Full title

Techniques for identifying cross-disciplinary and 'hard-to-detect' evidence for systematic review

Short title

Identifying hard-to-detect evidence

Authors and affiliation

A O'Mara-Eves, G Brunton, D McDaid, J Kavanagh, S Oliver, J Thomas

EPPI Centre, Institute of Education, University of London, UK

Corresponding author:

Dr Alison O'Mara-Eves

EPPI-Centre, Social Science Research Unit

Institute of Education, University of London

18 Woburn Square, London, United Kingdom, WC1H 0NR

Email: a.omara-eves@ioe.ac.uk

Tel: +44 (0)20 7612 6393

Fax: +44 (0)20 7612 6400

Abstract (max 200 words)

Driven by necessity in our own complex review, we developed alternative systematic ways of identifying relevant evidence where the key concepts are generally not focal to the primary studies' aims and are found across multiple disciplines—that is, hard-to-detect evidence. Specifically, we sought to identify evidence on community engagement in public health interventions that aim to reduce health inequalities. Our initial search strategy used text mining to identify synonyms for the concept "community engagement". We conducted a systematic search for reviews on public health interventions, supplemented by searches of trials databases. We then used information in the reviews' evidence tables to gather more information about the included studies than was evident in the primary studies' own titles or abstracts. We identified 319 primary studies cited in reviews after full-text screening. In this paper, we retrospectively reflect on the challenges and benefits of the approach taken. We estimate that more than a quarter of the studies that were identified would have been missed by typical searching and screening methods. This identification strategy was highly effective and could be useful for reviews of broad research questions or where the key concepts are unlikely to be the main focus of primary research.

Keywords

Review literature as topic; information retrieval; searching; text mining; screening

Background

Searching broad topic areas raises particular challenges as they can cover many disciplinary areas. For example, community engagement interventions cut across many disciplines, topics and outcome domains including housing, transport, social inclusion, accident prevention, and substance abuse (Popay, Attree, Hornby et al., 2007). Such breadth demands that reviewers familiarise themselves with the terminology and research sources (journals, archives of research, etc.) of varied disciplines. Also, searching broadly requires the location and screening of many reports in order to identify a much smaller quantity of relevant research evidence. While broad searches are important to ensure that all of the potentially relevant literature is located, such exhaustive searching is costly and time-consuming.

Perhaps even more troublesome for reviewers is the issue of topic focus, since systematic reviews can sometimes aim to examine cross-cutting issues, such as inequalities, in order to draw conclusions which go beyond the focus of individual primary studies. Thus, the main focus of a review often differs significantly from the questions asked in the primary research it contains; this means that issues of significance to the review may not be referred to in the titles and abstracts of the primary studies, even though the primary studies actually do enable reviewers to answer the question they are addressing. As an example, primary research is often conducted with samples that are predominantly from disadvantaged groups, but the study might not mention this in either title or abstract when reducing health inequalities is not the focus of the intervention. Relevant evidence is therefore likely to be missed by typical review methods that have an initial screening stage in which only titles and abstracts are reviewed. For example, in a review of health inequalities and young people, 31% of intervention evaluations excluded on the basis of title and abstract as not addressing health inequalities were subsequently found to be relevant when the full paper was assessed (Oliver et al 2008). For many reviews, the risk of missing studies through title and abstract-only screening as a result of the mismatch between the focus of the review and the primary studies can be reduced through supplementary searching (e.g., citation chasing or expert recommendations; Woodman, Harden, Thomas et al., 2010), but this may not be sufficient for cross-cutting, diffuse, or complex topics.

