviour and potentially cycles of addiction in these studies.

In the present review, analyses suggested that studies implementing a longitudinal design were likely to have higher effect sizes than those that implemented a repeated cross-sectional or other design. This is likely to reflect that many of the populations included in cross-sectional studies were transient, and not all respondents included in post-test measures would have received the full 'dose' of the intervention. Some studies, including Phillips and colleagues (2014), explicitly attributed population churn as being a characteristic of their study population and a potential factor in explaining low effect sizes. This highlights the need for community engagement studies working within neighbourhoods or communities of geography to (i) establish or measure the degree of population change in their study areas before the study begins; and (ii) incorporate this element of change into the study design itself.

Findings from the QCA extend previous ideas about lay delivery of interventions, extent of engagement and impacts related to different health issues. Lay delivery was found to be the most frequently recorded form of community engagement in the studies included in our original review. The current review builds on this through uncovering evidence suggesting that, within studies of coalitions, employing lay delivery had additional impact on successful outcomes.

In keeping with the results of the meta-analyses, those studies with lower extent of engagement were also those with lower effect sizes. Extent of engagement is an important moderator in determining intervention extent. Community engagement is therefore more than simply a box-ticking exercise, but involves active and sustained engagement over the course of the intervention.

Coalitions focused on injury or infection prevention or screening, sexual health, organ donation and cancer prevention were most successful in comparison to those focused on healthy eating, physical activity, mental health or substance use. These health topics represent a combination of interventions that either help beneficiaries reduce their own personal risk of developing communicable and non-communicable diseases or, in the case of organ donation, have an altruistic component in helping other community members. Many of the cancer prevention and sexual health interventions had an element of screening involved (e.g. for HIV/other sexually transmitted infections (STIs), STIs leading to cancer, or direct cancer screening), meaning that the intervention worked to encourage beneficiaries to adopt short-term/infrequent behaviours and actions, as opposed to changing regular habits or practices, such as changing diet or smoking activity.

### 6.4 Limitations of and gaps in the evidence

While the evidence provided descriptions of several modifiable processes of community engagement, no evaluations were located evaluating the impact of these processes on subsequent intervention development or health outcomes. The evidence supporting the impacts of community engagement strategies on health outcomes is in its infancy. Most reports are case studies (Minkler and Wallerstein 2010). Only Chung et al. (2010) report undertaking a before-after design in order to evaluate the processes of community engagement; however, this evaluated the impact of community engagement on intervention development rather than health outcomes, although the authors expect to evaluate this in future.

Other limitations were the lack of identified studies undertaking direct comparisons of community engagement as the intervention under study versus no community engagement. Further, few studies of high methodological quality were located.

### 6.5 Strengths of the review

This review represents the first attempt we are aware of to collate and evaluate specific modifiable processes of community engagement involving coalitions, producing conceptually sound and plausible results. It is a starting point to explore how community engagement is effective and what modifiable processes could be considered in setting up coalitions, as well as what to evaluate and report about community engagement to inform others about 'what works'. This highlights the importance of robust study design in evaluating community engagement interventions and the broader implications for the design of future interventions.

### 6.6 Limitations of the review

Some limitations arising from this review should be considered. The focused timelines required for guideline development meant that we were unable to examine studies from non-OECD countries. These studies have been grouped for future analysis. For the same reasons we could not look at studies of non-disadvantaged populations and other forms of community engagement, such as those utilising peer delivery alone. However a systematic review of peer delivery strategies is in process (J. Harris, personal communication) and a synthesis of studies examining social media interventions will be undertaken for Review 3 of this project.

As with our previous review, the current review identified a large number of heterogeneous studies of community engagement. Despite the large number of studies, there was a lack of extractable data in many studies suitable for statistical or qualitative synthesis. Once we stratified by study type, this resulted in a relatively small pool of studies, making it challenging to provide robust estimates of association.

Paradoxically, there was a fair lack of diversity in health topic. The literature was predominated by studies of healthy eating and physical activity, even though our subsequent analyses indicated that this might be a less effective focus for community engagement than injury and infection prevention.

The authors' descriptions of what constituted a coalition varied considerably, as evidenced by the variation in the extent of engagement across design, delivery and evaluation of interventions. We were reliant on descriptions of community engagement provided in published reports; despite contacting authors for further clarification, little supplemental evidence was provided. The uncertainty about the potential confounding due to lack of information on interventions and populations meant that we were unable to undertake robust multivariate models.

### 6.7 Applicability of the evidence to UK populations in the scope

It is important to note that only 3 of the 26 included studies were conducted in the UK, and several of the studies focused on ethnic groups specific to America (i.e. African-American, Latino and Hispanic). This makes it necessary to interpret findings with caution.

Implementing these interventions in the UK context requires collaboration and collective input from UK communities.

### 6.8 What these findings add to Review 1

The analyses undertaken in Review 2 both confirm and further refine those found in Review 1. For example, Review 2 findings confirm those of Review 1 that suggest that a high extent of community engagement across design, delivery and evaluation is associated with greater beneficial effects of health interventions, in comparison to either moderate or low extent of community engagement.

However, the more detailed analysis undertaken in the Review 2 meta-analysis suggests that health behaviour outcomes might be larger in studies focused specifically on injury or infection prevention/screening, rather than in the larger list of health topics suggested in Review 1. In addition, conclusions in Review 1 drawn from the examination of population subsets (i.e. women, men, children/young people, low-income populations) were not borne out in the more detailed analyses undertaken in Review 2. This is most likely due to the addition of studies from the original review of community engagement to the meta-analysis, which had two effects: to add power to the size of the sample for meta-analysis, thus making the findings more robust; and to reduce the amount of data to which findings from the framework synthesis and QCA can be attributed, due to a lack of coding on community engagement processes from the original review's included studies.

# 7. Conclusions and evidence statements

Taken together, the findings across all three syntheses in this review suggest that community-led or community collaboration projects which design, deliver and evaluate health interventions are associated with larger behavioural outcomes; and that behavioural outcomes are larger for projects focused on injury or infection prevention compared to other health issues. Where coalitions, collaborations and partnerships with community members include the use of bidirectional communication, collective decision making and training support for intervention provision, a higher extent of community engagement across the project's design, delivery and evaluation was also found. Effective configurations of engagement within collaborations and coalitions generally include peer or lay delivery; and projects with a low extent of engagement were aligned with lower effect sizes.

## 7.1 Evidence statements for effectiveness of community engagement via coalitions, collaborations or partnerships

Several evidence statements can be derived from the analyses we conducted, where data allowed. These have been structured to address the review's original research questions:

### *7.1.1 How effective are community engagement approaches at improving health and wellbeing and reducing health inequalities?*

The overall pooled estimate of effect size for coalition-based community engagement health interventions provides moderate evidence of positive effects across all three outcome domains, with the highest effect size observed for self-efficacy (based on five methodologically sound studies (+)[6-10] and five studies not deemed to be methodologically sound (-)[1-5]), followed by behaviours (based on twenty-three methodologically sound studies[6-10,32-49]) and twenty-six studies not deemed to be methodologically sound[1-5,11-31]) and then clinical outcomes (based on ten methodologically unsound studies[2,12,18-20,22,26,27,34,50] and nine studies deemed to be sound [8,33,39,42,47,49,51-3]).[9] While the effects recorded in the majority of studies were positive, and pooled estimates were statistically significantly above zero, there was, however, substantial heterogeneity in all effect size estimates, and many individual studies did not attain effect sizes that were statistically significantly above zero, leading us to conclude that there was only moderate evidence of effectiveness for this form of community engagement.

