

## RESEARCH ARTICLE

# Identifying evidence for five realist reviews in primary health care: A comparison of search methods

Claire Duddy<sup>1</sup>  | Nia Roberts<sup>2</sup> 

<sup>1</sup>Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

<sup>2</sup>Bodleian Health Care Libraries, University of Oxford, Oxford, UK

## Correspondence

Claire Duddy, Nuffield Department of Primary Care Health Sciences, University of Oxford, Radcliffe Primary Care Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, UK.  
Email: [claire.duddy@phc.ox.ac.uk](mailto:claire.duddy@phc.ox.ac.uk)

## Funding information

NIHR Systematic Reviews Fellowship, Grant/Award Number: NIHR-RM-SR-2017-08-018

The approach to identifying evidence for inclusion in realist reviews differs from that used in 'traditional' systematic reviews. Guidance suggests that realist reviews should be inclusive of diverse data from a range of sources, gathered in iterative searching cycles. Saturation is prioritised over exhaustiveness. Supplementary techniques such as citation snowballing are emphasised as potentially important sources of evidence. This paper describes the processes used to identify evidence in a selection of realist reviews focused on primary health care settings and examines the origin and type of evidence selected for inclusion. Data from five realist reviews were extracted from (a) reviewers' reference management libraries and (b) records kept by review teams. Although all reviews focused on primary health care, they used data from a wide range of document types and research designs, drawing on learning from multiple perspectives and settings, and sourced the documents containing this data in a variety of ways. Systematic searching of academic databases played an important role, supplementary search techniques such as snowballing were used to identify a significant proportion of documents included in the reviews. Our analysis demonstrates the diverse data sources used within realist reviews and the need for flexible, responsive efforts to identify relevant documents. Reviewers and information specialists should devise approaches to data gathering that reflect the individual needs of realist review projects and report these transparently.

## KEYWORDS

grey literature, information retrieval, literature searching, primary health care, realist review, realist synthesis

## Highlights

### What is already known

- Realist reviews (or realist syntheses) are an increasingly popular approach to evidence synthesis.
- Existing scoping reviews have identified significant variation in the conduct and reporting of searching processes undertaken in realist reviews.

### What is new

- This primary study presents an in-depth examination of the searching processes and documents included in a set of five realist reviews, providing a detailed description of how documents were selected in these projects.
- Our results demonstrate a diversity of approaches to gathering data for inclusion in realist reviews, and in the type of documents that can contribute data to realist analysis.

### Potential impact for JRSMB readers outside the authors' field

- As realist reviews become more common in fields outside health, there will be an increasing need to understand how guidance on searching for data for these reviews can be put into practice.
- This paper outlines important implications and areas to focus on for review teams undertaking searching for realist reviews.

## 1 | BACKGROUND

Realist review (or 'realist synthesis') is an increasingly popular approach to evidence synthesis.<sup>1,2</sup> A realist review differs from other forms of evidence synthesis or literature review in its aim of generating causal explanations for observed outcomes, based on included data. The approach may be adopted where there is a need to understand how or why outcomes occur (identify mechanisms), or to account for differences in observed outcomes in different circumstances (contexts). Realist reviews are inclusive of a wide range of evidence. Reviewers use data extracted from studies of different designs, alongside other documents (grey literature) to help build and test 'programme theories' that offer causal explanations. A glossary of 'realist' terminology that appears in this paper is provided in Table 1.

The approach to identifying data for inclusion in a realist review differs from that employed for a more 'traditional' systematic review. Methodological guidance for systematic reviews stipulates the need for highly sensitive searches applied in multiple databases, supplemented by additional strategies, including seeking out grey literature, scanning reference lists, handsearching and contacting experts in the subject area.<sup>3-5</sup> These approaches aim to maximise the information included in reviews (and, potentially, statistical power for any meta-analysis), while minimising reporting biases which could affect overall results.<sup>6</sup> Some empirical evidence exists to support these practices, demonstrating the existence and effects of such biases,<sup>7-11</sup> and attempting to assess the contribution of different approaches to study identification. Some of this evidence is contested,<sup>12-15</sup> and optimal approaches may well vary across subject areas and review types.<sup>16,17</sup>

Conversely, guidance and empirical work that emphasises the importance of 'extended', 'supplementary' or 'complementary' techniques to identify studies often