A primary goal of a systematic review is to minimise bias when examining the evidence and an important part of that is the avoidance of systematically missing relevant evidence (Higgins & Green, 2011). Broad review topics and those with hard-to-detect[1] evidence are at particular risk of systematically missing evidence. Indeed, it is well noted that systematic reviews of complex and heterogeneous evidence require special approaches to study identification (Greenhalgh & Peacock, 2005; Wentz, Roberts, Bunnv, et al., 2001; Woodman et al., 2010). Such approaches include searching a large number of databases which cover a range of disciplines. In the area of public health this could mean searching biomedical, nursing, social science, psychology, education, and built environment databases. This brings

---

[1] We use the term hard-to-detect rather than hard-to-find, as the latter term is usually used in reference to studies that are known to the reviewer but a full-text copy cannot be located or retrieved. In contrast, hard-to-detect studies are those that the reviewer has not yet identified as being relevant and might have difficulty in doing so because of information reported in the study's title and abstract.

challenges as key concepts are described using different terminology in different disciplines. Techniques to identify correct terminology requires use of database thesauri and indexes (e.g., Medical Subject Headings) (where available), 'snowballing' of search terms based on known includes, extensive background reading across disciplines, and consultation with topic experts. However, there are no clear guidelines on ways to identify evidence in these circumstances and previous attempts to devise approaches to electronic searching with high sensitivity and precision for public health topics have reported difficulties (e.g., Wentz et al., 2001).

One approach to identifying additional terms that has been proposed is the use of text mining (Ananiadou, Rea, Okazaki, Procter, & Thomas, 2009). In contrast with a traditional approach to developing a search strategy, in which the list of search terms is exclusively manually constructed by the reviewers, text mining can automatically identify key terms within relevant documents, which can be *added to* the reviewer-derived terms. Given the recent development of text mining approaches, there is little published literature on their use in developing search strings. The available evidence is promising; for example, Hausner, Waffenschmidt, Kaiser, and Simon (2012) describe the evaluation of textual analysis processes for developing search terms for a review on a clinical health topic (see also Simon, Hausner, Klaus, and Dunton, 2010 for an example in the nursing context). Moreover, these methods are increasingly being promoted for use in search string development (as evidenced by, for example, a workshop on "Text analysis tools for information retrieval" by Hausner and colleagues at the 19th Cochrane Colloquium in 2011) and are a part of the toolbox that some information scientists use as a way to validate search filters and check terms included in test search strategies (e.g., Hausner et al., 2012; C. Stansfield, personal communication on 21 June 2013).

Ananiadou et al. (2009) suggest three possible ways that text mining technologies can be employed in the context of improving the search strategy (i.e., term extraction, document clustering, and document classification). In this paper, we focus on one of these approaches, term extraction, as it is easily implemented and the software is freely available. This technique and the software used (TerMine; Frantzi, Ananiadou, & Mima, 2000) are described in the methods section below.

Methods

Objectives
The purpose of this paper is to demonstrate alternative systematic ways of identifying relevant evidence where the key concepts that are of interest cut across various disciplines and are generally not central to the primary studies' aims. That is, we aim to identify 'hard-to-detect' evidence. We demonstrate this through description and retrospective evaluation of techniques that we used to identify evidence for a review on community engagement in public health interventions to reduce health inequalities (O'Mara-Eves, Brunton, McDaid, et al., in press).

There are three parts to the identification process that are adapted for this purpose: the way in which we develop the search syntax for searching electronic databases ('Search terms'), the sources that were searched ('Research sources'), and the process through which documents are screened for inclusion in the review ('Screening'). The approaches used are described in the following sections.

Search terms

We used a variety of standard approaches when developing the search strategy (e.g. reviewer knowledge of terms, reading of background literature and exemplar papers and examining the search syntax of similar reviews). However, novel to this review, we additionally used a text mining method ("automatic term extraction") (Thomas, McNaught, Ananiadou, 2011) to identify synonyms for the concept "community engagement", to ensure that cross-disciplinary terms were detected. We ran the full-text documents of five reviews and discussion papers of community engagement (Popay et al., 2007; Campbell, Hughes, & Gilling, 2008; Mason et al., 2007; NICE, 2008; Swainston & Summerbell, 2007) that we knew to be relevant to the review through TerMine text mining software (Frantzi, Ananiadou, & Mima, 2000). Collectively, the reviews and discussion papers covered disciplines and topic areas as diverse as:

- economics
- employment/work/job creation
- healthcare and medicine
- housing/built environment
- local government
- neighbourhood/community renewal/regeneration/development
- nursing and allied health
- psychology
- public health/ health promotion
- social care
- social policy
- sociology
- transport
- the voluntary sector.

