Testing methodological developments in the conduct of narrative synthesis: a demonstration review of research on the implementation of smoke alarm interventions

Lisa Arai,[1] Nicky Britten, Jennie Popay, Helen Roberts,[1] Mark Petticrew, Mark Rodgers and Amanda Sowden

English In the context of systematic reviews, statistical meta-analysis of findings is not always possible. Where this is the case, or where a review of implementation evidence is required, narrative synthesis of data is typically undertaken. Drawing on recently developed guidance aimed at those undertaking data synthesis – and information on the implementation of domestic smoke detectors – we present findings from a demonstration of the tools and techniques that can be used in a narrative synthesis. The work demonstrates how this process can be made more transparent, and suggests that using the tools and techniques can improve the quality of narrative synthesis.

Français Dans le contexte de revues systématiques, une méta-analyse des conclusions n'est pas toujours possible. Lorsque c'est le cas, ou lorsqu'une revue de la mise en œuvre est nécessaire, on fait généralement une synthèse narrative des données. Sur la base de directives récemment élaborées ciblant les personnes qui entreprennent la synthèse des données – et l'information sur la mise en œuvre des détecteurs de fumée – nous présentons les conclusions d'une démonstration des outils et des techniques qui peuvent être utilisés dans une synthèse narrative. Le travail démontre comment on peut rendre ce processus plus transparent, et suggère que l'utilisation des outils et des techniques peut améliorer la qualité de la synthèse narrative.

Español En el contexto de análisis sistemático, los meta análisis estadísticos de resultados no son siempre posibles. Ya sea este el caso o si se requiere un análisis de evidencia de implementación, la síntesis de datos es típicamente emprendida. Basándonos en una guía desarrollada recientemente apuntada a la síntesis de datos emprendida - y en la información de la implementación de los detectores de humo domésticos- presentamos resultados de una demostración de las herramientas y técnicas que se pueden usar en una síntesis narrativa. El trabajo demuestra cómo este proceso se puede hacer más transparente y sugiere que usando las herramientas y técnicas se puede mejorar la calidad de la síntesis narrativa.

Key words narrative • synthesis • implementation • methodology

Introduction

There is widespread agreement that systematic reviews provide the best method we have to date for understanding whether an intervention or programme 'works'. However, it is also generally accepted that these reviews typically do not provide a clear recipe for turning a body of research into recommendations for policy or practice. Sometimes this will be for practical reasons. For example, there may be strong evidence to support the effectiveness of a certain reading scheme, but there may be insufficient knowledge and training opportunities for practitioners to roll it out. Sometimes it will be for cost-benefit reasons: the intervention may work, but the cost may be so high that no society is likely to accept it. Or there may be ethical problems: publishing the addresses of convicted paedophiles might constrain their activities, but infringe their human rights or have other negative consequences, such as driving them underground (Pawson, 2002).

In other situations the barrier to knowledge transfer may be the trustworthiness of the synthesis itself. This may be because a statistical meta-analysis is not possible where, for example, the studies included in the review are too diverse. Even in situations where a robust meta-analysis demonstrates that an intervention is effective, evidence on how best to implement it is rarely reviewed at the same time.

In both these cases – where a statistical meta-analysis of effectiveness data is not possible or where a review of implementation evidence is required – a narrative approach to synthesising the findings of multiple studies is typically undertaken. But unlike statistical approaches to synthesis, the processes for narrative synthesis (NS) have not been transparent to date, and rely on a high level of trust in assumptions made by the authors.

This article draws on work funded by the UK Economic and Social Research Council (ESRC) that aimed to produce and test guidance for those conducting NS. The full background to this guidance is described elsewhere (Popay et al, 2006). Here we describe a practical example of the application of the guidance to a body of evidence on factors influencing the implementation of interventions to improve the use/functioning of domestic smoke alarms. This is potentially an important intervention, given the steep social class gradient in deaths from house fires (Audit Commission, 2007). A second demonstration of the application of the guidance to a body of evidence on the effectiveness of these interventions is reported elsewhere (Rodgers et al, 2007).

It is important to emphasise that the results of these two demonstration syntheses are not intended to make a substantive contribution to the evidence base on smoke alarm interventions. Our aims are largely methodological. For this reason the demonstration work did not involve all stages of a systematic review – indeed a good-quality review about the effects of smoke alarms already exists (DiGuiseppi and Higgins, 2000). We focus only on the synthesis stage. The specific objectives of the methodological work reported here were to:

- illustrate the decision-making processes involved in the conduct of a NS;
- identify factors that can inform choices about the use of particular tools and techniques during synthesis;

- provide examples of how particular tools and techniques can be used in the synthesis of evidence on the implementation of specific interventions; and
- demonstrate the type of outcomes achieved by NS and highlight some of the limitations of this approach to synthesis.

The article is divided into four sections. First, we briefly consider the nature of evidence on implementation before describing the methods we used in the demonstration synthesis. We then describe the NS process, providing both an appraisal of the specific tools and techniques used as well as reporting on the substantive findings. Finally, we reflect on the implications of the work for the future conduct of NS.

The nature of implementation evidence

Effectiveness studies are primarily concerned with the impact of an intervention.[2] In contrast, the primary aim of implementation studies is to consider how or why interventions have particular impacts, including what went wrong when interventions did not have the anticipated impact (van Meter and van Horn, 1975). These studies focus on how factors/processes, operating at the level of systems (which might include international, national, regional or local level systems, depending on the intervention), organisations and individuals, impinge on the implementation of an intervention.

There is great diversity in the design of implementation studies. They may be cross-sectional or longitudinal, and they may stand alone or be embedded in impact studies, including trials. Such studies commonly involve multiple methods and may involve routine data on the 'reach' of the intervention, new surveys or other methods generating quantitative data and any of a range of qualitative methods including in-depth interviews producing narrative data, participant and non-participant observation, focus groups, draw and write techniques for children and ethnographic fieldwork. Moreover, process studies may be summative and/or formative, providing iterative feedback, as an evaluation proceeds, to those involved in implementation in order to contribute to the further development of an intervention. For these reasons, reviewers will almost certainly have to deal with considerable heterogeneity in the design of implementation studies and the type of data/findings they produce.

All implementation studies have in common the aim of providing a description of the intervention being evaluated, but the nature of this description can vary. Denzin and Lincoln (1994) have drawn a distinction between 'thin' and 'thicker' descriptions in qualitative research, arguing that 'thick' description has greater explanatory potential. Providing a 'thin' description of the range of factors and/or processes that may have impacted on the implementation of a particular intervention is relatively straightforward. In contrast, developing a credible explanation of how and why a particular intervention has an effect is much more challenging. In addition, the basis for making explanatory claims in quantitative research is different from that used in qualitative research. As Brock (2003) has argued, part of the problem is that the range of factors that can impinge on a particular intervention is very wide and in many studies the number of cases studied is relatively few. In quantitative research these factors may be referred to as 'moderators' and/or 'mediators'.