TABLE 1 Glossary of realist terminology

| Term   | Definition  |
|--|---|
| Context  | Any condition that triggers and/or modifies the behaviour of a mechanism; important features of the circumstances in which a programme exists or phenomenon happens which 'trigger' the mechanisms that generate outcomes   |
| Context-mechanism-outcome configuration (CMOC) | A diagrammatic or narrative description offering a theoretical explanation of the causal relationship between some particular context(s), mechanism(s) and outcome(s). Multiple CMOCs may exist within a 'programme theory' (see below)   |
| Mechanism                                      | The underlying process by which outcomes are generated in particular contexts. Mechanisms are usually descriptions of the tendencies, reasoning and behaviour of agents involved in a process or programme, and their responses to important context(s)   |
| Programme theory                               | A diagrammatic or narrative description of a theory that describes what an intervention comprises and how it is expected to work, or the process(es) by which outcomes of interest are thought to come about. A realist programme theory is expressed in terms of relationships between important context(s), mechanism(s) and outcome(s), or CMOCs (see above) |

focuses on 'non-traditional' reviews.<sup>18–21</sup> For example, guidance related to realist reviews recommends the use of techniques such as citation tracking, handsearching and contacting authors and subject experts, in addition to database searching.<sup>22–24</sup> Similar recommendations are made for subject areas where it is anticipated that evidence will be more diverse or difficult to locate.<sup>25–33</sup>

Research focused on primary health care settings is one such area. Evidence synthesis in primary care research has been prioritised by the NIHR School of Primary Care Research via the establishment of the 'Evidence Synthesis Working Group'.<sup>34</sup> The realist approach to conducting reviews has been recognised within this collaboration as offering an appropriate methodology to address the complexity of many problems in primary care.<sup>35</sup> Three of the reviews included in this analysis were conducted by research teams working within this group.

In contrast to other types of evidence synthesis, existing methodological guidance for realist reviews is clear that reviews can be inclusive of a diverse range of material, including multiple study designs and grey literature, across disciplines.<sup>22–24</sup> The development and testing of realist programme theories rests on the inclusion of data on contexts, mechanisms and outcomes relating to research questions; different types of data may provide evidence for each of these components.<sup>36</sup> From a realist perspective, all document types, and all study designs have the potential to contribute useful data for programme theory development and testing, regardless of quality.<sup>37</sup> However, while encouraging this inclusivity, the same guidance is clear that the overall balance of sensitivity and specificity in searching for a realist review may take into account the need to achieve 'theoretical saturation' based on a sample of the available literature, as opposed to comprehensiveness or exhaustiveness in evidence identification. In this context, 'saturation' has been described as a matter of subjective judgement on the part of reviewers, based on their assessment that 'sufficient evidence is found such that it is reasonable to claim that the theory is coherent and plausible'.<sup>22</sup> Differences in reviewers' judgement, and in the focus of different review questions, may lead to differences in approaches to searching for and selecting evidence.

Evidence identification for realist review is further complicated by the recommendation of 'iterative' searching that might proceed in multiple phases. There are four main recommended components to the approach to identify data for inclusion:

1. 'Background searching' for familiarisation with the literature and the formulation and refinement of specific research questions.
2. Searching 'to track programme theories', that is, to inform early theory development by identifying existing explanations for how an intervention works or phenomenon comes about.

3. The 'main' search to identify empirical evidence for 'theory testing', that is, empirical evidence that can refine, refute or confirm developing realist programme theories (via the development of 'context-mechanism-outcome configurations').
4. Additional searches as required, seeking additional studies to help further refine theories and respond to emergent information needs.<sup>23,24</sup>

These are not necessarily intended to proceed as linear steps, but rather iteratively, in response to programme theory development as it progresses over the course of the review. In light of these potential complexities, transparent reporting of the processes used to identify the data included in a realist review is essential, allowing end-users to judge the rigour of review processes.<sup>22</sup>

Two published reviews of existing practice in realist reviews have identified wide variation in approaches to identifying evidence for inclusion.<sup>1,38</sup> One broad review included 54 reviews and identified variation in searching methods, including many reviews which had adopted narrower approaches than guidance suggests. For example, some reviews relied only, or primarily, on searching a small number of databases. In other cases, realist syntheses were attempted based on traditional systematic review methods, without additional searching.<sup>38</sup> A recent scoping review of searching methods in realist reviews included 34 reviews, and found similar diversity of approaches, and frequently, incomplete reporting of searching processes.<sup>1</sup> Some variation is likely attributable to differences in resources and funding, as noted by some authors<sup>34</sup> and acknowledged in the RAMESES standards.<sup>35</sup> The emergence of 'rapid realist review'<sup>36</sup> also raises important questions about the extent and efficiency of searching processes and methods in realist reviews: are some searching methods more fruitful than others? How can reviewers and information specialists determine the most appropriate approaches to searching within project time and resource constraints? None of the reviews included in this study were badged as 'rapid', but the projects varied in length and resourcing.