TerMine is a term extraction program that identifies the key terms and compound terms in a body of text and produces a list of terms in the papers in order of their C-value. The C-value is a statistical measure of the frequency and significance of term occurrence and indicates the salience of the term within a document (Frantzi et al., 2000). As an example, Table 1 shows the top 30 terms (out of 1,907 terms) and their C-scores as identified by TerMine in the Popay et al. (2007) paper. We examined the list of terms generated by the software to determine those that were synonyms or related terms to community engagement. We only looked at terms with a C-value of 5.0 or above; **t**his threshold was chosen as it was the

common value below which mined terms seemed to lose relevance across the five papers. Also, if the term has greater salience as indicated by higher C-value, then it is reasonable to believe that the term has greater traction in the literature. In other words, a low C-value indicates that the term is peripheral to the focus of the document. Given that the documents are known to be relevant to our review topic, then it is likely that the more salient terms within a document are also more likely to be salient in terms of the concept of interest.

It is important to emphasise that the terms uniquely identified through TerMine were added to the terms already identified by the reviewers through other methods—that is, text mining complemented rather than substituted the typical approaches to developing the search syntax. In other words, this approach differs from typical search syntax development through the inclusion of the text mining-identified terms as an *additional* source of terms. The search syntax is presented in Appendix A.

<<Add Table 1 about here>>

Research sources
We used three types of sources when identifying research for inclusion in our review: databases of systematic reviews, databases of trials relevant to public health, and author contacts.

First, we identified systematic reviews of public health and health promotion interventions through searching various websites and databases devoted to systematic reviews. The aim of this step was to capitalise on the systematic searches that have already been carried out for other reviews by identifying relevant primary studies included in those reviews. We searched a range of registers, websites, and databases for systematic reviews that discuss how some or all of their included studies contain interventions that utilise community engagement.

The systematic review registers, websites, and databases that we searched were:

1. Database of promoting health effectiveness reviews (DoPHER).
2. Cochrane database of systematic reviews (CDSR).
3. Database of abstracts of reviews of effects (DARE).
4. Campbell Library.
5. NIHR Health Technology Assessment (HTA) programme website.
6. Health Technology Assessment (HTA) database hosted by CRD.

These sources were selected as they are known to have a wide range of up-to-date systematic reviews relevant to public health and health promotion. In addition, Woodman et al. (2010) found that specialist review databases including DARE and DoPHER were time efficient places to search for reviews on health inequalities.

The reviews were used to identify included primary studies that are relevant to the scope of this project; the systematic reviews themselves were not included in the synthesis in this project. Importantly, when attempting to identify primary studies within reviews, we used information in the systematic review evidence tables and full-text document to gather more

information about the included studies than was evident in the primary studies' own titles or abstracts.

Second, searches of the systematic review resources were supplemented by searches of the Trials Register of Promoting Health Interventions (TRoPHI) database and the NHS Economic Evaluation Database (NHS-EED). The studies in these databases are themselves the product of systematic searches. Importantly, the studies in the TRoPHI database have been previously coded by researchers in the process of conducting other systematic reviews, regarding aspects of the studies such as participant characteristics and intervention details (e.g., intervention deliverer). We were therefore able to draw on the information coded in TRoPHI to supplement information in the title and abstract when screening.

Third, we contacted authors of a small number of key studies that were excluded on methodological grounds to ask them if they have outputs that would meet our inclusion criteria, or if they could provide further information about the study to assess its suitability for inclusion.

This approach differs from typical search processes mainly in terms of focusing on the actual contents of the systematic reviews (not just their reference lists) and searching databases of primary studies that have been indexed against keywords that are relevant to this review. Typical approaches to searching for primary studies (i.e., as distinct from reviews of review-level evidence, which deliberately seek to include reviews in the synthesis) tend to focus on identifying primary studies through sources including database, trials registers, websites, hand-searching journals, and citation chasing techniques. That is, they do not generally look at the information included in the full text of any systematic reviews that are identified, even if they do scan their reference lists.