[9] All evidence described here based on post-test outcomes that weren't measured in units of 'change'.

1. Harper et al. 2009 (-)
2. Resnicow et al. 1992 (-)
3. Russell et al. 2010 (-)
4. Secker-Walker et al. 2000 (-)
5. Zoellner et al. 2013 (-)
6. Andrews et al. 2007 (+)
7. Campbell et al. 1999 (+)
8. Islam et al. 2013 (+)
9. Perry et al. 1996 (+)
10. Resnicow et al. 2004 (+)
11. Andrews et al. 2012 (-)
12. Brownson et al. 1996 (-)
13. Brownson et al. 2005 (-)
14. Daniel et al. 1999 (-)
15. Eades et al. 2012 (-)
16. Foerster et al. 1998 (-)
17. Hancock et al. 2001 (-)
18. Hoelscher et al. 2010 (-)
19. Kneipp et al. 2011 (-)
20. Kumpusalo et al. 1996 (-)
21. Lewis et al. 1993 (-)
22. Macaulay et al. 1997 (-)
23. Martin et al. 2013 (-)
24. McAlister et al. 1992 (-)
25. Mendoza et al. 2009 (-)
26. O'Loughlin et al. 1999 (-)
27. Phillips et al. 2014 (-)
28. Schensul et al. 2009 (-)
29. Schwarz et al. 1993 (-)
30. Wright et al. 1997 (-)
31. Zhou et al. 2003 (-)
32. Andersen et al. 2013 (+)
33. Becker et al. 2005 (+)
34. Chen et al. 2013 (+)
35. Davis et al. 1995 (+)
36. Dedobbeleer et al. 2001 (+)
37. Dodge et al. 2013 (+)
38. Dzewaltowski et al. -2010 (+)
39. Kong et al. 2013 (+)
40. Lindenberg et al. 2002 (+)
41. LoSciuto et al. 1999 (+)
42. Nafziger et al. 2001 (+)
43. Rhodes et al. 2011 (+)
44. Schinke et al. 2000 (+)
45. Schorling et al. 1997 (+)
46. Shelley et al. 2008 (+)
47. Simmons et al. 1998 (+)
48. Winkleby et al. 2004 (+)
49. Yanek et al. 2001 (+)
50. Davidson et al. 1994 (-)
51. Kieffer et al. 2013 (+)
52. Lindqvist et al. 1999 (+)
53. Robinson et al. 2003 (+)

To examine specific aspects of this research question, we addressed some sub-questions discussed below.

#### *7.1.1.1 Do health outcome effects differ for 'distal' (e.g. self-efficacy), 'intermediate' (e.g. health behaviour), or 'proximal' (clinical/physiological measure) outcomes?*

Only three studies (two of unsound and one of sound methodological quality[1-3]) collected outcomes across all three health domains. We observed weak evidence of a correlation between effect-sizes across domains among those (few) studies that included measures across these domains.

1. Islam et al. 2013 (+)
2. Resnicow et al. 1992 (-)
3. Zoellner et al. 2013 (-)

***7.1.1.2 Do direct comparisons of community engagement (i.e. studies that test a community engagement intervention versus the same intervention without community engagement) differ in health outcome effects from indirect comparisons of community engagement (e.g. those that test community engagement versus usual care)?***

Because of the low number of studies with direct comparisons in both sets of studies combined (n=3), and of our analysis strategy, which examined effect sizes separately across different outcome domains, we were unable to synthesise the results to address this research question. Thus we found no evidence of an effect of community engagement as a sole strategy on outcomes.

***7.1.1.3 What is the relationship between the extent of community engagement (high, moderate or low) and health outcome effects?***

We found no evidence supporting the hypothesis that studies with higher levels of community engagement also tended to have higher effect sizes for clinical or self-efficacy outcomes, and only weak evidence for an effect on behavioural outcomes (based on 23 methodologically sound studies[27-49] and twenty-six studies not deemed to be methodologically sound[1-26]).

However, we uncovered moderate evidence based on 20 studies deemed methodologically sound[5,20,27-34,36,38,39,41-49] and 14 studies considered to be methodologically unsound[1,4,5,6-8,10,14,15,19,20,22,26,37] of a link between extent of engagement and outcome effects for a subset of behavioural outcomes, those studies with a longitudinal design. Here extent of engagement was found to explain a substantial part of between-study variance for this group of studies. For clinical outcomes or self-efficacy outcomes, extent of engagement did not help to explain differences in effect size between studies.

1. Andrews et al. 2012 (-)
2. Brownson et al. 1996 (-)
3. Brownson et al. 2005 (-)
4. Daniel et al. 1999 (-)
5. Eades et al. 2012 (-)
6. Foerster et al. 1998 (-)
7. Hancock et al. 2001 (-)
8. Harper et al. 2009 (-)
9. Hoelscher et al. 2010 (-)
10. Kneipp et al. 2011 (-)
11. Kumpusalo et al. 1996 (-)
12. Lewis et al. 1993 (-)
13. Macaulay et al. 1997 (-)
14. Martin et al. 2013 (-)
15. McAlister et al. 1992 (-)
16. Mendoza et al. 2009 (-)
17. O'Loughlin et al. 1999 (-)
18. Phillips et al. 2014 (-)
19. Resnicow et al. 1992 (-)
20. Russell et al. 2010 (-)
21. Schensul et al. 2009 (-)
22. Schwarz et al. 1993 (-)
23. Secker-Walker et al. 2000 (-)
24. Wright et al. 1997 (-)
25. Zhou et al. 2003 (-)
26. Zoellner et al. 2013 (-)
27. Andersen et al. 2013 (+)
28. Andrews et al. 2007 (+)
29. Becker et al. 2005 (+)
30. Campbell et al. 1999 (+)
31. Chen et al. 2013 (+)
32. Davis et al. 1995 (+)
33. Dedobbeleer et al. 2001 (+)
34. Dodge et al. 2013 (+)
35. Dzewaltowski et al. -2010 (+)
36. Islam et al. 2013 (+)
37. Kong et al. 2013 (+)
38. Lindenberg et al. 2002 (+)
39. LoSciuto et al. 1999 (+)
40. Nafziger et al. 2001 (+)
41. Perry et al. 1996 (+)
42. Resnicow et al. 2004 (+)
43. Rhodes et al. 2011 (+)
44. Schinke et al. 2000 (+)

45. Schorling et al. 1997 (+)
46. Shelley et al. 2008 (+)
47. Simmons et al. 1998 (+)
48. Winkleby et al. 2004 (+)
49. Yanek et al. 2001 (+)

### *7.1.2 Across disadvantaged groups, how effective are community engagement approaches at encouraging people to participate in activities to improve their health and wellbeing and realise their capabilities?*

To address this research question, several sub-questions were posed:

#### ***7.1.2.1 Do health outcome effects differ for different age groups?***

Overall, we observed no evidence to suggest that studies that focused on younger or older groups achieved larger or smaller effect sizes across most domains. In the case of studies that collected behavioural outcomes through repeated cross-sectional designs, we did find weak evidence that studies focused on children achieved higher effect sizes than studies that did not, based on two studies deemed methodologically sound (+)[14,15] and 13 studies not deemed to be methodologically sound (-)[1,2,3,4,5,6,7,8,9,10,11,12,13]. However, we were unable to discount other factors, including lower population turnover, in studies with a focus on children.