The variation in current practice and potentially open-ended and stop-start nature of evidence identification for realist reviews raises questions for project planning and funding applications (including in relation to engaging information specialists), and in selecting the most efficient and fruitful strategies to identify the richest evidence for theory development.

## 2 | METHODS

This study was designed to closely examine the approaches taken to gather evidence in a series of realist

reviews all focused on primary health care settings. Our aim was to identify and explore the similarities and differences in the approaches taken in each review and to draw on our own experience of involvement in these reviews to consider the implications for future review projects, and methodological research in this area.

The five realist reviews included in this study all focus on different topics (early visiting services, female genital mutilation, laboratory test ordering, antipsychotic medication review and social prescribing link workers) within primary health care settings. These reviews were

identified for inclusion via the personal involvement of CD and/or NR in each review. The lead reviewers on each project agreed to contribute their data to this study. The detailed methods employed in each review are reported in their own study protocols and published reports (where applicable).<sup>39–48</sup> For this study, we collected data from each review team relating to the final list of documents included in each review. In the five included reviews, ‘included’ documents were those that contributed data to the programme theory developed in each review. Across the reviews, documents were

**TABLE 2** Document type and study design categories

| Document type categories                                   |   |
|--|---|
| Research paper   | Published and peer-reviewed academic journal article reporting original primary research or secondary research  |
| Research report  | Published or unpublished reports of research projects, including programme evaluations and reports produced by think tanks, research organisations and charities  |
| Opinion/commentary   | Documents outlining individual or group views or perspectives, usually from experts or subject specialists, including editorials and letters  |
| News   | News items sourced from newspapers and news websites, or news or update items posted on other organisational websites   |
| Policy document  | Documents describing existing positions, policy or guidelines (including clinical guidelines), and/or their implementation, or consultation or recommendation documents relating to proposed policy   |
| Conference materials                                       | Abstracts, posters or other presentation formats delivered at conferences or meetings   |
| Social media   | Documents contributing data that are in the form of blog posts, tweets or other items shared via social media platforms   |
| Other  | All other document types that contributed data to the reviews, itemised for each review below as required. These include, for example, presentations/lectures (aside from those delivered at conferences or similar meetings), dissertations/theses and books |
| Study design categories (for documents reporting research) |   |
| Qualitative interviews or focus groups                     | Qualitative research where data were collected via interviews or focus groups   |
| Other qualitative  | Qualitative research where data were collected via other methods, for example, ethnography (observation), document analysis, participatory methods, process evaluation  |
| Survey/questionnaire                                       | Cross-sectional quantitative research where data were collected via a survey or questionnaire   |
| Quantitative observational                                 | Cross-sectional or longitudinal quantitative research, including analysis of routine or specially collected data using cohort, case–control, time series, before and after and cross-sectional designs  |
| Randomised controlled trial                                | Studies in which subjects are randomly assigned to two or more groups, with the aim of assessing the outcomes associated with a specific intervention   |
| Case study/case series                                     | Observational study providing in-depth and detailed description and analysis of an individual or series of cases (involving people or organisations)  |
| Mixed methods evaluation                                   | Studies employing mixed qualitative and quantitative methods that evaluate the development or implementation of an intervention   |
| Economic evaluations                                       | Health economics studies including cost-effectiveness analysis, cost-utility analysis or cost-benefit analysis  |
| Evidence synthesis   | Secondary research, including literature reviews of all kinds (systematic, narrative, realist reviews and others)   |
| Other  | All other study designs not covered by these categories, itemised for each review below as required. These include, for example, decision analysis, consensus methods and other experimental study designs  |

included where they contributed relevant empirical data or theoretical perspectives that could be used to test or refine causal explanations in the form of 'context mechanism outcome configurations'.<sup>22,23</sup>

## 2.1 | Data collection

For each review, we collected data relating to the type and origin of each included document. All data were extracted and recorded independently by both CD and NR, to reduce the risk of errors. Disagreements were resolved by discussion, and if necessary, by consulting the lead authors of each included review.