Screening

The outcome of the searches was a database of references and documents which were screened using the review's inclusion criteria. There were multiple stages of screening to avoid hasty exclusion of studies:

1. Title and abstract screening of systematic reviews.
2. Full-text screening of systematic reviews.
3a. Evidence table screening of primary studies within reviews.
3b. Title and abstract screening of primary studies located through TRoPHI and NHS-EED with loose inclusion criteria.
4. Full-text screening of primary studies.

When screening the systematic reviews (stages 1 and 2), we were over-inclusive when applying the inclusion criteria. That is, we were inclusive when assessing the hard-to-detect concept of 'community engagement', while 'health inequalities' was not assessed at these stages. Likewise, to avoid hasty exclusion when screening the titles and abstracts from the primary studies located through electronic databases (stage 3b), we did not screen for the hard-to-detect concepts of 'community engagement' and 'health inequalities'. To demonstrate how the inclusion criteria became progressively more specific to the aims of our

review, we present the inclusion criteria for screening stages 1 and 4 below. Note that the cut-off date for inclusion in the review across all stages was 1990.

Reviews that were screened on the basis of the title and abstract (stage 1) were included if they:

1. Were published after 1990 (in line with previous related reviews);
2. Were a systematic review (i.e., describe search strategies and inclusion criteria used);
3. Included outcome, economic or process evaluation studies;
4. Described one or more interventions relevant to community engagement;
5. Were written in English;
6. Measured and reported health or community outcomes.

Primary studies that were screened on the basis of the full-text document (stage 4) were included if they:

1. Reported primary research;
2. Were not a Masters thesis;
3. Included intervention outcome, economic and/or process evaluations;
4. Focused on community engagement as the main approach;
5. Contained a control or comparison group;
6. Characterised study populations/reported differential impacts of social determinants of health captured by the PROGRESS-Plus framework (i.e., Place of residence, Race/ethnicity, Occupation, Gender, Religion, Education, Socio-economic position, Social capital, plus other characteristics that attract discrimination such as age, disability, and sexual orientation); and
7. Reported health or health-related (including cost) effectiveness outcomes and/or process data.

The flow of literature through the review is presented in Figure 1.

<<Insert Figure 1 about here. Figure label:

Figure 1 - flow of systematic reviews (blue) and primary study reports (red) to the map>>

This approach to screening differed from typical approaches in that we looked at information beyond the titles and abstracts of citations (namely, in the information reported in the full-text of systematic reviews that describe the primary studies) to assist in the decision of whether to include or exclude a citation. Typical approaches only consider the information included in the title and abstracts of citations found through searching.

Evaluation of the approach
After screening was completed and the final pool of included studies was identified, we assessed the efficacy of the strategies used.

1. To gain an overall sense of the effectiveness of the evidence identification strategy, we compared the yield of our screening process against a pre-determined expected

yield. The latter was based on the number of studies included in reviews of similar topics (1, 8) and was estimated to be about 40 studies.

2. To assess the effectiveness of using text mining to expand our search terms to those broader than our original conception, we determined whether additional search terms were generated through text mining.

3. To assess the effectiveness of using the evidence tables and broader information in the systematic reviews to identify hard-to-detect concepts, we reviewed the titles and abstracts of all included studies to see if they mentioned anything related to both community engagement and health inequalities. If the title and abstract did not mention anything related to either concept, the study was marked as "Not identifiable as relevant to community engagement and health inequalities in title and abstract"; these studies represent those that would have been missed if we had adhered to typical searching and screening methods.

Results

Overall effectiveness

The search process was remarkably successful: we identified 319 primary studies for inclusion in the review. This is almost eight times more than expected before searching began, based on the number of studies included in reviews of similar topics (Popay et al., 2007; Swainston & Summerbell, 2007).

In addition, terms from community engagement which were identified through standard approaches (e.g., background reading) were poorly represented in the titles and abstract of studies included in the map. For example, only 8 out of the 319 papers in the map used these terms in their title or abstract: "community participation" or "community engagement" or "community empowerment" or "community mobilisation" or "community partnership". This suggests that using a typical study identification approach (i.e., using reviewer-known terms in the searches and only screening titles and abstracts) would have not yielded as many identified studies.