In qualitative research, the potential for generalising from a single study arises from the extent to which researchers link their analysis to wider theoretical understandings. For example, Hilary Graham (1993) linked her analysis of qualitative data on white working-class women's smoking behaviour to theories about coping behaviour, arguing that smoking provides a coping mechanism in difficult life circumstances. While this provides the basis for generalising to other white working-class women, Graham's later work (Graham and Blackburn, 1998) suggests that the findings do not apply across ethnic groups.

On the basis of his review of the literature on implementation studies, Brock (2003) concludes that the field as a whole lacks cohesion, with no consensus having been reached over what a good-quality implementation study should look like. Consequently, with regard to systematic reviews, implementation research raises at least three methodological challenges: how best to search for relevant studies and what inclusion criteria to use; how to assess the methodological quality and therefore the reliability of studies using multiple methods; and how to approach the synthesis of findings from what may be very heterogeneous studies.

This article describes methodological work on a potentially useful approach to the synthesis of findings from multiple implementation studies, which we term 'narrative synthesis'. This approach relies primarily on the use of words and text to summarise and explain the findings of the synthesis (Popay et al, 2006). Given that it is focused on only the synthesis stage of a systematic review, the approach does not address the other methodological questions highlighted: how to search for and quality appraise mixed method implementation studies.

Method

Identifying studies

Most of the studies included in the NS were drawn from an earlier exploratory review of evidence on the implementation of interventions aiming to reduce accidents among children and young people (Popay et al, 2003; Arai et al, 2005; Roberts et al, 2006; Roen et al, 2006). This earlier review included primary studies identified in a Cochrane review (DiGuiseppi and Higgins, 2000) as well as new studies found via a comprehensive search strategy. The demonstration synthesis was restricted to evidence on the implementation of interventions aiming to improve the uptake and functioning of domestic smoke alarms. A limited search was also undertaken to identify any new studies published since the original review was undertaken. Time and resources did not permit a synthesis of all of the studies identified so a purposive sample was taken on the basis of the 'thickness' of the data on factors influencing implementation (see Roen et al, 2006, for more detail). As a result of this process, two members of the review team conducted a NS based on seven papers. Details of these seven papers, and the interventions reported on, are shown below in Table 1.

Table 1: The papers included in the NS of evidence on the implementation of programmes aiming to improve uptake/functioning of domestic smoke alarms

Author and year	Location	Type of intervention
Campbell DeLong Resources Inc (2003)	Oregon, US	Fire safety awareness campaign (aimed primarily at landlords)
Camit (1998)	New South Wales, Australia	Provision of discounted smoke alarms, plus written fire safety information and alarm demonstration
Camit (2002)	New South Wales, Australia	Provision of discounted smoke alarms, plus written fire safety information and alarm demonstration
DiGuiseppi et al (1999)	London, UK	Provision and installation of free smoke alarms, plus written fire safety information
McConnell et al (1996)	Memphis, US	Fire safety education, including use of video, written and other material
Roberts et al (2004)	London, UK	Qualitative work with children and adults relating to provision and installation of smoke alarms, plus written fire safety information
Young et al (1999)	New South Wales, Australia	Provision of discounted smoke alarms, plus written fire safety information and alarm demonstration

The main elements of NS

The guidance used for this demonstration (Popay et al, 2006) identifies four elements to the process of NS:

- developing a theory of how the intervention works, why and for whom;
- developing a preliminary synthesis;
- exploring relationships within and between studies;
- assessing the robustness of the synthesis.

For each element of the synthesis process the guidance describes a number of tools and techniques that may be useful. These were identified through a systematic search of relevant methodological papers, books, guidelines and websites. The choice of specific tools and techniques at each point in the process will depend on the type of data being synthesised.

Some approaches to evidence synthesis – for example, realistic synthesis – focus entirely on testing theories of how an intervention works, why, for whom and in what circumstances (Pawson, 2006; Petticrew and Roberts, 2006;). Weiss calls this a 'theory of change' (Weiss, 1997). In systematic reviews, theory development and testing would be one element of the review process as a whole and, strictly speaking, should be undertaken early in a review before the synthesis proper begins. In the guidance we suggest that this process can inform decisions about the review question and the types of studies to include. In terms of the NS, a 'theory of change' can also contribute to the interpretation of the review's findings and will be valuable in assessing how widely applicable those findings may be. We included theory development and testing in the NS framework as a result of reviewers' comments on a full draft of the NS guidance – too late for it to be included in the demonstration review reported here. For this reason, the demonstration review focused only on the latter three elements of the NS framework: the preliminary synthesis; exploring relationships; and assessing robustness.

The NS was carried out separately by two members of the study team, each consciously attempting to test the usefulness and clarity of the guidance on the conduct of NS. The tools and techniques that appeared to be useful and relevant to each element of the synthesis were selected. Two principal factors determined which tools and techniques were used: the preferences and judgement of those synthesising the data (based largely on their familiarity with, and previous experience of, specific tools and techniques) and, as noted above, the nature of the data itself. The results of this exercise are shown in Table 2, which distinguishes between those tools and techniques that were used in the worked examples (Yes), not used and probably not relevant (No), or not used but potentially relevant (Potentially yes). Only tools and techniques actually used in the demonstration synthesis are described in this article. Any individual synthesis is unlikely to use the whole portfolio of tools and techniques, and fuller descriptions of all of these are provided in the guidance (Popay et al, 2006).

Extracting data from included studies

Once the studies to be included had been selected, we had to extract relevant data. However, as we discovered in our earlier review, the process of extracting data on implementation raises a number of issues for reviewers. These are:

- locating the data in the text;
- establishing the nature and type of implementation data;
- ascertaining their provenance and reliability;
- extracting the data in preparation for analysis.