1. Brownson et al. 2005 (-)
2. Lewis et al. 1993 (-)
3. Hoelscher et al. 2010 (-)
4. Zhou et al. 2003 (-)
5. Mendoza et al. 2009 (-)
6. Brownson et al. 1996 (-)
7. Schensul et al. 2009 (-)
8. O'Loughlin et al. 1999 (-)
9. Secker-Walker et al. 2000 (-)
10. Kumpusalo et al. 1996 (-)
11. Wright et al. 1997 (-)
12. Phillips et al. 2014 (-)
13. Macaulay et al. 1997 (-)
14. Dzewaltowski et al. -2010 (+)
15. Nafziger et al. 2001 (+)

#### ***7.1.2.2 Do health outcome effects differ for studies targeting men only versus those targeting women only?***

Overall, we observed no evidence to suggest that studies that focus on either gender had either lower or higher effect sizes for behavioural or clinical outcomes. There was weak evidence suggestive of a study-level gender difference in effect size in terms of self-efficacy outcomes, where studies which focused on females had higher effect sizes than those that did not, although this was based on a very small sample of five studies deemed methodologically sound (+)[4,6,7,9,10] and five studies deemed to be methodologically unsound (-)[1,2,3,5,8]. However, we were not able to rule out the influence of the unit of measurement (binary vs continuous) or the methodological quality of the study as explanatory factors. The relationship between gender and self-efficacy may be an area for future investigation, although the current weak evidence means that this result should not be used to influence guidelines around the issue.

1. Zoellner et al. 2013 (-)
2. Secker-Walker et al. 2000 (-)
3. Harper et al. 2009 (-)
4. Islam et al. 2013 (+)
5. Resnicow et al. 1992 (-)
6. Perry et al. 1996 (+)
7. Resnicow et al. 2004 (+)
8. Russell et al. 2010 (-)
9. Andrews et al. 2007 (+)
10. Campbell et al. 1999 (+)

***7.1.2.3 Do health outcome effects differ for studies specifically developed for low-income groups versus those that are not?***

The framework synthesis findings in Review 1 suggested that a high extent of community engagement was seen in interventions for low-income groups. However, subsequent meta-analyses provided no evidence of differences in the number of processes recorded between low-income groups versus those with any other type of disadvantage, although this may be due to the methodological limitations of the studies (some of which were not included in the meta-analyses). In total, in the meta-analyses, nine studies focused on low-income groups, of which four were of sound methodological quality[6-9] and five were judged to be unsound[1-5]. There were some indications that studies with a focus on low-income groups that collected clinical outcomes[1,4,7,9] had lower effect sizes in this domain compared to those that did not, although we were not able to rule out the influence of study methodological quality in driving this result.

1. Hoelscher et al. 2010 (-)
2. Martin et al. 2013 (-)
3. O'Loughlin et al. 1999 (-)
4. Phillips et al. 2014 (-)
5. Wright et al. 2013 (-)
6. Dodge et al. 2013 (+)
7. Dzewaltowski et al. 2010 (+)
8. Kloek et al. 2006 (+)
9. Kong et al. 2013 (+)

*7.1.3 What processes and methods facilitate the realisation of community and individual capabilities and assets amongst disadvantaged groups?*

To examine this question in further detail, we asked the following sub-question:

***7.1.3.1 Which potentially modifiable processes of community engagement have been evaluated?***

While evidence evaluating the processes of *intervention implementation* was located, no studies provided evaluations of the processes of *community engagement*. We identified descriptions of several processes of community engagement, providing weak evidence of processes, including bidirectional communication[1-5,8,9,11-14,17-23], collective decision making[1-5,9,10,12-14,16-23], training support for intervention provision[1,2,4-7,9-17,20-23], allowing adequate time for relationship development[3,9,12-14], negotiation/reflection[12], conflict resolution skills[12,21], arranging meetings to suit community members' needs[4,17], use of external facilitators[3,21], administrative support[15] and interagency working[15].

1. Kneipp et al. 2011 (-)
2. Martin et al. 2013 (-)
3. Zoellner et al. 2013 (-)
4. Kong et al. 2013 (+)
5. Islam et al. 2013 (+)
6. Russell et al. 2010 (-)
7. Eades et al. 2012 (-)
8. Andersen et al. 2013 (+)
9. Rhodes et al. 2011 (+)
10. Chen et al. 2013 (+)
11. Andrews et al. 2012 (-)
12. Berg et al. 2009 (+)
13. Bonell et al. 2010 (+)
14. Cohen et al. 2013 (-)
15. Dzewaltowski et al. 2010 (+)
16. Hoelscher et al. 2010 (-)
17. Kieffer et al. 2013 (+)
18. Lassen et al. 2011 (-)
19. Parikh et al. 2010 (+)
20. Phillips et al. 2014 (-)
21. Plescia et al. 2008 (-)
22. Segal et al. 2009 (+)

23. Woods et al. 2013 (-)

***7.1.3.2 Are potentially modifiable processes of community engagement associated with health outcome effects?***

When we created a cumulative score based on the number of community engagement processes across studies included in both the original and the current reviews, we found moderate evidence from 20 studies deemed methodologically sound[15-34] and 14 studies considered to be methodologically unsound[1-14] that each additional engagement process was associated with larger effect size; but only for a model that included behavioural outcomes measured longitudinally. However, we were unable to identify a single modifiable process that significantly helped to explain differences in effect size. Evidence was thus considered weak because, although several studies described these processes, none were evaluated; and while no single processes of engagement appeared to be related to effect size, longitudinal studies with more processes of engagement described were associated with larger effects. Future evaluation of these processes are needed to provide robust evidence; however, this analysis acts as a starting point for making practical suggestions about how communities and providers can work together to undertake community engagement.

1. Kneipp et al. 2011 (-)
2. Foerster et al. 1998 (-)
3. Harper et al. 2009 (-)
4. Martin et al. 2013 (-)
5. Resnicow et al. 1992 (-)
6. McAlister et al. 1992 (-)
7. Hancock et al. 2001 (-)
8. Zoellner et al. 2013 (-)
9. Schwarz et al. 1993 (-)
10. Andrews et al. 2012 (-)
11. Daniel et al. 1999 (-)
12. Russell et al. 2010 (-)
13. Eades et al. 2012 (-)
14. Kong et al. 2013 (+)
15. Islam et al. 2013 (+)
16. Shelley et al. 2008 (+)
17. Schinke et al. 2000 (+)
18. Dodge et al. 2013 (+)
19. Winkleby et al. 2004 (+)
20. Resnicow et al. 2004 (+)
21. Perry et al. 1996 (+)
22. LoSciuto et al. 1999 (+)
23. Yanek et al. 2001 (+)
24. Lindenberg et al. 2002 (+)
25. Becker et al. 2005 (+)
26. Dedobbeleer et al. 2001 (+)
27. Andersen et al. 2013 (+)
28. Schorling et al. 1997 (+)
29. Campbell et al. 1999 (+)
30. Simmons et al. 1998 (+)
31. Davis et al. 1995 (+)
32. Rhodes et al. 2011 (+)
33. Andrews et al. 2007 (+)
34. Chen et al. 2013 (+)

*7.1.4 Are there unintended consequences from adopting community engagement approaches?*

No evidence was found which suggested unintended consequences from adopting community engagement approaches. However, weak evidence from 50 (24 methodologically unsound and 26 methodologically sound) studies examined in the qualitative comparative analysis suggested that a low extent of community engagement was aligned with lower effect sizes[1-51].