### 2.1.1 | Types of document

For each review, we examined the title, abstract and, where required, the full text of each included document and categorised each in relation to its type. Documents reporting research were then further categorised according to study design. Wherever a document reporting research described mixed methods, we recorded the primary study design, that is, the method contributing the most data to the findings of the paper. However, where a document reported the findings from an explicitly mixed methods evaluation, this was recorded as a separate category. The final set of study design categories were developed iteratively, with new categories added as required to ensure coverage of the breadth of document types and study designs included in the set of reviews. The full list of categories used is given below in Table 2.

### 2.1.2 | Origin of documents

We gathered data on the origin of each included document by examining reviewers' reference collections (e.g.,

EndNote libraries), written methods, and personal records of searching processes. For each document, we aimed to identify:

- the date of retrieval and stage of searching/review,
- the method by which it was found,
- the identity of the member of the review team who retrieved it (e.g., the lead reviewer, or information specialist).

Where documents were identified by searching databases or other resources, we also recorded the name of the first source in which it was identified, that is, the first database or search engine in which it was identified as a potential document to be screened for inclusion. We confirmed the order in which databases were searched using our own records of the searching and de-duplication process followed in each review. Where the origin of any document was unclear, we consulted with reviewers themselves and asked for their recollection of the origin of documents. Wherever possible, we verified these recollections by repeating searches or related processes to re-identify documents. For example, if a reviewer indicated that they identified a record by 'snowballing' or consulting the reference list of another included document, we double checked the reference list for verification that this was possible and likely.

## 3 | RESULTS

### 3.1 | Included realist reviews

This comparison included five realist reviews, each focused on developing understanding of a different aspect of primary health care services. The main characteristics of the included reviews are outlined below in Table 3. The five review projects varied in their resource levels, ranging in team size from two to nine people, and in length from 18 months to 5 years. All but one review had access to an information specialist to design and run searches. All of

**TABLE 3** Characteristics of included realist reviews

| Review                             | Topic (in primary health care setting) | Information specialist involvement | No. of search phases | Total documents screened | Total included documents |
|------------------------------------|--|------------------------------------|----------------------|--------------------------|--------------------------|
| Review A (Abrams <sup>44</sup> )   | Early visiting services                | Yes                                | 3                    | 938                      | 70                       |
| Review B (Dixon <sup>48</sup> )    | Female genital mutilation              | Yes                                | 4                    | 4677                     | 124                      |
| Review C (Duddy <sup>45</sup> )    | Laboratory test ordering               | Yes                                | 3                    | 1583                     | 145                      |
| Review D (Grünwald <sup>47</sup> ) | Antipsychotic medication review        | No                                 | 4                    | 15,924                   | 55                       |
| Review E (Tierney <sup>46</sup> )  | Social prescribing link workers        | Yes                                | 4                    | 2638                     | 118                      |



the reviewers reported conducting multiple searches across various review stages, including background or scoping searches, systematic-style searches for empirical evidence and various additional searches in response to emergent information needs. The review teams screened and ultimately included varying numbers of documents.

### 3.2 | Types of included documents

The reviews included in this paper used data sourced from a wide variety of types of document. In four of the five included reviews, data extracted from published, academic research papers dominate. However, in one review (Review A), only 51% of included papers fell into this category,<sup>44</sup> and in one other (Review E) these documents made up only 24% of included documents.<sup>46</sup> All of the reviews included data extracted from numerous other types of document, in varying proportions. Notably, all of the reviews included some data drawn from opinion pieces or commentary (ranging from 2% to 12% of included documents), and two reviews included a small amount of material from social media platforms (Reviews A and E, with 6% and 1% of documents of this kind, respectively). A summary of the document types included in each review is shown below in Figure 1. For definitions of these document types, see Table 2.

#### 3.2.1 | Included study designs

We also collated information on the study designs employed in those included documents that reported research, either as published papers in academic journals,

or otherwise in reports (including evaluation documents). Amongst these documents, qualitative study designs, or mixed methods designs with qualitative elements contributed most data to the reviews, comprising between 23% (Review C) and 80% (Review E) of included documents. In most cases, these qualitative papers utilised interviews or focus groups to collect data, but other types of qualitative research (including ethnography and document analysis) were also represented. All of the reviews also drew on data generated by a variety of other methods, including case reports/series and quantitative research designs. In particular, all of the reviews included data extracted from documents reporting research that utilised surveys or questionnaires. In Reviews A, B, C and D, between 16% and 25% of documents reported the findings of this type of study.