The use of text mining to facilitate search strategy development for broad topics

Table 2 presents the list of terms that were considered for inclusion in the search syntax which were identified through text mining of the full-text documents of known relevant reviews (Popay et al., 2007; Campbell, Hughes, & Gilling, 2008; Mason et al., 2007; NICE, 2008; Swainston & Summerbell, 2007). All of these had C-scores greater than 5. Text mining revealed useful synonyms and terms associated with "community engagement" that had not previously been considered (e.g., "local involvement networks" and "participation compacts"). It is noteworthy that most of the terms that were not related to community engagement tended to relate to economic concepts (e.g., "cost effectiveness" and "economic appraisal"). In many cases, however, the terms had already been identified by the reviewers as relevant, which served to validate the usefulness of including those terms. The resulting syntax, which included both reviewer and text mining-identified terms, was more sensitive than we would have anticipated using either approach on its own. As an example, the syntax

used to search the Cochrane Database of Systematic Reviews is presented in Appendix A, with the terms uniquely identified through text mining in bold font.

The use of systematic reviews to identify primary studies

We estimate that 91 out of 319 (28.5%) of the studies that were identified would not have been detected if we had relied solely on screening titles and abstracts. The results are presented in Table 3. These studies were identified because of information reported about the studies in a systematic review or coded in the TRoPHI database that was not mentioned in the studies' title or abstract.

<<Insert Table 2 about here>>

<<Insert Table 3 about here>>

Discussion

We highly recommend that reviewers use automatic term extraction to identify additional search terms particularly where the topic is broad or multidisciplinary, or where the reviewers have methodological expertise but less specialist topic knowledge (as is often the case, (Shepherd, 2012). The process is straightforward and the software is freely available on the internet[2], yet the value added by reducing the risk of systematically missing evidence that uses different terminology is very large. The only (minor) challenge of implementing text mining was setting a reasonable threshold for the C-value, which would need to be determined on a review-by-review basis depending on the yield of terms identified by the text mining software.

In terms of the search sources, we took a different approach in this review because previous experience has shown us that it can be difficult to identify community engagement approaches and health inequalities on the basis of the abstract of a study. We knew that if we followed typical methods, the full text of nearly all citations retrieved through database searching would need to be retrieved in order to check for these hard-to-detect topics. This is because both concepts encompass a wide range of terms and their relevance may therefore not be clear from the title and abstract of a paper.

We therefore decided to utilise existing systematic reviews in a way that is unusual, at least in our experience. Since systematic reviews usually contain detailed and structured summaries of the studies they include (often in the form of extensive tables in appendices), we took the view that these summaries would be a useful source of potentially relevant studies, and would often contain more pertinent information for our purposes than is present in many abstracts.

The results of our retrospective evaluation suggest that this strategy was far more successful than we had expected. We conclude that we would not have been able to find the range of

---

[2] TerMine is freely available through the National Centre for Text Mining at http://www.nactem.ac.uk/software/termine/.

studies that we did using traditional searching methods. Although reviews commonly report identifying studies through sources other than electronic databases (such as through hand searching, reference checking, author contact, and citation chasing), our experience suggests that the number and proportion identified is not as large as we detected here. Moreover, we believe that this approach spared us from having to retrieve potentially many thousands of full text reports whose relevance was ambiguous based on the information provided in the title and abstract. The extent to which time or effort in the reviewing process was saved is not entirely clear. This is a retrospective evaluation and, unfortunately, we did not collect data on the time taken for each of the parts of the process. This data would be useful in determining the efficiency of these approaches in future applications of these approaches. We do note, however, that the success of the approach in terms of identifying many more included studies than expected led to the request of an extension to time (five extra months) and funding of the project. The larger than expected number of studies had two main implications for the review. Firstly, we had to retrieve more than twice as many documents as we had expected. Secondly, the task of selecting, data-extracting, and analysing so many studies substantially increased. As such, whilst the actual process of identifying the studies may or may not have impacted on the time and resource required, the outcome of our effective approaches (i.e., increased number of studies to deal with) certainly had an impact.