In relation to the first, it should be noted that implementation is rarely the focus of published reports of interventions, particularly those published in high-impact peer-reviewed journals where the focus is much more likely to be on the effectiveness of interventions, and where brevity is an editorial requirement. The same point can

Table 2: Tools and techniques for the synthesis of implementation data

Tools and techniques	Use for NS of implementation data?
Preliminary synthesis	
Textual descriptions	Yes
Tabulation	Yes
Groupings and clusterings	Yes
Constructing a common rubric	No
Thematic analysis	Yes
Content analysis	No
Vote counting	No
Exploring relationships	
Variability in outcomes	Potentially yes
Variability in study design	Potentially yes
Variability in study population	Potentially yes
Moderator variables and subgroup analyses	No
Idea webbing and subgroup analyses	Yes
Conceptual triangulation	Potentially yes
Translation	Yes
Case descriptions	Potentially yes
Visual representation	No
Investigator triangulation and methodological triangulation	Potentially yes
Assessing robustness	
Weight of evidence	Potentially yes
Best evidence synthesis	Potentially yes
Checking with authors	Potentially yes
Critical reflection	Yes

be made about descriptions of the intervention itself. These are also rarely available in detail in peer-reviewed journals, which limits comparison of the effectiveness of different types of interventions. Occasionally, additional documents are referred to which describe the intervention, and its implementation. Commonly this is not the case but some implementation data may be found interspersed throughout the text or, more frequently, they are located predominantly in the discussion section, where authors provide an explanation for the effectiveness, or otherwise, of an intervention.

Once these data have been identified, reviewers need to establish their nature and type. Most data of this kind consist of simple narrative observations made by the authors, on the factors affecting the implementation of an intervention. It appears on the basis of our experience to be relatively rare for primary data (direct quotations

from people targeted by interventions, for instance) to be reported to support these observations.

Ascertaining the provenance and reliability of data on implementation can be particularly problematic. We found that it is rare to find details of where these data (either in the form of narrative author observations or quotations) come from, or on what they are based. It is often unclear whether authors are drawing on fieldworkers' observations or on unreported data from interviews, or simply providing 'informed hunches'. Where data are 'thicker', providing more detail and context (perhaps where qualitative data have been presented), additional unreported data on implementation may be available from the authors. These issues have obvious implications for methodological quality appraisal. In our experience there is either insufficient information available on study design or methods for a judgement about the quality of implementation data to be made, or it is clear that the quality is poor.

Finally, how should those attempting a synthesis extract the data in preparation for analysis? Where authors have made simple observations about implementation, these can be taken whole from the text (and the presentation/reporting of this will depend on the number and complexity of such observations). Quotations or other qualitative data can also be taken directly from the text and the data entered into tables with appropriate headings (for example: author/paper; location of implementation data in the report; nature and scope of data). In some circumstances, where a full reporting of an implementation study has been included, the data to be extracted may take the form of concepts derived from a full analysis of qualitative data or equivalent results from the analysis of quantitative survey or routine data.

The process and product of the NS

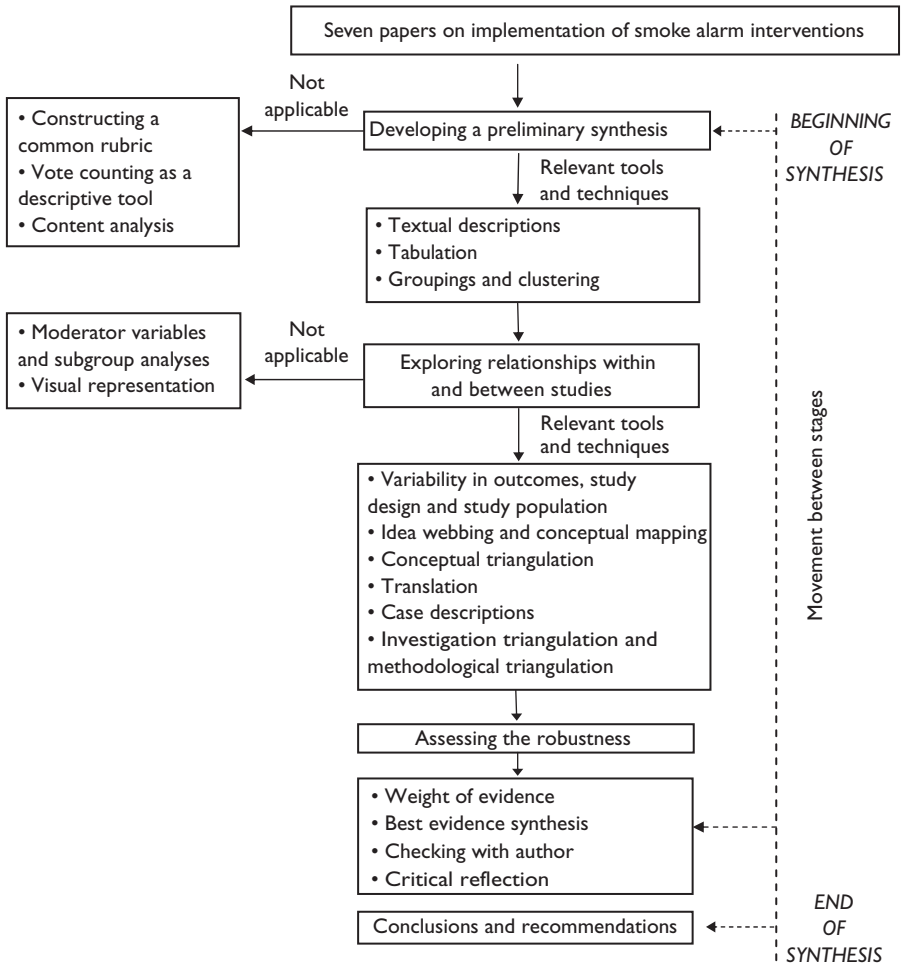
In this section we describe each of the three elements of the NS of evidence on the implementation of interventions aiming to increase the uptake and maintenance of domestic smoke alarms – the development of a preliminary synthesis, exploration of relationships within and between studies, and the assessment of the robustness of the synthesis results – illustrating in detail the application of various tools and techniques. A flow chart summarising the synthesis process as a whole is presented in Figure 1. Some of the applications have been limited by the relatively small evidence base for the topic selected, as a number of the tools and techniques listed below require a larger body of evidence.

Developing a preliminary synthesis

Textual description

This tool was used by both reviewers at an early stage in the synthesis process in order to summarise the papers and begin to extract information in a systematic way. Textual descriptions offer the potential to include more detail than, for example, tabulations. However, as the following two examples illustrate, textual descriptions can vary in

Figure 1: Synthesis process



the way information is presented – with the first reviewer, for example, highlighting key words. The type and amount of information included in the two examples also varies. Reviewers should develop a standard format to ensure consistency of reporting in these types of descriptions.