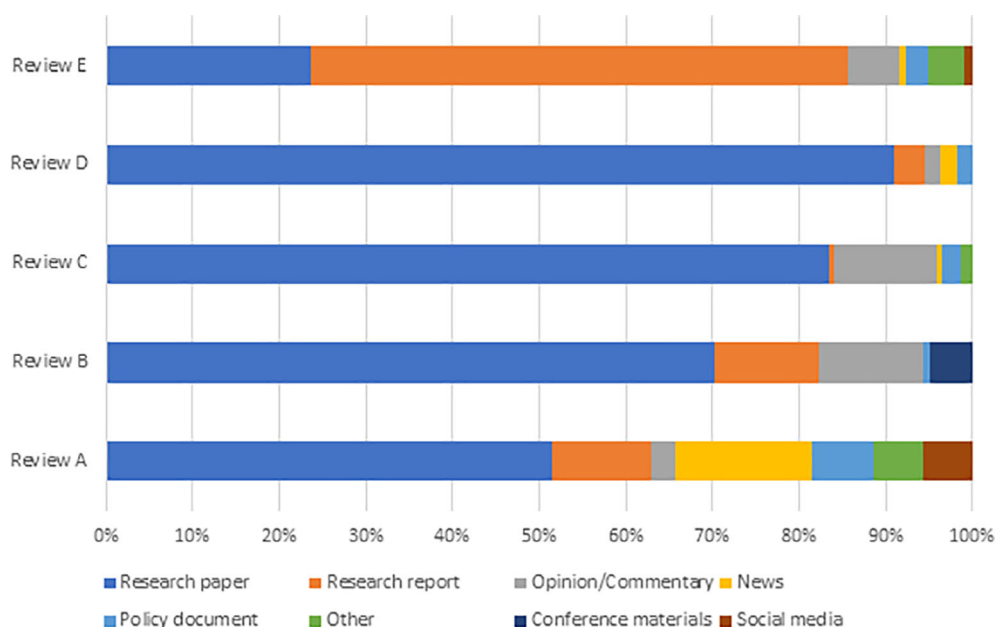
All of the included reviews also utilised data extracted from secondary research, with the proportion of included documents of this type ranging from 8% (Review E) to 33% (Review D). Evidence syntheses of all kinds, including quantitative, qualitative and mixed methods systematic reviews and narrative syntheses, contributed additional data to the reviews included in this study.

Figure 2 below summarises the study designs of the included documents that reported research studies.