A potential limitation is that we may have missed useful information contained in individual studies that were not included in a systematic review. In particular, this method might miss the very latest research, as there is an inevitable delay between the searches conducted by reviewers and the publication of systematic reviews. In addition to missing studies due to time-lag, studies may be missed due to stricter inclusion criteria or use of different definitions of key concepts in the systematic review. To minimise this risk, we supplemented the review searches with searches of two trials databases (TRoPHI and NHS-EED). We recommend that reviewers adopting this approach should consider ways in which they can ensure that evidence missed by the systematic reviews are captured through supplementary processes such as searching trials databases and registers, author contacts, and forward citation chasing (see Greenhalgh & Peacock, 2005, for useful approaches for complex topics).

It is also important to consider the aims of the review in terms of coverage of the literature. In our review, we aimed to generate a dataset of studies that is not systematically biased in what it contains or what is missed, rather than necessarily having an exhaustive dataset (the latter is arguably an impossible ideal). Our approach to searching and screening should hopefully avoid systematic bias in what is missed by focusing on systematic reviews that themselves have transparent, systematic, and (presumably) unbiased search strategies. In this way, the reviewer needs to consider whether the approaches demonstrated here are suitable for the aims of their review.

Indeed, whilst we found that the combination of adaptations to the three parts to study identification (i.e., search terms, research sources, and screening) employed here were useful, other reviews might only adopt one or two of the adaptations, depending on the context of the review. Our review included separate syntheses that examined processes and explored theories in addition to a synthesis of the effectiveness of the interventions, and so our sample

needed to encompass variation to a greater extent than reviews with more narrowly-defined PICO (population, intervention, comparison, outcome) elements, or those with a greater need for identifying every relevant study (Gough, Thomas, & Oliver, 2012). Narrower PICO-driven reviews might need to incorporate more searches of primary evidence sources to achieve greater representativeness in the sample compared to coverage of variation in the literature.

A particular concern was the identification of economic evidence, because a limitation of the NHS-EED database is that it only includes literature of specific relevance to the UK National Health Service. We therefore additionally looked for papers that were linked to the included citations—i.e., papers published separately from the effectiveness analyses that reported different aspects of the same study. Through this approach, we identified a handful of economic studies that were helpful to the economic review but had not been included in the NHS EED database. This may also reflect a time lag in uploading less UK-centric economic studies to this database.

Another possible limitation is the restriction of our review to English language documents. We see no reason, however, that the approaches taken here could not be extended to other languages using the same translation processes used in a typical search procedure. Indeed, this approach may be helpful to reviewers who are not fully fluent in a second language and may help with developing a list of key terms and phrases in another language.

A final point is that the usefulness of the text mining approach is likely to depend on the appropriateness of the initial references identified. The term extraction algorithm depends on the content of the documents supplied to it, so it should never be used on its own but rather in conjunction with the expertise and usual processes that are followed when developing a search strategy. Using a larger number of documents and documents from a variety of research teams and perspectives will most likely yield a greater number and variety of terms. However, because we are proposing that the text mining is used to identify *additional* search terms for the search strategy, rather than replacing typical approaches to developing a list of search terms (e.g., reviewer knowledge, background reading, and drawing on search terms used in other relevant reviews), the only potential harm of running the text miner on an unhelpful set of initial references is wasted time. The search syntax itself should be no worse than had the text mining not been attempted because no new terms will be added to the search syntax beyond those already identified by the reviewers. A possibility for expanding the text mining approach (not used or evaluated here) could be to analyse additional relevant references as they are found, to see whether new terms are identified. If they are, then a supplementary search might be conducted to detect any further studies that use the 'new' terms. If no further terms are identified, then this could support the view that a saturation of salient terms has been achieved in the search strategy.