Example 1:
In McConnell et al (1996), the target population was new heads of households in public housing residences of the Memphis Housing Authority (MHA) in the US and they were predominantly female African-Americans living with children. The MHA policy is to ensure that a functioning smoke detector is located in every unit when rented, but a spot check of 325 units in 1992 found that less than 8% had a

working smoke detector. The 35-minute intervention (delivered during mandatory orientation sessions for new MHA heads of household) consisted of the following components: a pre-test; a videotape accompanied by brief presentations delivered by one of 36 uniformed fire fighters, one MHA supervisor or one civilian educator; a contract of behaviour between tenant and housing authority; a post-test; and a fire-safety reminder card. The outcomes were fire incidence data (after possibly 15 months, timescale not clear); residents' evaluations of the programme; changes in their fire safety knowledge; and their commitments to fire safety behaviours. The method of evaluation was an uncontrolled comparison between trained and untrained residents, using contemporary and historical comparison groups. The evaluation data were all quantitative. The results showed a lower incidence of fires in trained residents compared with untrained residents (one fire for every 4,312 renter months in trained residents compared with one fire for every 780 renter months in untrained residents; a relative risk of 5.5). Comparing trained residents with untrained residents over the nine-year baseline period gave a relative risk of 4.8. Comparisons between newer and older residents from the MHA records suggested that newer residents were more likely to experience fires, thus countering the suggestion that the results can be explained by the fact that the trained residents were also new residents. No data were provided on the proportion of working smoke detectors post intervention.

Example 2:
Young et al (1999) and Camit (1998, 2002) report on the effectiveness and implementation of a smoke alarm promotion campaign in New South Wales, Australia, oriented to the needs of Arabic, Chinese and Vietnamese communities. Qualitative data were collected in focus groups and interviews. Survey data were also collected. Their main observations in relation to implementation were that, among the target community, there was a lack of awareness of the need for smoke alarms. Living in rented property where the landlord was thought to be unsympathetic to the need for a smoke alarm also created barriers to the installation of smoke alarms.

Tabulation

Both reviewers produced tables, but both felt that tabulation and textual descriptions essentially provided very similar data, possibly using the same headings but laid out differently. In a table, however, it is easier to compare data across different studies (for example, see Table 3). Conversely, tables are not the best medium for describing implementation data because these are often in a narrative form and therefore take up more space than numerical data.

Groupings and clusters

Organising the included studies into groups can assist in the development of a preliminary synthesis as well as begin to identify patterns within and across groups – the second element of a NS (see Table 4). This technique is particularly useful when there are larger numbers of papers. The type of groups identified is likely to depend

Table 3: Example of tabulation

Author and year	Location and setting	Target population	Method	Main findings
Roberts et al (2004)	London, UK Urban	58 adults and 41 children in community (in the qualitative study) 2,145 households exposed to intervention	Focus groups and interviews	Problems with smoke alarms (sensitivity, false alarming) identified as major barriers to implementation
Camit (1998, 2002); Young et al (1999)	NSW, Australia Mixed	Chinese, Vietnamese, Arabic-speaking (numbers not given)	Focus groups	Implementation successful using multifaceted, language-appropriate approach
Campbell DeLong Resources Inc (2003)	Oregon, US Mixed	All residents, but focus on Latino-speaking (sample population varied according to element of intervention)	Interviews	Successful implementation heavily dependent on landlords' attitudes to smoke alarms

Table 4: Example of grouping

Grouping according to:		
(1) Location	**(2) Focus of report**	**(3) Population**
UK: DiGiuseppi et al (1999); Roberts et al (2004)	Broad, general factors affecting programme: DiGiuseppi et al (1999); Camit (1998, 2002)	Ethnically mixed: Camit (1998, 2002); Young et al (1999); Roberts et al (2004)
US: Campbell DeLong Inc (2003); McConnell et al (1996)	Individual factors affecting programme: McConnell et al (1996); Young et al (1999); Campbell DeLong Inc (2003); Roberts et al (2004)	Ethnically mixed and low income: Campbell DeLong Inc (2003); McConnell et al (1996); DiGiuseppi et al (1999); Roberts et al (2004)
Australia: Camit (1998, 2002); Young et al (1999)		

on the reviewers' categorisations, but may also depend on the study type (whether it is describing a trial or a campaign, for example). The result is similar to tabulation, except that the relationships (or differences) between groups of studies can be made more explicit when they are clustered in this way.

Translating data: thematic analysis

Thematic analysis is a common technique for the analysis of qualitative data in primary research. In a synthesis, it can be used to identify systematically recurrent or salient themes or concepts across a number of studies. In the demonstration synthesis, different reviewers, and the same reviewer at different times, identified different themes, on an inductive basis, by reading and rereading the papers. They fell into two categories: those themes identified by the authors from their qualitative data; and aspects of the interventions that seemed to act as barriers or facilitators, in the view of the authors or the reviewer. One example of each is given below.

Example 1:
This thematic analysis was originally presented as a table, which set out the themes developed on the basis of the evidence extracted from the included studies under three main headings: the smoke alarm, the individual and the community. The themes developed under each heading can be summarised as follows.

The smoke alarm:

- The design of smoke alarms has serious implications for their use.

The individual:

- Individuals perceive themselves to be invulnerable to fire.
- There is a general lack of awareness of fire safety and of the importance of smoke alarms.
- Cultural factors can have an effect on smoke alarm use and functioning.
- Factors such as age and poverty can affect smoke alarm installation.

The community:

- Landlords play a role in promoting/hindering good fire safety awareness.
- Neighbours', and others', attitudes affect smoke alarm use.
- Type of tenure affects smoke alarm use.
- The success of campaigns to promote installation of smoke alarms can be dependent on community characteristics

Example 2:
Differences between reviewers do not suggest that the synthesis is flawed, but rather draw attention to different ways of interpreting the same data. Both reviewers identified a typology including barriers and levers (Table 5), as did some of the study authors. Whether the data are seen ecologically or in stages, the idea of barriers and levers is common to both. In a final synthesis, specific factors that act as barriers/levers, the notion of stages (temporality) and the organisation of these factors within domains at different levels (ecological perspective) could be brought together.