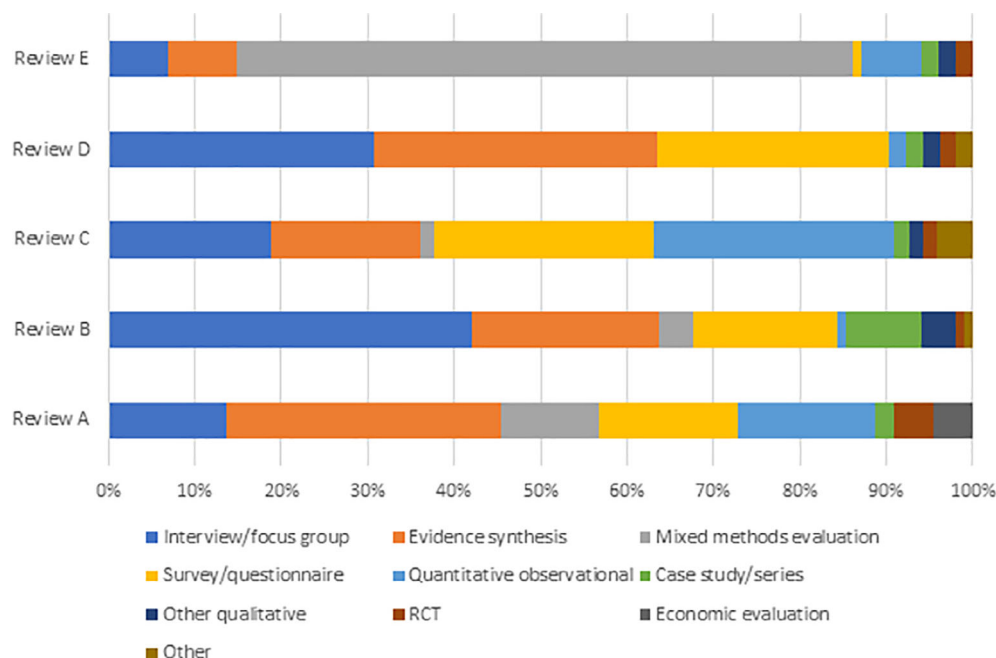
### 3.3 | Sources of included documents

#### 3.3.1 | Search phases

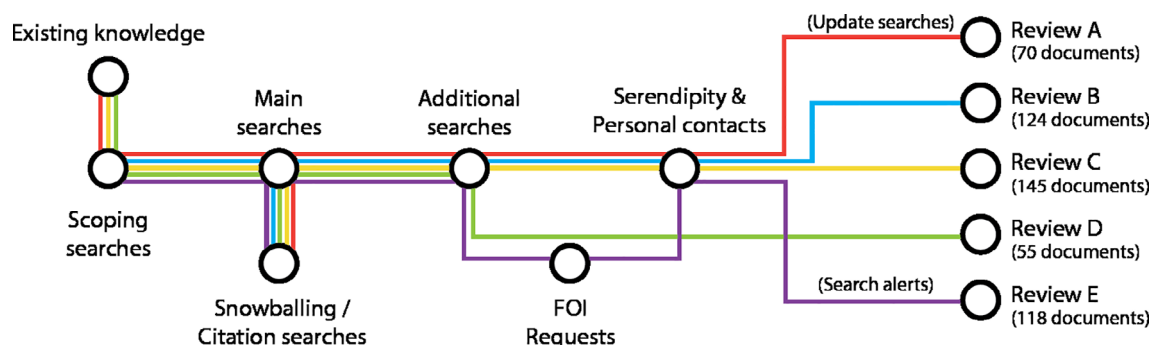
Searching was conducted across multiple stages in all five of the included reviews. The stages followed in each review are summarised below in Figure 3.



**FIGURE 1** Document types included in each review [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 2** Study designs employed in included documents reporting research [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 3** Searching phases in each review [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

All of the reviews included some form of initial, informal ‘scoping’ or ‘orienting’ search phase, during which reviewers gathered documents to help them to develop their understanding of the topic area, identify existing theoretical perspectives and/or inform initial programme theory development. In all but one of the reviews (Review A), some of the documents identified at this stage were ultimately included in the final review, that is, they contributed some data to programme theory development.

In all reviews, the majority of included documents were identified during the next searching stage, variously characterised as the ‘main’, ‘formal’ or ‘systematic’ phase of searching. The searches conducted at this stage bear the closest resemblance to the searching conducted in a ‘traditional’ systematic review: all reviews sought to gather relevant documents via a detailed, systematic-style search across multiple academic databases and other

information sources. However, in all of the reviews, significant numbers of documents were also added using additional, supplementary searching techniques, especially snowballing or citation tracking from the studies identified by the original search (see below).

All reviews employed an additional, later searching phase and/or other methods of gathering additional data following the main search phase. These ‘secondary’ or ‘iterative’ searches were used to identify additional data for programme theory development, and to identify documents containing data relating to existing substantive theories in the literature. These searches often had a different topic focus as reviewers sought to identify different (but related) material to that identified in the main searches. Four of the included reviews identified a small number of additional documents via personal contacts or networks, or serendipitously (e.g., while researching other topics). One review project included a separate

phase of update searches, rerunning the main search to identify more recently published material prior to submitting for publication.

### 3.3.2 | Search methods

Each review employed a range of searching methods to gather documents for inclusion. Figure 4 below summarises the proportion of documents obtained via different routes in each review.

In four of the five reviews, most documents contributing data were sourced via searching databases or search engines. The proportion obtained in this way ranged from 18% in Review E to 88% in Review A. Snowballing, or citation chaining (both backwards and in some cases, forwards) was also utilised in all of the included reviews to identify a significant proportion of additional included documents (between six and 44% of included documents across the five reviews). Other methods of obtaining documents that were employed in the reviews included drawing on existing knowledge and searching specific websites identified as likely to host relevant grey literature. In one review (Review E), Freedom of Information requests were employed to seek access to unpublished documents related to the intervention under study and this method identified a quarter of the documents contributing data to that review.<sup>46</sup> As noted above, in four of the included reviews, the reviewers reported that some additional documents were identified via personal contacts or entirely serendipitously (e.g., while searching for material or attending events related to other, unrelated projects).