1. Andersen et al. 2013 (+)
2. Andrews et al. 2007 (+)
3. Andrews et al. 2012 (-)
4. Becker et al. 2005 (+)

5. Brownson et al. 1996 (-)
6. Brownson et al. 2005 (-)
7. Campbell et al. 1999 (+)
8. Chen et al. 2013 (+)
9. Daniel et al. 1999 (-)
10. Davis et al. 1995 (+)
11. Dedobbeleer et al. 2001 (+)
12. Dodge et al. 2013 (+)
13. Dzewaltowski et al. 2010 (+)
14. Eades et al. 2012 (-)
15. Foerster et al. 1998 (-)
16. Hancock et al. 2001 (-)
17. Harper et al. 2009 (-)
18. Hoelscher et al. 2010)(-)
19. Islam 2013 et al. (+)
20. Kneipp et al. 2011 (-)
21. Kong et al. 2013 (+)
22. Kumpusalo et al. 1996 (-)
23. Lewis et al. 1993 (-)
24. Lindenberg 2002 (+)
25. LoSciuto et al. 1999 (+)
26. Macaulay et al. 1997 (-)
27. Martin et al. 2013 (-)
28. McAlister et al. 1992 (-)
29. Mendoza et al. 2009 (-)
30. Nafziger et al. 2001 (+)
31. O'Loughlin et al. 1999 (-)
32. Perry et al. 1996 (+)
33. Phillips et al. 2014 (-)
34. Resnicow et al. 1992 (-)
35. Resnicow et al. 2004 (+)
36. Rhodes et al. 2011 (+)
37. Robinson et al. 2003 (+)
38. Russell et al. 2010 (-)
39. Schensul et al. 2009 (-)
40. Schinke et al. 2000 (+)
41. Schorling et al. 1997 (+)
42. Schwarz et al. 1993 (-)
43. Secker-Walker et al. 2000 (-)
44. Shelley et al. 2008 (+)
45. Simmons et al. 1998 (+)
46. Winkleby et al. 2004 (+)
47. Wright et al. 1997 (+)
48. Yanek et al. 2001 (+)
49. Zhou et al. 2003 (-)
50. Zoellner et al. 2013 (-)

### *7.1.5 What processes identified in the literature are more aligned with effective interventions, and which (if any) are more aligned with non-effective interventions?*

Finally, findings from our QCA supported those of the meta-analyses by providing tentative evidence that more community engagement was aligned with higher effect sizes; conversely, less community engagement across design, delivery and evaluation was aligned with lower effect sizes[1-50] (see the list in point 4 above).

*Applicability of evidence statements*

Across all of the evidence statements, the evidence can be considered to be partially applicable, given the large proportion of non-UK focused studies.

**Note:** We did not provide details of studies and their risk of bias ratings for evidence statements from the meta-analyses where high levels of heterogeneity prevented us from providing pooled estimates of effect size; this included the results from meta-regression. Details of studies and their risk of bias ratings for evidence statements based on meta-regression were not provided for reasons of consistency. In many cases the results were based on either (i) covariates that had been modelled continuously (representing a step change) where the result was based on all studies, rather than a difference between two groups; or (ii) multivariate analyses where the result represented the impact of one study-level characteristic controlling for another.

# 8. References

*Papers included in the review are indicated with an asterisk**


*Andersen E, Høstmark AT, Holme I, Anderssen SA (2013) Intervention effects on physical activity and insulin levels in men of Pakistani origin living in Oslo: A randomised controlled trial. *Journal of Immigrant and Minority Health,* 15: 101-110.

*Andrews JO, Bentley G, Crawford S, Pretlow L, Tingen MS (2007) Using community-based participatory research to develop a culturally sensitive smoking cessation intervention with public housing neighborhoods. *Ethnicity and Disease,* 17(2): 331-337.

*Andrews AM, Zhang N, Magee JC, Chapman R, Langford AT, Resnicow K (2012) Increasing donor designation through black churches: results of a randomized trial. *Progress in Transplantation,* 22: 161-167.

Arblaster L, Lambert M, Entwistle M, Forster M, Fullerton D, Sheldon T, Watt I (1996) A systematic review of the effectiveness of health service interventions aimed at reducing inequalities in health. *Journal of Health Services Research and Policy*, 1(2): 93-103.

*Becker DM, Yanek LR, Johnson WR, Garrett, D, Moy TF, Reynolds SS, Blumenthal RS, Vaidya D, Becker LC (2005) Impact of a community-based multiple risk factor intervention on cardiovascular risk in black families with a history of premature coronary disease. *Circulation,* 111(10): 1298-1304.

*Berg M, Coman E, Schensul JY (2009) Youth action research for prevention: A multi-level intervention designed to increase efficacy and empowerment among urban youth. *American Journal of Community Psychology,* 43(3-4): 345-359.

Bergstrom H, Hagstromer M, Hagberg J, Elinder LS (2013) A multi-component universal intervention to improve diet and physical activity among adults with intellectual disabilities in community residences: A cluster randomised controlled trial. *Research in Developmental Disabilities* 34: 3847-3857.

*Bonell C, Sorhaindo A, Allen ES, Strange VJ, Wiggins M, Fletcher A, Oakley AR, Bond LM, Flay BR, Patton GC, Rhodes T (2010) Pilot multimethod trial of a school-ethos intervention to reduce substance use: Building hypotheses about upstream pathways to prevention. *Journal of Adolescent Health,* 47(6): 555-563.

Borenstein M (2009) Effect sizes for continuous data. In: Harris MC, Iedges LV, Valentine JC, (eds) *The handbook of research synthesis and meta-analysis*. New York, NY: Russell Sage Foundation, pages 221-235.

Borenstein, M, Hedges LV, Higgins JP, Rothstein HR (2011) *Introduction to meta-analysis*. New York: John Wiley and Sons.

Brenner B, Manice M (2011) Community engagement in children's environmental health research. *Mount Sinai Journal of Medicine*, 78(1): 85-97.

*Brownson RC, Smith C, Pratt M, Mack NE, Jackson-Thompson J, Dean CG, Dabney S, Wilkerson JC (1996) Preventing cardiovascular disease through community-based risk reduction: The Bootheel Heart Health Project. *American Journal of Public Health,* 86(2): 206-213.

*Brownson RC, Hagood, L, Lovegreen SL, Britton, B, Caito NM, Elliott MB, Emery J, Haire-Joshu D, Hicks D, Johnson B, McGill JB, Morton S, Rhodes G, Thurman T, Tune D (2005) A multilevel ecological approach to promoting walking in rural communities. *Preventive Medicine,* 41(5): 837-842.