Conclusion
Text mining helped to identify relevant search terms for a broad topic that is inconsistently referred to in the literature. Using the information in systematic reviews and specially-

indexed databases helped to identify interventions and populations relevant to our review that would otherwise not have been detected. The identification strategy was effective and is useful for reviews of broad research questions or where the key concepts are unlikely to be the main focus of primary research. The strategy maintained a transparent and systematic approach to identifying relevant evidence.

Acknowledgements

References

Ananiadou S, Okazaki N, ProcterR, Rea B, Thomas J. (2009). Supporting systematic reviews using text mining. *Social Science Computer Review,* 27(4): 509-523.

Campbell F, Hughes L, Gilling T (2008) *Reaching out: community engagement and health.* London: Improvement and Development Agency (IDeA).

Frantzi, K., Ananiadou, S. and Mima, H. (2000) Automatic recognition of multi-word terms. *International Journal of Digital Libraries,* 3(2): 117-132.

Gough D, Thomas J, Oliver S. (2012). Clarifying differences between review designs and methods. *Systematic Reviews,* 1(28). doi:10.1186/2046-4053-1-28.

Greenhalgh T, Peacock R (2005) Effectiveness and efficiency of search methods in systematic reviews of complex evidence: Audit of primary sources. *BMJ,* doi:10.1136/bmj.38636.593461.68.

Hausner E, Waffenschmidt S, Kaiser T, Simon M. (2012). Routine development of objectively derived search strategies. *Systematic Reviews,* 1:19. doi:10.1186/2046-4053-1-19.

Higgins JPT, Green S (editors) (2011) *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Mason et al (2007) *NICE community engagement: Rapid review of the economic evidence for community engagement and community development approaches in interventions or initiatives seeking to address wider determinants of health.* York: University of York.

NICE (2008) *Community engagement to improve health.* York: NICE.

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

Oliver S, Kavanagh J, Caird J, Lorenc T, Oliver K, Harden A, Thomas J, Greaves A and Oakley A (2008) *Health promotion, inequalities and young people's health: a systematic review of research.* London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London

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.

Shepherd J (2012) Judgment, resources and complexity: A qualitative study of the experiences of systematic reviewers of health promotion. *Evaluation and the Health Professions*, published online May 21, 2012, doi: 10.1177/0163278712447222.

Simon M, Hausner E, Klaus SF, Dunton N. (2010). Identifying nurse staffing research in Medline: Development and testing of empirically derived search strategies with the PubMed interface. *BMC Medical Research Methodology,* 10(76). doi:10.1186/1471-2288-10-76.

Swainston K & Summerbell C (2007) *The effectiveness of community engagement approaches and methods for health promotion interventions: rapid review phase 3.* Teesside: NICE National Collaborating Centre.

Thomas J, McNaught J, Ananiadou S (2011) Applications of text mining within systematic reviews. *Research Synthesis Methods*, 2 (1): 1-14.

Wentz R, Roberts I, Bunn F, Edwards P, Kwan I, Lefebvre C (2001) Identifying controlled evaluation studies of road safety interventions: Searching for needles in a haystack. *Journal of Safety Research,* 32: 267– 276.

Woodman J, Harden A, Thomas J, Brunton J, Kavanagh, J, Stansfield C (2010) Searching for systematic reviews of the effects of social and environmental interventions: a case study of children and obesity. *Journal of the Meedical Library Association,* 98(2): 140–146.

Table 1 – Example of the TerMine output: Top 30 terms identified in the Popay et al. (2007) review and their C-scores

Rank	Term	C-Score
1	social determinant effectiveness review	452.00
2	community engagement	363.60
3	social determinant	172.80
4	determinant effectiveness review	133.79
5	comparator area	87.00
6	social capital	78.89
7	community development	77.21
8	outside scope	73.00
9	community member	68.00
10	community engagement initiative	67.67
11	community involvement	66.88
12	community group	65.00
13	sure start	64.69
14	new deal	64.19
15	young people	58.00
16	percentage point	54.50
17	neighbourhood renewal	53.28
18	renewal.net case	51.00
19	ndc area	44.00
20	focus group	42.00
21	local government	41.71
22	engagement initiative	40.67
23	indirect community engagement initiative	40.00
24	local authority	39.69
25	quality study population/	38.04
26	direct community engagement initiative	38.00
27	indirect ce initiative	34.87
28	national evaluation	34.64
29	research report	33.50
30	community participation	33.10