### 3.3.3 | Search sources

For those documents retrieved via searches of databases and/or search engines (i.e., not by supplementary or additional search techniques), we identified the first source (database or search engine) of each. Although all of the reviews focused on the same broad topic area, and so many had sources in common, a slightly different set of sources was searched in each review, reflecting differences in each review's question and focus. Each project's resources and time available also influenced selection. Table 4 below summarises the sources searched in each review and indicates the number of included documents retrieved via searching each source and at different stages of the review. Note that the numbers presented here do not include documents retrieved via other methods (e.g., citation chaining).

Across all five reviews, and all searching phases, the sources that contributed the most documents for inclusion were MEDLINE (or PubMed), Google and Google Scholar. Searches in other databases contributed fewer documents; additional databases yielded between 8% (in Review A) and 45% (in Review B) of documents obtained via searching (between 7% and 31% of all included documents). The main exception may be in Review B, where searches in Embase and the Web of Science Core Citation Indexes contributed 12 (10% of the total) and 15 (12%) documents to the review, respectively.<sup>48</sup>

### 3.3.4 | Information specialist involvement

An information specialist (librarian) with expertise in developing comprehensive literature searches was

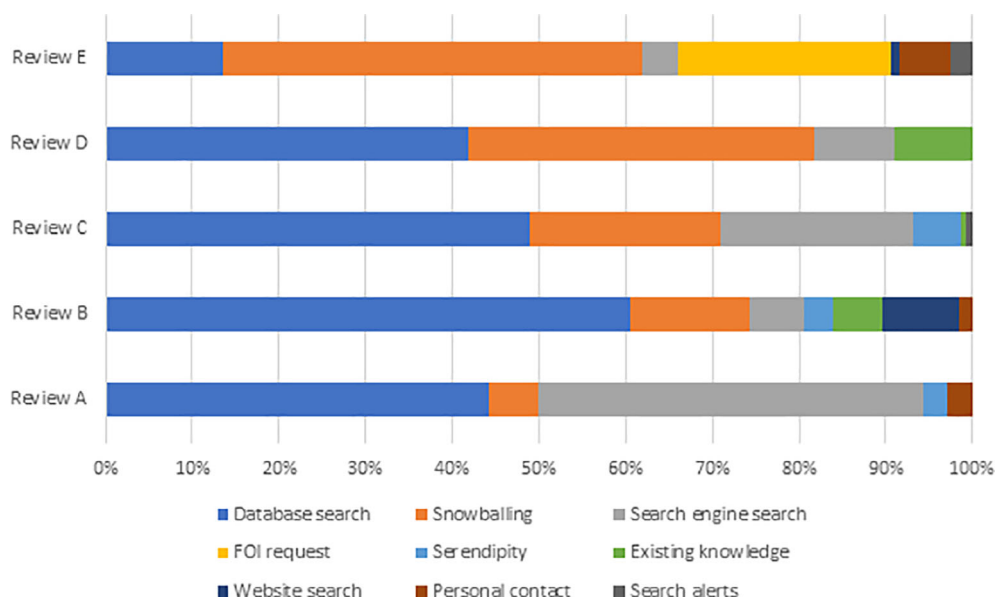


FIGURE 4 Documents sourced via different searching methods in each review [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



TABLE 4 Sources searched in each review and included documents retrieved from each