Brunton G, Caird J, Stokes G, Stansfield C, Kneale D, Richardson M, Thomas J (2015) *Review 1: Community engagement for health via coalitions, collaborations and partnerships: A systematic review and meta-analysis*. London: EPPI Centre, Social Science Research Unit, UCL Institute of Education.

*Campbell MK, Demark-Wahnefried W, Symons M, Kalsbeek WD, Dodds J, Cowan A, Jackson B, Motsinger B, Hoben K, Lashley J, Demissie S, McClelland JW (1999) Fruit and vegetable consumption and prevention of cancer: The Black Churches United for Better Health project. *American Journal of Public Health,* 89(9): 1390-1396.

*Chen MS Jr, Fang DM, Stewart SL, Ly MY, Lee S, Dang JH, Nguyen TT, Maxwell AE, Bowlus CL, Bastani R (2013) Increasing hepatitis B screening for Hmong adults: Results from a randomized controlled community-based study. *Cancer Epidemiology, Biomarkers and Prevention,* 22: 782-791.

Chinn S (2000) A simple method for converting an odds ratio to effects size for use in meta-analysis. *Statistics in Medicine*, 19: 3127-3131.

Chung B, Jones L, Dixon EL Miranda J, Wells K; Community Partners in Care Steering Council (2010) Using a community partnered participatory research approach to implement a randomized controlled trial: Planning the design of Community Partners in Care. *Journal of Health Care for the Poor and Underserved*, 21(3): 780-795.

*Cohen DA, Han B, Derose KP, Williamson S, Marsh T, McKenzie TL (2013) Physical activity in parks: A randomized controlled trial using community engagement. *American Journal of Preventive Medicine,* 45(5): 590-597.

Cohen, J (1992) A power primer. *Psychological Bulletin*, 112(1): 155-159.

*Daniel M, Green LW, Marion SA, Gamble D, Herbert CP, Hertzman C, Sheps SB (1999) Effectiveness of community-directed diabetes prevention and control in a rural Aboriginal population in British Columbia, Canada. *Social Science and Medicine,* 48(6): 815-832.

*Davidson LL, Durkin MS, Kuhn L, O'Connor P, Barlow B, Heagarty MC (1994) The impact of the Safe Kids/Healthy Neighborhoods Injury Prevention Program in Harlem, 1988 through 1991. *American Journal of Public Health,* 84(4): 580-586.

*Davis SM, Lambert LC, Gomez Y, Skipper B (1995) Southwest cardiovascular curriculum project: Study findings for American Indian elementary students. *Journal of Health Education,* 26(Suppl. 2): S72-S81.

*Dedobbeleer N, Desjardins S (2001) Outcomes of an ecological and participatory approach to prevent alcohol and other drug 'abuse' among multiethnic adolescents. *Substance Use and Misuse,* 36(13): 1959-1991.

Department of Health (2002) *Shifting the balance of power: The next steps*. London: Department of Health.

Department of Health (2005) *Commissioning a patient-led National Health Service*. London: Department of Health.

Department of Health (2006a) *Our health, our care, our say: A new direction for community services*. Norwich: The Stationery Office.

Department of Health (2006b) *A stronger local voice: A framework for creating a stronger local voice in the development of health and social care services*. London: Department of Health.

Department of Health (2006c) *Health reform in England: Update and commissioning framework*. London: Department of Health.

Department of Health (2010) *Health Inequalities National Support Team resources*. http://www.institute.nhs.uk/commissioning/general/health_inequalities_national_support_team_resources.html (accessed 16 November 2014)

*Derose KP, Marsh T, Mariscal M, Pina-Cortez S, Cohen DA (2014) Involving community stakeholders to increase park use and physical activity. *Preventive Medicine*, 64: 14-19.

*Dodge KA, Goodman WB, Murphy RA, O'Donnell K, Sato J (2013) Randomized controlled trial of universal postnatal nurse home visiting: Impact on emergency care. *Pediatrics,* 132 (Suppl. 2): S140-146.

*Dzewaltowski D, Rosenkranz R, Geller K (2010) HOP N after-school project: An obesity prevention randomized controlled trial. *International Journal of Behavioral Nutrition and Physical Activity,* 7: 90.

*Eades SJ, Sanson-Fisher RW, Wenitong M, Panaretto K, D'Este C, Gilligan C, Stewart J (2012) An intensive smoking intervention for pregnant Aboriginal and Torres Strait Islander women: A randomised controlled trial. *Medical Journal of Australia,* 197(1): 42-46.

Elder JP, Wildey M, de Moor C, Sallis J F.Jr, Eckhardt L, Edwards C, Erickson A, Golbeck A, Hovell M, Johnston D (1993) The long-term prevention of tobacco use among junior high school students: classroom and telephone interventions. *American Journal of Public Health,* 83(9): 1239-1244.

*Foerster SB, Gregson, J, Beall DL, Hudes M, Magnuson H, Livingston S, Davis MA, Joy AB, Garbolino T (1998) The California Children's 5 a Day-Power Play! Campaign: Evaluation of a large-scale social marketing initiative. *Family and Community Health,* 21(1) 46-64.

Gist ME, Mitchell TR (1992) Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management Review*, 17(2): 183-211.

*Hancock L, Sanson-Fisher R, Perkins J, McClintock A, Howley P, Gibberd R (2001) Effect of a community action program on adult quit smoking rates in rural Australian towns: The CART project. *Preventive Medicine,* 32(2): 118-127.

Harbord RM, Harris RJ, Sterne JA (2009) Updated tests for small-study effects in meta-analyses. *Stata Journal*, 9(2): 197-210.

*Harper GW, Bangi AK, Sanchez B, Doll M, Pedraza A (2009) A quasi-experimental evaluation of a community-based HIV prevention intervention for Mexican American female adolescents: The SHERO's program. *AIDS Education and Prevention,* 21(5 Suppl.): 109-123.

Higgins JP and Green S (eds) (2008) *Cochrane handbook for systematic reviews of interventions*, version 5. Chichester: Wiley-Blackwell.

*Hoelscher D, Springer A, Ranjit J, Perry CL, Evans AE, Stigler M, Kelder SH (2010) Reductions in child obesity among disadvantaged school children with community involvement: The Travis County CATCH trial. *Obesity,* 18(Suppl. 1): S36-44.

*Islam NS, Zanowiak JM, Wyatt LC, Chun K, Lee L, Kwon SC, Trinh-Shevrin C (2013) A randomized-controlled, pilot intervention on diabetes prevention and healthy lifestyles in the New York City Korean community. *Journal of Community Health*, 38: 1030-1041.

Jamal F, Bertotti M, Lorenc T, Harden A (2013) Reviewing conceptualisations of community: reflections on a meta-narrative approach. *Qualitative Research*, 2013: doi:10.1177/1468794113509262. http://qrj.sagepub.com/content/early/2013/11/06/1468794113509262.abstract. Accessed: 6 June 2014.

Jones L, Koegel P, Wells KB (2008) Bringing experimental design to community-based participatory research. In Minkler M, Wallerstein N (eds) *Community-based participatory research: From processes to outcomes*. San Francisco: John Wiley and Sons.