Table 5: Barriers and levers to use/functioning of smoke alarms

	Barriers	Levers
(1) Barriers and levers to acquisition of smoke alarms		
General	• Problems accessing communities • Suspicion of 'authority' or local government	• Gaining trust of key community 'players' and leaders • Emphasising separation from distrusted authority/forming alliance with trusted partners
Specific to smoke alarm campaigns	• Lack of awareness of benefits of smoke alarms • Perceived cost of smoke alarms • Perception that household is not at risk of fire (due to type of house/characteristics of household members)	• Running well-coordinated, culturally appropriate awareness campaign • Giveaway/availability of reduced-price alarms • Awareness campaign
(2) Barriers and levers to installation of smoke alarms		
General	• Anxiety about damage to property	• Landlord approval/permission for installation, or landlord example of installation
Specific to smoke alarm campaigns	• Inability/unwillingness to install alarm • Anxiety about letting installer into the house	• Installation of alarm by project worker endorsed by community
(3) Barriers and levers to continued use of smoke alarms		
Specific to smoke alarm campaigns	• False alarms • Problems with maintenance	• Education about triggers for false alarms • Reinstallation of alarm • Project workers offer to maintain alarms • Education about maintenance • Development of better alarms

Exploring relationships within and between studies

Exploring the influence of heterogeneity

One reviewer attempted to explore the influence of heterogeneity in the included studies, and focused attention on the characteristics of the different studies and their potential relationships to the findings. The other reviewer did not find this technique as useful as some of the others. While there are several possible sources of heterogeneity, in the demonstration synthesis we focused on variability in the nature of the interventions.

Example of variability in interventions:
The interventions were different: an educational intervention for residents with already installed smoke alarms (McConnell et al, 1996); the sale of smoke alarms (Young et al, 1999); and the free installation of smoke alarms (Roberts et al, 2004,). However, one aspect of the interventions is potentially explanatory in relation to their outcomes. The intervention in the McConnell study was drawn up following focus

groups and interviews with individual residents in which they were asked about the best approaches to be used. Although the research team evaluating the intervention linked to the qualitative study reported by Roberts and colleagues had invested considerable time in ensuring community involvement, Roberts et al concluded that take-up and maintenance of smoke alarms might have been improved if a qualitative study of people's perceptions of fire risk had been conducted prior to and informed the intervention. The Young study was not drawn up in consultation with the local community. Thus the involvement of the local population and the development of a tailored intervention may well have influenced the success of the McConnell study.

Idea webbing and conceptual mapping

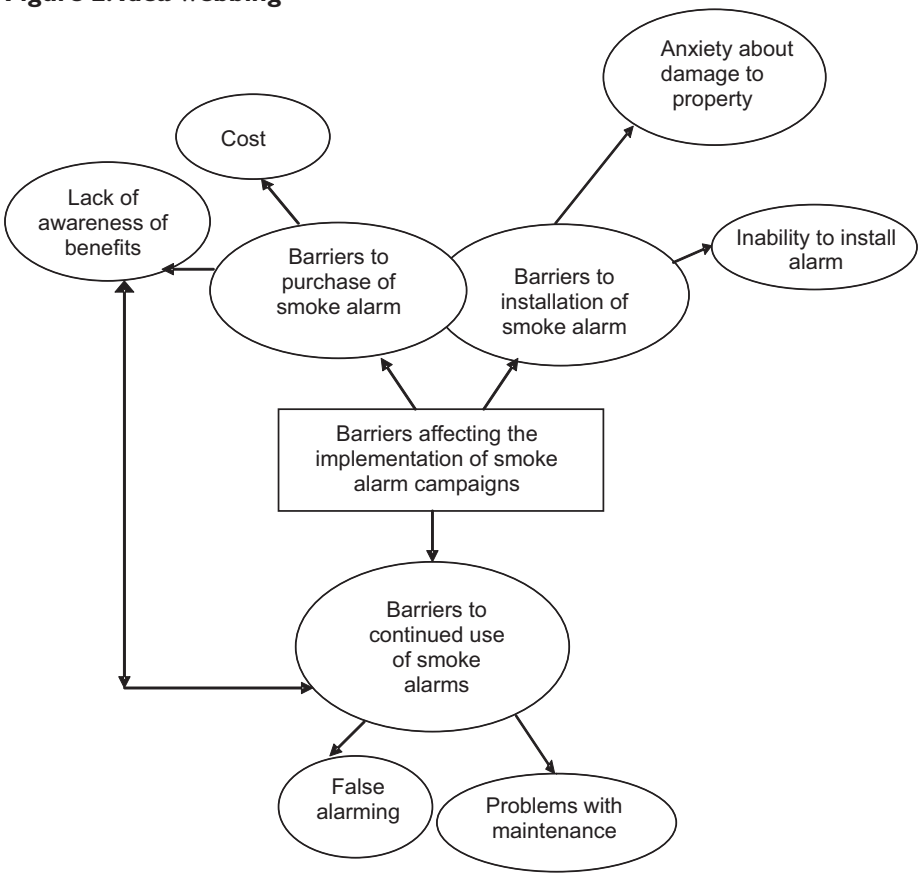
These techniques were used interchangeably by both reviewers (see Figure 2 for an example of one reviewer's idea web and Figure 3 for an example of a conceptual map). The use of figures and diagrams was found to be helpful for exploring potential relationships between different aspects of the implementation of interventions. While the diagrams may help reviewers in structuring their thoughts, their value to others will depend on how complex and detailed the figures are, and how easy they are to 'read'. A clear and accessible diagram may be useful as a preliminary product of the synthesis by, for example, linking themes identified in a thematic analysis. In the demonstration synthesis, the use of the figures allowed the specific aspects of different interventions to be mapped onto one another in a way that assisted the reviewers in thinking about how they might be related.

Translation as an approach to exploring relationships

Translation as a process for synthesis is typically associated with the work of Noblit and Hare (1988) on meta-ethnography. In this approach, the translation of studies into one another aims to maintain the central concepts within each study in their relation to other key concepts in that study. The question for the reviewer is whether a concept or theme in one study represents the same concept or theme in another, even if they have been labelled differently. In the demonstration synthesis, one reviewer with experience of meta-ethnography felt that translation was the most useful technique for exploring relationships within and/or between study findings.

Example of reciprocal translation:
Three concepts seemed to offer themselves for translation across studies. The first is landlord commitment/lack of commitment. The difficulties with landlords discussed in the Young paper seem to be the exact opposite of the commitment demonstrated by the MHA in the McConnell et al (1996) study, but not explicitly commented on by the authors. The second concept is risk perception: feeling oneself at high risk (Roberts et al, 2004), or underestimating the risk of fire (Young et al, 1999) (although non-adopters of smoke alarms in the Roberts study considered fire a high risk, but tended to balance this against it being less of a risk for them personally, or against other health harms, such as nuisance alarms). The McConnell et al (1996)

Figure 2: Idea webbing



intervention presumably increased residents' estimates of their own risk but no information is provided about this. The third and less robust concept is residents' level of trust. In the McConnell study, residents were involved in the development of the intervention and presumably this generated a certain amount of trust (even though the education was mandatory). In the Roberts study, the fact that some residents were uncomfortable with strangers coming into their homes to fit smoke alarms, and suspicious of anything offered free, suggests a lack of trust. These three concepts may just represent the potential for translation between the three studies. However, they may also start to characterise the elements necessary for a successful intervention: landlord commitment, risk perception and residents' level of trust.