Table 2 - Terms that were considered for inclusion in the search syntax that were identified through text mining

1. Area-based initiative	2. Citizen participation	3. Coalition member
4. Community action	5. Community based	6. Community capacity building
7. Community champions	8. Community coalition	9. Community collaborator
10. Community committee	11. Community control	12. community development
13. Community empowerment	14. community engagement	15. Community engagement networks
16. Community health	17. Community involvement	18. Community member
19. Community mobilization	20. Community organisation	21. Community participation
22. Community regeneration	23. Community relation	24. Community representative
25. Community support	26. Cost effectiveness	27. Cost health benefit
28. Development approach	29. Economic appraisal	30. Engagement initiative
31. Health councils	32. Health determinant	33. Health development
34. Health impact	35. Health improvement	36. Health inequality
37. Health promotion	38. Integrated local development programme	39. Lay community
40. Lay people	41. Lay person	42. Local area agreement
43. Local governance	44. Local involvement network	45. Local strategic partnership
46. Minority	47. Neighbourhood committee	48. Neighbourhood renewal
49. Participation compact	50. Participator	51. Participatory action
52. Partnership working	53. Peer participation	54. Priority setting
55. Public engagement	56. Public health	57. Public health intervention guidance
58. Public involvement	59. Public participation	60. Rapid participatory assessment
61. Resident involvement	62. Service user	63. Service user
64. Social determinant effectiveness	65. Stakeholder	66. Volunteer

Table 3 - Number of primary studies ultimately included in the review that were identifiable by title and abstract, by search source

Identifiable as relevant to community engagement and health inequalities in title and abstract?	Identified through systematic reviews	Identified through TRoPHI or NHS EED
Yes	187 (58.6%)	39 (12.2%)
No	85 (26.6%)	6 (1.9%)

Note. Two additional studies (0.6%) were recommended by experts.

Appendix A: Syntax used to search the Cochrane Database of Systematic Reviews

*Note. **Bold** text indicates terms identified uniquely through text mining.*

"disadvantage" OR "disparities" OR "disparity" OR "equality" OR "equity" OR "gap" OR "gaps" OR "gradient" OR "gradients" OR "health determinants" OR "health education" OR "health inequalities" OR "health promotion" OR "healthy people programs" OR "inequalities" OR "inequality" OR "inequities" OR "inequity" OR "preventive health service" OR "preventive medicine" OR "primary prevention" OR "public health" OR "social medicine" OR "unequal" OR "variation"
AND
 "change agent" OR "citizen" OR "community" OR "**champion**" OR "collaborator" OR "disadvantaged" OR "lay community" OR "lay people" OR "lay person" OR "member" OR "minority" OR "participant" OR "patient" OR "peer" OR "public" OR "representative" OR "resident" OR "service user" OR "stakeholder" OR "user" OR "volunteer" OR "vulnerable"
AND
"capacity building" OR "**coalition**" OR "collaboration" OR "committee" OR "**compact**" OR "control" OR "co-production" OR "councils" OR "delegated power" OR "democratic renewal" OR "development" OR "empowerment" OR "engagement" OR "forum" OR "governance" OR "health promotion" OR "initiative" OR "**integrated local development programme**" OR "intervention guidance" OR "involvement" OR "juries" OR "**local area agreement**" OR "**local governance**" OR "**local involvement networks**" OR "**local strategic partnership**" OR "mobilisation" OR "mobilization " OR "**neighbourhood committee**" OR "**neighbourhood managers**" OR "**neighbourhood renewal**" OR "**neighbourhood wardens**" OR "networks" OR "organisation" OR "panels" OR "participation" OR "**participation compact**" OR "participatory action" OR "partnerships" OR "pathways " OR "**priority setting**" OR "public engagement" OR "public health" OR "**rapid participatory assessment**" OR "regeneration" OR "relations" OR "support"