*Kieffer EC, Caldwell CH, Welmerink DB, Welch KB, Sinco BR, Guzman JR (2013) Effect of the healthy MOMs lifestyle intervention on reducing depressive symptoms among pregnant Latinas. *American Journal of Community Psychology*, 51(1-2): 76-89.

*Kieffer EC, Welmerink DB, Sinco BR, Welch KB, Rees Clayton EM, Schumann CY, Uhley VE (2014) Dietary outcomes in a Spanish-language randomized controlled diabetes prevention trial with pregnant Latinas. *American Journal of Public Health*, 104(3): 526-533.

*Kim KH, Linnan, L, Campbell MK, Brooks, C, Koenig HG, Wiesen, C (2008) The WORD (wholeness, oneness, righteousness, deliverance): A faith-based weight-loss program utilizing a community-based participatory research approach. *Health Education and Behavior*, 35(5): 634-650.

*Kloek GC, van Lenthe FJ, van Nierop PWM, Koelen MA, Mackenbach JP (2006) Impact evaluation of a Dutch community intervention to improve health-related behaviour in deprived neighbourhoods. *Health and Place,* 12(4): 665-677.

*Kneipp SM, Kairalla JA, Lutz BJ, Pereira D, Hall AG, Flocks J, Beeber L, Schwartz T (2011) Public health nursing case management for women receiving temporary assistance for needy families: A randomized controlled trial using community-based participatory research. *American Journal of Public Health,* 101(9): 1759-1768.

*Kong AS, Sussman AL, Yahne C, Skipper BJ, Burge MR, Davis SM (2013) School-based health center intervention improves body mass index in overweight and obese adolescents. *Journal of Obesity,* http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3622307/ (accessed 27 January 2015).

*Kumpusalo E, Neittaanmaki L, Halonen P, Pekkarinen H, Penttila I, Parviainen M (1996) Finnish Healthy Village Study: Impact and outcomes of a low-cost local health promotion programme. *Health Promotion International,* 11(2): 105-115.

*Lassen A, Bruselius-Jensen M, Sommer HM, Thorsen AV, Trolle E (2007) Factors influencing participation rates and employees' attitudes toward promoting healthy eating at blue-collar worksites. *Health Education Research*, 22(5): 727-736.

*Lassen A, Thorsen AV, Sommer HM, Fagt S, Trolle E, Biltoft-Jensen A, Tetens I (2011) Improving the diet of employees at blue-collar worksites: Results from the 'Food at Work' intervention study. *Public Health Nutrition,* 14(6): 965-974.

*Lewis CE, Raczynski JM, Heath GW, Levinson R, Hilyer JC, Cutter GR (1993) Promoting physical activity in low-income African-American communities: The PARR project. *Ethnicity and Disease,* 3(2): 106-118.

*Lindenberg CS, Solorzano RM, Bear D, Strickland O, Galvis C, Pittman K (2002) Reducing substance use and risky sexual behavior among young, low-income, Mexican-American women: Comparison of two interventions. *Applied Nursing Research,* 15(3): 137-148.

*Lindqvist K, Timpka T, Schelp L, Ahlgren M (1999) Evaluation of an inter-organizational program for prevention of work-related injuries in a WHO Safe Community. *Work: A Journal of Prevention, Assessment and Rehabilitation,* 13(2): 89-96.

Lipsey M, Wilson D (2001) *Practical meta-analysis*. Thousand Oaks, CA: Sage.

*LoSciuto L, Hilbert SM, Fox MM, Porcellini L, Lanphear A (1999) A two-year evaluation of the Wood Rock Youth Development Project. *The Journal of Early Adolescence,* 19(4): 488-507.

*Macaulay AC, Paradis G, Potvin L, Cross EJ, Saad-Haddad C, McComber A, Desrosiers S, Kirby R, Montour LT, Lamping DL, Leduc N, Rivard M (1997) The Kahnawake Schools Diabetes Prevention Project: Intervention, evaluation, and baseline results of a diabetes primary prevention program with a native community in Canada. *Preventive Medicine,* 26(6): 779-790.

*Martin KS, Wu R, Wolff M, Colantonio AG, Grady J (2013) A novel food pantry program: Food security, self-sufficiency, and diet-quality outcomes. *American Journal of Preventive Medicine,* 45: 569-575.

*McAlister AL, Ramirez AG, Amezcua C, Pulley LV, Stern MP, Mercado S (1992) Smoking cessation in Texas-Mexico border communities: A quasi-experimental panel study. *American Journal of Health Promotion,* 6(4): 274-279.

*Mendoza JA, Levinger DD, Johnston BD (2009) Pilot evaluation of a walking school bus program in a low-income, urban community. *BMC Public Health,* 9(1): 122.

*Minkler M, Wallerstein N (eds) (2010) *Community-based participatory research for health: From process to outcomes*. New York: John Wiley and Sons.

*Nafziger AN, Erb TA, Jenkins PL, Lewis C, Pearson TA (2001) The Otsego-Schoharie healthy heart program: Prevention of cardiovascular disease in the rural US. *Scandinavian Journal of Public Health*, 29(56 suppl.): 21-32.

National Institute for Health and Care Excellence (2008) *Community engagement*. NICE Public Health Guidance 9. London: National Institute for Health and Care Excellence. http://www.nice.org.uk/guidance/ph9/resources/guidance-community-engagement-pdf (accessed 10 November 2014).

National Institute for Health and Care Excellence (2012) *Methods for the development of NICE public health guidance*. 3rd ed. London: National Institute for Health and Care Excellence. http://www.nice.org.uk/article/pmg4/resources/non-guidance-methods-for-the-development-of-nice-public-health-guidance-third-edition-pdf (accessed 10 November 2014).

*O'Loughlin JL, Paradis G, Gray-Donald K, Renaud L (1999) The impact of a community-based heart disease prevention program in a low-income, inner-city neighborhood. *American Journal of Public Health,* 89(12): 1819-1826.

O'Mara-Eves A, Brunton G, McDaid D, Oliver S, Kavanagh J, Jamal F, Matosevic T, Harden A, Thomas J (2013) Community engagement to reduce inequalities in health: A systematic review, meta-analysis and economic analysis. *Public Health Research*, 1(4).

Oliver SR, Rees RW, Clarke-Jones L, Milne R, Oakley AR, Gabbay J, Stein K, Buchanan P, Gyte G (2008) A multidimensional conceptual framework for analysing public involvement in health services research. *Health Expectations*, 11(1): 72-84.

*Parikh P, Simon E, Fei K, Looker H, Goytia C, Horowitz CR (2010) Results of a pilot diabetes prevention intervention in East Harlem, New York City: Project HEED. *American Journal of Public Health,* 100(Suppl. 1): S232-239.

*Perry CL, Williams CL, Veblen-Mortenson, S, Toomey TL, Komro KA, Anstine PS, McGovern PG, Finnegan JR, Forster JL, Wagenaar AC, Wolfson M (1996) Project Northland: Outcomes of a communitywide alcohol use prevention program during early adolescence. *American Journal of Public Health,* 86(7): 956-965.