This kind of translation could eventually lead to the development of theory to explain the results of the NS: this would form one output of the synthesis and might inform the development of future interventions.

At this stage the reviewers constructed a tabular synthesis, bringing together the earlier analyses based on tabulations, thematic analysis and translation (see Table 6).

Figure 3: Conceptual mapping

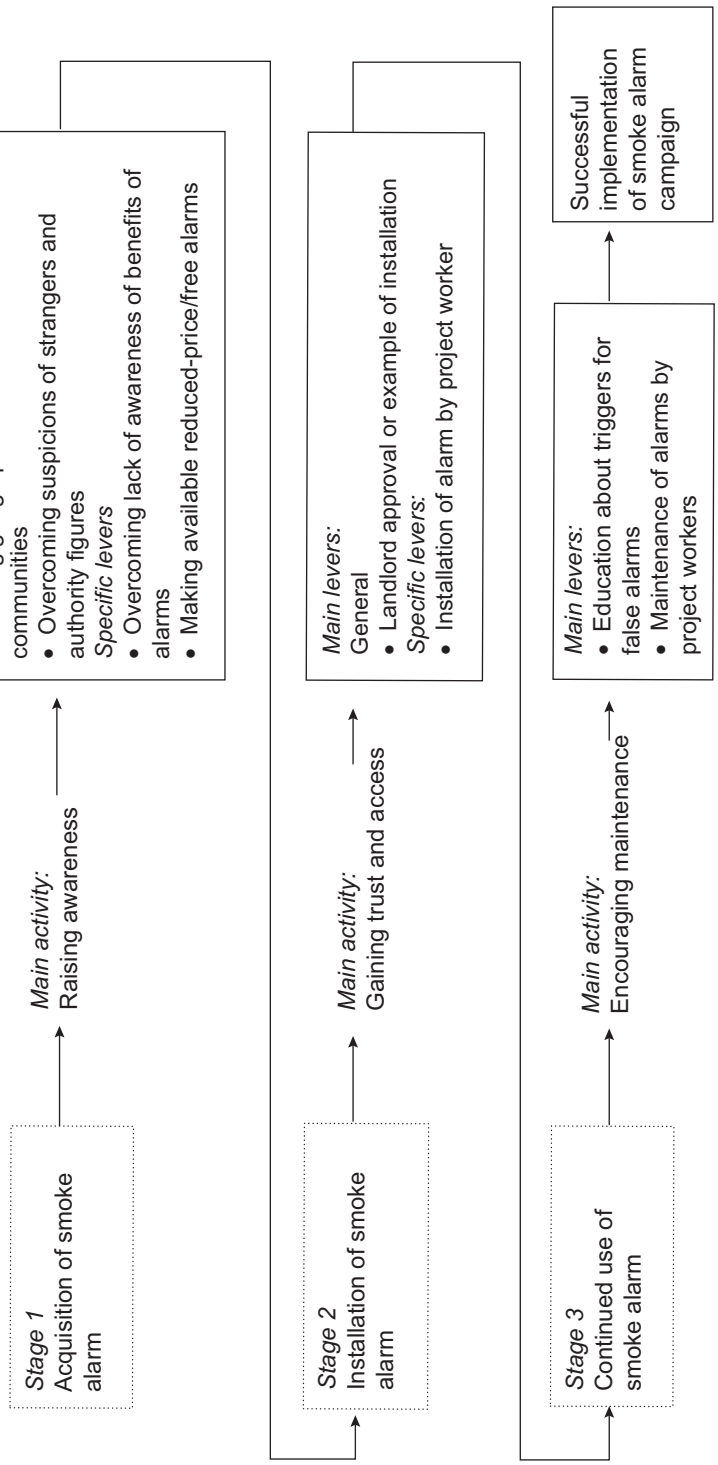


Table 6: Synthesis tabulation

Location and author	Context		Explanatory factors	
	Targeted stage of intervention	Level	Landlord commitment	Risk perception
Australia: Camit (1998, 2002); Young et al (1999)	Acquisition Installation	Community	Concerns about damage to rented property	Lack of perceived importance
US: McConnell (1996)	Maintenance	Heads of households (tenants)	Housing authority carried out intervention	Powerful video made in study setting
UK: Di Giuseppi et al (1999); Roberts et al (2004)	Acquisition Installation	Tenants/ community	Intervention supported by local authority and tenants' association	In disabling smoke alarms, tenants are aware of risk, but balance immediate and longer-term risks to well-being
US: Campbell DeLong (2003)	Maintenance	Tenants	Population vulnerable to negligent landlords; good property management will solve problem	Lack of understanding of purpose of smoke alarm

The reviewers' task was to bring together the preliminary synthesis product in such a way as to produce a coherent and evidence-based narrative.

Evans (2002) has argued, however, that it is important to distinguish between 'descriptive synthesis' and 'interpretive synthesis' and is critical of the heavy reliance placed by some reviewers on synthesis by tabulation. For commentators such as Evans, the relationship between the visual representation of data (the descriptive synthesis) and the narrative elaboration of the patterns identified (the interpretative synthesis) is critical to the quality of a NS. In this context, therefore, the reviewers produced a narrative account of the 'story' told by the tabulated synthesis findings, as follows:

> The interventions described in these papers were all implemented in the context of disadvantaged populations, mostly living in rented property. The interventions were targeted at three different stages: acquisition, installation or maintenance of domestic smoke alarms. They were aimed at communities, individual tenants and other parties including landlords, estate agents and local shopkeepers. Two aspects of implementation appear to affect the success or otherwise of the interventions, although there are other relevant factors. The commitment of landlords is beneficial at all stages: by providing

> alarms free of charge, or to overcome tenants' fears about damage to property; to support or pay for installation; and to encourage and/or enforce maintenance. Landlord commitment is insufficient on its own, but negligent landlords are a major obstacle at all stages. Tenants' appropriate perception of their own risk of domestic fires, which may be enhanced by a tailored and site-specific intervention, is necessary but insufficient at two stages: to motivate acquisition (in situations where the landlord does not supply smoke detectors); and to achieve maintenance. From these studies, it was not possible to reach a conclusion about the role of risk perception in the third stage, that of smoke alarm installation.

It is important to stress that this synthesis was conducted on a sample of studies rather than being the product of a comprehensive search. It is possible that a more comprehensive synthesis would have identified a wider range of factors shaping the outcome associated with smoke alarm interventions and, as a result, produced a more comprehensive explanatory account of these outcomes. In particular, this synthesis did not examine the often rational choices made by tenants in balancing risks, whether these are allowing a stranger into their home to fit an alarm or trading off noise, nuisance and relationships with neighbours against an oversensitive alarm – choices revealed by a number of studies in this field. Additionally, this NS did not begin with the identification of a theoretical framework: if it had, this would have provided a structure for bringing together the various outputs from the tools and techniques used.

Assessing the robustness of the synthesis

Comparison with earlier review findings

In our original NS guidance we suggested that the results of the synthesis might be compared with those of an earlier related review (Popay et al, 2003). In the demonstration synthesis, this comparison was a useful contribution to the assessment of robustness. Differences between reviews may be explained in terms of difference in the synthesis process and/or included studies. The results of the demonstration NS reported here were compared with the earlier review of the same evidence base, which had involved a simple thematic analysis of the findings of included studies, with no attempt at synthesis using other tools or techniques. This identified features present in papers containing a 'thicker' description of implementation processes. The authors concluded that where interventions had been successfully implemented, the programmes were likely to have the following features:

- a relatively detailed description of the intervention, its strengths and weaknesses, and its suitability for the targeted population;
- some consideration of the social context within which the trials take place;
- some recognition of the discrepancy between the design and orientation of an intervention and its implementation in an everyday setting;
- an exploration of the reasons for anomalous results and findings;

- some description of the specific factors affecting implementation. These might include an understanding the people and the community receiving the intervention, and the role of community leaders and other key local figures to programme success.

These are useful, although rather general, insights into factors affecting implementation. In fact, this is less a synthesis than a list of insights about implementation drawn from the reports.

In comparison, the results of the NS reported here suggest potential explanatory factors in implementation that might account for the success or failure of particular interventions in different contexts. These explanatory factors were not identified in the earlier review but emerged in the process of using the various tools and techniques.

The NS guidance describes a number of tools that may aid in the assessment of the robustness of the synthesis results. These were not used in the synthesis described here for a number of reasons. The 'weight of evidence' and 'best evidence synthesis' approaches could not be used as they required information on the methodological quality of included studies, which was not available. The latter would also require a larger number of included studies.

Checking with the authors of primary studies was not undertaken because of limited time. Checking with authors is potentially valuable but it does depend on the accessibility and generosity of authors in providing further information. Our experience of this in other studies has been almost entirely positive (Popay et al, 2003).

Critical reflection and conclusion

The synthesis reported here was based on a small number of studies selected purposively on the basis that they provided 'thicker' evidence on implementation (Popay et al, 2003). This approach to inclusion was better suited to an exercise like this one where we were testing the processes of NS rather than seeking to provide a definitive synthesis product or 'answer'. However, it should be remembered that at present, many, perhaps most reviews are based on small numbers of studies with similarly scant data on implementation; normally, however, this would be because more studies have not been found. Methodological work such as this does not claim to be an exhaustive synthesis and there are several other factors that may be expected to impinge on the uptake and functioning of domestic smoke alarms, such as the level of community engagement and the effect of policy or legislation on installation.

The work reported here demonstrates how the process of NS can be made more transparent and suggests that using some of the tools and techniques described in the guidance produced by the authors of this article can improve the quality of this type of synthesis. It has also highlighted a number of other issues in the conduct of NS. For example, combining the work of two reviewers can provide information on inter-researcher differences. The reviewers in the demonstration synthesis used the different tools and techniques in parallel ways, not necessarily identifying precisely

the same themes or concepts. Indeed, one of the reviewers carried out two thematic analyses at different times, and identified different themes in her later reflections. This could be seen as leading to a lack of reliability in conclusions, although here the use of tools and techniques allowed greater transparency than is usually possible.

The product of the synthesis will reflect the experience of the reviewers as well as their familiarity with tools and techniques. The reviewer with previous experience of meta-ethnography, for example, found it easy to use the technique of translation to compare the concepts of different studies. Thus the range of tools and techniques provides options to choose from, and individual researchers will have their own preferences depending on their own style. For example, some will be more comfortable with visual techniques such as concept mapping, while others will be more comfortable with tabular representations. No attempt was made in the demonstration project to achieve consensus between the reviewers, as we were interested in documenting their separate attempts and describing differences. It seems likely that greater previous research experience and greater immersion in the data will produce a more thorough synthesis. Use of several tools and techniques promotes deeper engagement with the data to be reviewed than use of a single tool or technique, and we would encourage others to use as wide a range of techniques as time allows.

In comparison with the earlier exploratory review, use of the guidance enabled a more systematic approach to the different papers and a more nuanced appreciation of the evidence. As a justification for NS is based, in part, on its claim to address the potential for bias, this demonstration review indicates that the use of specific tools and techniques can provide transparency of process. A more robust product is likely to be achieved if at least two reviewers work independently and then compare their findings to produce a mutually agreed (or a transparently divergent) final version.

The work also raises questions about the role of the subject 'expert' in NS in particular and systematic reviewing in general. Narrative synthesis is not, in our view, simply a technical fix that can be attained through a recipe book approach. It seems likely that the complex judgements and processes of interpretation required in the review and synthesis process will be greatly assisted if at least one member of the review team has good knowledge of the subject of the review. Although this remains a question for future research, one of the authors of this article (HR), while not an expert on fire injury, was an author of one of the included studies, and this may have resulted in the inclusion of detail that could have been missed by reviewers less familiar with the study.

The methodological work reported here goes some way towards demonstrating the ways in which findings from disparate studies may be combined in a more systematic way in order to be helpful for those policy makers and planners who are enjoined to make better use of research evidence. The strength of this approach to NS is that it makes the process more open and transparent than is usually the case. Areas for further development include the challenge of searching for, and quality appraising, studies that adopt a mixed method approach to evaluating factors impacting on the implementation of interventions/programmes. There is also a need for further testing of the NS guidance involving larger numbers of studies and embedded within a comprehensive systematic review process.

Acknowledgements

The research reported here was funded by the ESRC (grant reference number H33250019) within the ESRC Methods Programme. We are grateful to Professor Angela Dale, the director of the programme, for her support throughout the project. We are also grateful to the international panel of experts in research synthesis. We would also like to pay tribute to Professor Sally Baldwin, who developed the proposal for this work with us, and who died as a result of an accident in October 2003.

Notes

[1] At the time of writing this article, Lisa Arai and Helen Roberts were based in the Child Health Research and Policy Unit, City University, London.