*Phillips G, Bottomley C, Schmidt E, Tobi P, Lais S, Yu G, Lynch R, Lock K, Draper A, Moore D, Clow A, Petticrew M, Hayes R, Renton A (2014) Well London Phase-1: Results among adults of a cluster-randomised trial of a community engagement approach to improving health behaviours and mental well-being in deprived inner-city neighbourhoods. *Journal of Epidemiology and Community Health,* 68(7): 606-614.

*Plescia M, Herrick H, Chavis L (2008) Improving health behaviors in an African American community: The Charlotte Racial and Ethnic Approaches to Community Health project. *American Journal of Public Health,* 98(9): 1678-1684.

Popay J, Attree P, Hornby D, Milton B, Whitehead M, French B, Kowarzik U, Simpson N, Povall S (2007) Community engagement in initiatives addressing the wider social determinants of health: A rapid review of evidence on impact, experience and process. Lancaster: University of Lancaster, http://www.nice.org.uk/guidance/ph9/documents/social-determinants-evidence-review-final2 (accessed 11 November 2014).

Ragin CC (2006) Set relations in social research: Evaluating their consistency and coverage. *Political Analysis,* 14(3): 291-310.

Ragin, CC, Drass KA, Davey S (2006) *Fuzzy-set/qualitative comparative analysis 2.0.* Tucson, AZ: Department of Sociology, University of Arizona.

*Resnicow K, Cohn L, Reinhardt J, Cross D, Futterman R, Kirschner E, Wynder EL, Allegrante JP (1992) A three-year evaluation of the Know Your Body program in inner-city schoolchildren. *Health Education and Behavior,* 19(4): 463-480.

*Resnicow K, Kramish Campbell M, Carr C, McCarty F, Wang T, Periasamy S, Rahotep S, Doyle C, Williams A, Stables G (2004) Body and soul: A dietary intervention conducted through African-American churches. *American Journal of Preventive Medicine,* 27(2): 97-105.

*Rhodes SD, McCoy TP, Vissman AT, DiClemente RJ, Duck S, Hergenrather KC, Foley KL, Alonzo J, Bloom FR, Eng E (2011) A randomized controlled trial of a culturally congruent intervention to increase condom use and HIV testing among heterosexually active immigrant Latino men. *AIDS and Behavior,* 15(8): 1764-1775.

Rihoux B, Ragin CC (eds) (2008) *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques*. Los Angeles: Sage.

Ritchie J, Spencer L (1994) Qualitative data analysis for applied policy research. In Bryman A, Burgess RG (eds) *Analyzing qualitative data*. London: Routledge, pages 173–194.

*Robinson TN, Killen JD, Kraemer HC, Wilson DM, Matheson DM, Haskell WL, Pruitt LA, Powell TM, Owens AS, Thompson NS, Flint-Moore NM, Davis GJ, Emig KA, Brown RT, Rochon J, Green S, Varady A (2003) Dance and reducing television viewing to prevent weight gain in African-American girls: The Stanford GEMS pilot study. *Ethnicity and Disease,* 13(1, Suppl. 1): 65-77.

*Russell KM, Champion VL, Monahan PO, Millon-Underwood S, Zhao Q, Spacey N, Rush NL, Paskett ED (2010) Randomized trial of a lay health advisor and computer intervention to increase mammography screening in African American women. *Cancer Epidemiology, Biomarkers and Prevention,* 19(1): 201–210.

*Schensul JJ, Radda K, Coman E, Vazquez E (2009) Multi-level intervention to prevent influenza infections in older low income and minority adults. *American Journal of Community Psychology,* 43(3-4): 313-329.

*Schinke SP, Tepavac, L, Cole KC (2000) Preventing substance use among Native American youth: Three-year results. *Addictive Behaviors,* 25(3): 387-397.

*Schorling JB, Roach, J, Siegel, M, Baturka, N, Hunt DE, Guterbock TM, Stewart HL (1997) A trial of church-based smoking cessation interventions for rural African Americans. *Preventive Medicine,* 26(1): 92-101.

*Schwarz DF, Grisso JA, Miles C, Holmes JH, Sutton RL (1993) An injury prevention program in an urban African-American community. *American Journal of Public Health,* 83(5): 675-680.

*Secker-Walker RH, Flynn BS, Solomon LJ, Skelly JM, Dorwaldt AL, Ashikaga T (2000) Helping women quit smoking: Results of a community intervention program. *American Journal of Public Health,* 90(6): 940-946.

*Segal SP, Silverman CJ, Temkin TL (2011) Outcomes from consumer-operated and community mental health services: A randomized controlled trial. *Psychiatric Services,* 62(8): 915-921.

*Shelley D, Fahs M, Yerneni R, Das D, Nguyen N, Hung D, Burton D, Chin M, Chang MD, Cummings KM (2008) Effectiveness of tobacco control among Chinese Americans: A comparative analysis of policy approaches versus community-based programs. *Preventive Medicine,* 47(5): 530-536.

Sheridan K, Tobi P (2010) Towards a Community Engagement strategy. *British Journal of Healthcare Management* 16(3):123-128 . http://www.bjhcm.co.uk/cgibin/go.pl/library/article.cgi?uid=46819;article=BJHCM_16_3_123_128;format=pdf. Accessed: 6 June 2014.

Sheridan K, Adams-Eaton F, Trimble A, Renton A, Bertotti M (2011) Community Engagement using World Café: The Well London Experience Groupwork and Wellbeing, 20(3): 32-50 http://www.scie-socialcareonline.org.uk/profile.asp?guid=3517bea3-0be6-4b88-81a2-e01fec64a1ec. Accessed: 6 June 2014.

*Simmons D, Fleming C, Voyle J, Fou F, Feo S, Gatland B (1998) A pilot urban church-based programme to reduce risk factors for diabetes among Western Samoans in New Zealand. *Diabetic Medicine,* 15(2): 136-142.

*Sussman AL, Montoya C, Werder O, Davis S, Wallerstein N, Kong AS (2013) An adaptive CBPR approach to create weight management materials for a school-based health center intervention. *Journal of Obesity*, 2013, http://www.hindawi.com/journals/jobe/2013/978482/ (accessed 27 January 2015).

Thomas J, Brunton J, Graziosi S (2010) *EPPI-Reviewer 4.0: Software for research synthesis.* EPPI-Centre Software. London: Social Science Research Unit, Institute of Education, University of London.

Thomas J, Harden A, Newman M (2012) Synthesis: Combining results statistically and appropriately. In Gough D, Oliver S, Thomas J (2012) *An introduction to systematic reviews*. London: Sage.

Thomas J, O'Mara-Eves A, Brunton G (2014) Using qualitative comparative analysis (QCA) in systematic reviews of complex interventions: A worked example. *Systematic Reviews*, 3:67, www.systematicreviewsjournal.com/content/3/1/67 (accessed 17 October 2014).

Thompson S, Sharp S (1999) Explaining heterogeneity in meta-analysis: A comparison of methods. *Statistics in Medicine*, 18: 693–708.

Wallerstein N, Duran B (2010) Community-based participatory research contributions to intervention research: The intersection of science and practice to improve health equity. *American Journal of Public Health,* 100(Suppl. 1): S40-S46.

Wallerstein N, Oetzel J, Duran B, Tafoya G, Belone L, Rae R (2008) What predicts outcomes in CBPR? In Minkler M and Wallerstein N (eds) *Community-based participatory research: From processes to outcomes*. San Francisco: John Wiley and Sons.