[2] The word 'intervention' is used here to refer to a wide range of policies and practices that seek to change the circumstances of individuals and/or groups. These may include interventions aiming to improve health status as well more complex programmes aiming to improve a wide range of outcomes including, for example, social and economic circumstances or educational attainment.

References

Arai, L., Roen, K., Roberts, H. and Popay, J. (2005) 'It might work in Oklahoma but will it work in Oakhampton? Context and implementation in the effectiveness literature on domestic smoke detectors', *Injury Prevention*, vol 11, no 3, pp 148-51.

Audit Commission (2007) *Better safe than sorry: Preventing unintentional injury to children*, London: Audit Commission, www.healthcarecommission.org.uk/nationalfindings/publications.cfm

Brock, T. (2003) *Developing a methodological framework for the design of implementation studies*, New York, NY: MDRC.

Camit, M. (1998) *Smoke alarms wake you up if there is a fire: Report on the Arabic, Chinese and Vietnamese communities smoke alarm purchase and installation program 1998*, Gladesville, NSW, Australia: NSW Multicultural Health Communication Service.

Camit, M. (2002) 'Smoke alarms wake you up if there's a fire: a smoke alarm campaign targeting Arabic, Chinese, and Vietnamese communities in New South Wales', *Social Marketing Quarterly*, vol 8, no 1, pp 52-4.

Campbell DeLong Resources Inc (2003) *Barriers to maintaining working smoke alarms: Service area demographic study and interview research*, Portland, OR: Campbell DeLong Resources Inc, www.tvfr.com/dept/cs/campaigns/smokealarm/CDRI_Smoke%20Alarm%20Research%20Report.pdf

Denzin, N.K. and Lincoln, Y.S. (1994) *Handbook of qualitative research*, London: Sage Publications.

DiGuiseppi, C. and Higgins, J.P.T. (2000) 'Systematic review of controlled trials of interventions to promote smoke alarms', *Archives of Disease in Childhood*, vol 82, no 5, pp 341-8.

DiGuiseppi, C., Slater, S., , I., Adams, L., Sculpher, M., Wade, A. and McCarthy, M. (1999) 'The "Let's Get Alarmed!" initiative: a smoke alarm giveaway programme', *Injury Prevention*, vol 5, no 3, pp 177-82.

Evans, D. (2002) 'Systematic reviews of interpretive research: interpretive data synthesis of processed data', *Australian Journal of Advanced Nursing*, vol 20, no 2, pp 22-6.

Graham, H. (1993) *When life's a drag: Women, smoking and disadvantage*, London: HMSO.

Graham, H. and Blackburn, C. (1998) 'The socio-economic patterning of health and smoking behaviour among mothers with young children on income support', *Sociology of Health and Illness*, vol 20, no 2, pp 215-40.

McConnell, C.F., Dwyer, W.O. and Leeming, F.C. (1996) 'A behavioral approach to reducing fires in public housing', *Journal of Community Psychology*, vol 24, no 3, pp 201-12.

Noblit, G.W. and Hare, R.D. (1988) *Meta-ethnography: Synthesizing qualitative studies*, Newbury Park, CA: Sage Publications.

Pawson, R. (2002) *Does Megan's law work? A theory-driven systematic review*, Working Paper 8, London: ESRC UK Centre for Evidence Based Policy and Practice, http://evidencenetwork.org/Documents/wp8.pdf

Pawson, R. (2006) *Evidence-based policy: A realist perspective*, London: Sage Publications.

Petticrew, M. and , H. (2006) *Systematic reviews in the social sciences: A practical guide*, Oxford: Blackwell.

Popay, J., Arai, L., , H. and Roen, K. (2003) *Exploring methodological and practical issues in the systematic review of factors affecting the implementation of child injury prevention initiatives*, London: Health Development Agency.

Popay, J., , H., Sowden, A., Petticrew, M., Arai, L., Britten, N., Rodgers, M. and Britten, N. with Roen, K. and Duffy, S. (2006) *Guidance on the conduct of narrative synthesis in systematic reviews: Final report*, Swindon: ESRC Methods Programme, www.esrc.ac.uk/ESRCInfoCentre/index.aspx

, H., Arai, L., Roen, K. and Popay, J. (2006) 'What evidence do we have on implementation?', in A. Killoran, C. Swann and M.P. Kelly (eds) *Public health evidence: Tackling health inequalities*, Oxford: Oxford University Press, pp 299-308.

Roberts, H., Curtis, K., Liabo, K., Rowland, D., DiGuiseppi, C. and Roberts, I. (2004) 'Putting public health evidence into practice: increasing the prevalence of working smoke alarms in disadvantaged inner city housing', *Journal of Epidemiology and Community Health*, vol 58, no 4, pp 280-5.

Rodgers, M., Petticrew, M., Sowden, A., Arai, L., Britten, N., Popay, J. and Roberts, H. (2007) 'Testing the guidance on the conduct of narrative synthesis in systematic reviews: effectiveness of interventions to promote smoke alarm ownership and function', unpublished paper.

Roen, K., Arai, L., Roberts, H. and Popay, J. (2006) 'Extending systematic reviews to include evidence on implementation: methodological work on a review of community-based initiatives to prevent injuries', *Social Science and Medicine*, vol 63, no 4, pp 1060-71.

van Meter, D.S. and van Horn, C.E. (1975) 'The policy implementation process: a conceptual framework', *Administration and Society*, vol, 6, no 4, pp 445-88.

Weiss, C.H. (1997) 'How can theory-based evaluation make greater headway?', *Evaluation Review*, vol 21, no 4, pp 501-24.

Young, M., Camit, M. and Mihajlovic, M. (1999) 'A smoke alarm campaign in Arabic, Chinese and Vietnamese communities', *NSW Public Health Bulletin*, vol 10, no 10, pp 133-5.

Lisa Arai, Social Science Research Unit, Institute of Education, University of London, UK
l.arai@ioe.ac.uk
Nicky Britten, Peninsula Medical School, Universities of Exeter and Plymouth, UK
nicky.britten@pms.ac.uk
Jennie Popay, Institute for Health Research, Lancaster University, UK
j.popay@lancaster.ac.uk
Helen Roberts, Social Science Research Unit, Institute of Education, University of London,
h.roberts@ioe.ac.uk
Mark Petticrew, MRC Social and Public Health Sciences Unit, University of Glasgow, UK
mark@msoc.mrc.gla.ac.uk
Mark Rodgers, mr14@york.ac.uk and **Amanda Sowden**, ajs18@york.ac.uk
Centre for Reviews and Dissemination, University of York, UK