Wermert AM, Mehl A, Opalek JM, Shaffer LE (2012) Implementation and early outcomes of a peer-led traffic safety initiative for high school students. *Journal of Trauma Nursing*, 19(2): 94-101.

White I, Thomas J (2005) Standardised mean differences in individually-randomised and cluster randomised trials, with applications to meta-analysis. *Clinical Trials* 2: 141–151.

*Winkleby MA, Feighery E, Dunn M, Kole S, Ahn D, Killen JD (2004) Effects of an advocacy intervention to reduce smoking among teenagers. *Archives of Pediatrics and Adolescent Medicine,* 158(3): 269-275.

*Woods G, Levinson AH, Jones G, Kennedy RL, Johnson LC, Tran ZV, Gonzalez T, Marcus AC (2013) The Living Well by Faith health and wellness program for African Americans: An exemplar of community-based participatory research. *Ethnicity and Disease,* 23(2): 223-229.

*Wright AL, Naylor A, Wester R, Bauer M, Sutcliffe E (1997) Using cultural knowledge in health promotion: Breastfeeding among the Navajo. *Health Education and Behavior,* 24(5): 625-639.

*Wright K, Giger JN, Norris K, Suro Z (2013) Impact of a nurse-directed, coordinated school health program to enhance physical activity behaviors and reduce body mass index among minority children: A parallel-group, randomized control trial. *International Journal of Nursing Studies,* 50: 727-737.

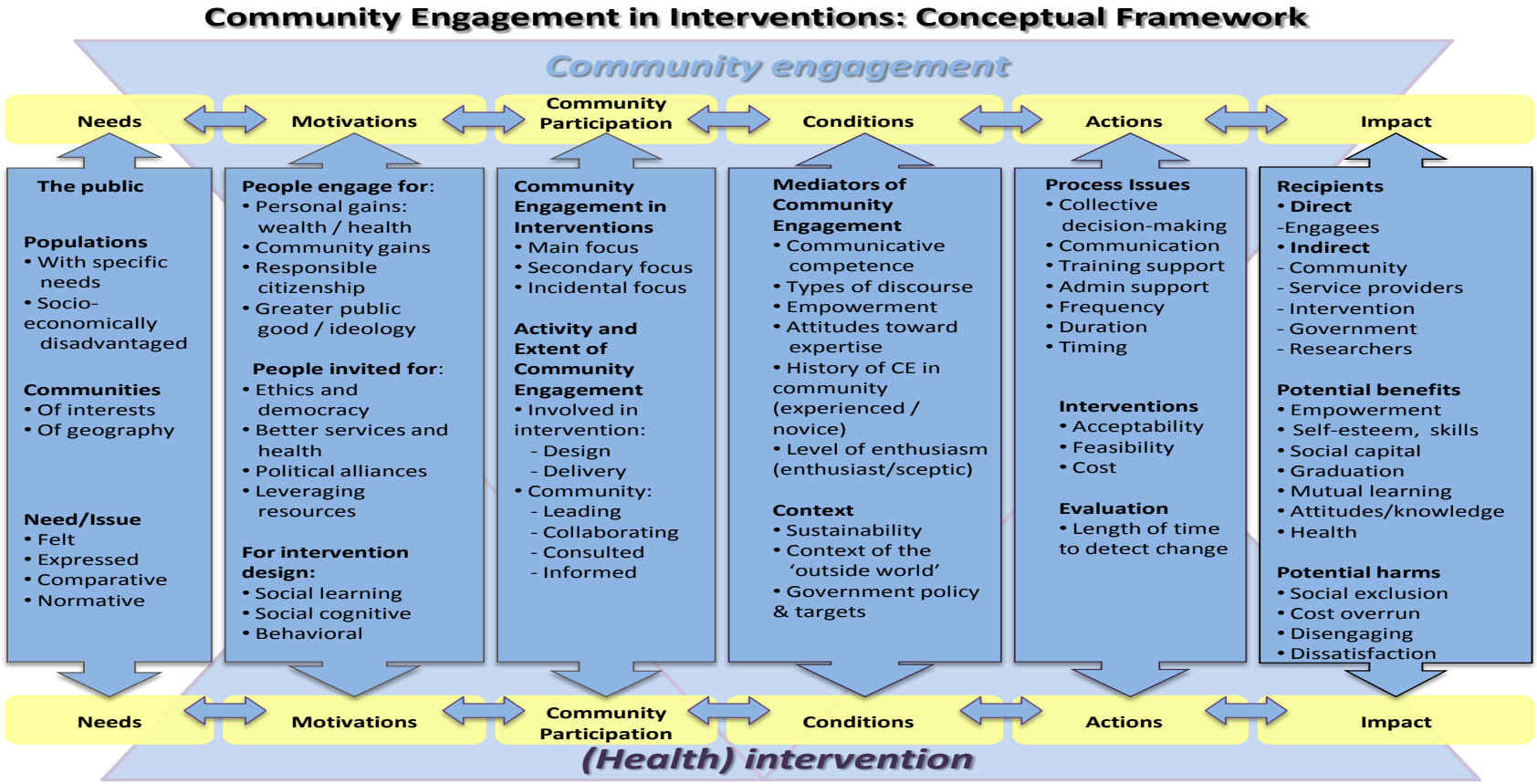
*Yanek LR, Becker DM, Moy TF, Gittelsohn, J, Koffman DM (2001) Project Joy: Faith based cardiovascular health promotion for African American women. *Public Health Reports,* 116(Suppl. 1): 68-81.

*Zhou F, Euler GL, McPhee SJ, Nguyen T, Lam T, Wong C, Mock J (2003) Economic analysis of promotion of hepatitis B vaccinations among Vietnamese-American children and adolescents in Houston and Dallas. *Pediatrics,* 11(6): 1289-1296.

*Zoellner J, Hill JL, Grier K, Chau C, Kopec D, Price B, Dunn C (2013) Randomized controlled trial targeting obesity-related behaviors: Better Together Healthy Caswell County. *Preventing Chronic Disease,* 10: E96.

# Appendices

## Appendix 1: Conceptual framework



First produced in 2015 by:
Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre)
Social Science Research Unit
UCL Institute of Education, University College London
18 Woburn Square
London WC1H 0NR
Tel: +44 (0)20 7612 6397
http://eppi.ioe.ac.uk/
http://www.ioe.ac.uk/ssru/
ISBN: 978-1-907345-78-4

The Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) is part of the Social Science Research Unit (SSRU), UCL Institute of Education, University College London.

The EPPI-Centre was established in 1993 to address the need for a systematic approach to the organisation and review of evidence-based work on social interventions. The work and publications of the Centre engage health and education policy makers, practitioners and service users in discussions about how researchers can make their work more relevant and how to use research findings.

Founded in 1990, the Social Science Research Unit (SSRU) is based at the UCL Institute of Education, University College London. Our mission is to engage in and otherwise promote rigorous, ethical and participative social research as well as to support evidence-informed public policy and practice across a range of domains including education, health and welfare, guided by a concern for human rights, social justice and the development of human potential.



This document is available in a range of accessible formats including large print. Please contact the UCL Institute of Education for assistance:

telephone: +44 (0)20 7947 9556 email: info@ioe.ac.uk