| Review                             | No (%) included documents retrieved via searching (databases or web) | Summary of scoping searches (databases or web) | Main searches (sources retrieving eight or more included documents across all reviews) |                |        |        |                 |   |  | Summary of additional searches (databases or web) |
|------------------------------------|--|--|--|----------------|--------|--------|-----------------|---|--|---|
|                                    |  |  |  |                |        |        |                 |   |  |   |
|                                    |  |  | MEDLINE  | Google Scholar | CINAHL | Scopus | Other databases |   |  |   |
| Review A (Abrams <sup>44</sup> )   | 62 (89%)   | 2 sources; 0 included documents                | 13   | 27             | 3      | 1      | 0               | 4 (Embase, HMIC, WoS)                           | 2 sources; 14 included documents                 |   |
| Review B (Dixon <sup>48</sup> )    | 83 (67%)   | 9 sources; 21 included documents               | 11   | 0              | 4      | 2      | 0               | 4 (ASSIA, Embase, WoS)                          | 7 sources; 41 included documents                 |   |
| Review C (Duddy <sup>45</sup> )    | 103 (71%)  | 2 sources; 15 included documents               | 49   | 1              | 17     | 4      | 3               | 11 (CENTRAL, EconLit, HMIC, NICE Evidence, WoS) | 2 sources; 3 included documents                  |   |
| Review D (Grünwald <sup>47</sup> ) | 28 (51%)   | 1 source; 0 included documents                 | 11   | 0              | 0      | 0      | 4               | 8 (Embase, PubMed, PsycINFO)                    | 1 source; 5 included documents                   |   |
| Review E (Tierney <sup>46</sup> )  | 21 (18%)   | 2 sources; 7 included documents                | 5  | 0              | 0      | 1      | 0               | 7 (ASSIA, HMIC, WoS)                            | No additional database or search engine searches |   |

Note: In reviews A, B, C and E, the searches in this database included the following indexes: SCI-EXPANDED, SSCI, A&HCI, CPCIL-S, CPCIL-SSH, BKCI-S, BKCI-SSH, ESCI. Abbreviation: WoS, web of science.

involved throughout four of the five included review projects. In one review (Review C), the lead reviewer herself is also an information specialist (C.D.).<sup>41,45</sup> Differences are apparent across the reviews: where there is greater reliance on documents retrieved via searching, the information specialist has contributed more to retrieval, and vice versa. Table 5 below summarises the proportion of documents retrieved by the information specialist, or by the lead reviewer in each case.

## 4 | DISCUSSION

### 4.1 | Summary of findings

This in-depth examination of five realist reviews has revealed several important similarities and differences across the included projects. All of the included reviews followed existing guidance and recommendations in their overall approach to searching. Although they used different terminology to describe different steps, all five projects included multiple stages of searching and employed appropriate methods in relation to the specific aims of each stage.<sup>23,24</sup>

Despite their focus on a common setting (primary health care), we found significant diversity in both the type of evidence included in this selection of reviews, and in the methods employed to identify that evidence. This finding echoes the results of previous reviews that have examined searching methods employed in realist reviews.<sup>1,38</sup> In this set of reviews, documents reporting the results of research and/or evaluations dominated, with a wide variety of qualitative, quantitative and mixed study designs contributing data to each. All of the included reviews also used data extracted from a variety of other types of documents, including those reporting opinion or commentary and news, and other forms of

TABLE 5 Documents retrieved by the information specialist versus reviewers

|                       | % Included documents retrieved by information specialist(s) | % Included documents retrieved by reviewer(s) |
|-----------------------|---|---|
| Review A              | 44%   | 56%   |
| Review B              | 61%   | 39%   |
| Review C <sup>a</sup> | 100%  | 100%  |
| Review D              | 0%  | 100%  |
| Review E              | 11%   | 89%   |

<sup>a</sup>In Review C, the lead reviewer was also an information specialist.

'grey literature', including policy documents, conference materials and social media posts. Across the reviews, most documents were identified by searching, including searching in bibliographic databases but also in Google Scholar and on the web. However, 'supplementary' or 'complementary' searching methods also made significant contributions to each review, contributing between 11% and 79% of documents across the set, with snowballing or citation chaining playing an important role.

## 4.2 | Implications for reviewers and information specialists

In this set of reviews, the strategies employed to identify relevant documents that could contribute data, and the type of those included documents, reflected the nature of the existing literature on the specific topics under investigation. For example, two of the included reviews (Reviews A and E) focused on assessing 'what works, for whom and in what circumstances?' for relatively new interventions in primary care. These reviews included the highest proportion of documents that could be termed 'grey literature' (49% and 76% of included documents, respectively), reflecting the fact that little formal research had been published on the relevant interventions when searching was conducted. Review A focused on 'early visiting services' and relied heavily on the results of web searches (using Google), while Review E explored social prescribing services and drew a significant proportion of its data from unpublished service evaluation documents obtained via Freedom of Information (FOI) requests.<sup>49</sup> Similarly, Review B focused on a topic area (female genital mutilation) where a significant amount of important research and outreach work has been conducted by third sector organisations. To ensure this 'grey' material was included, specific searches of relevant websites were conducted, and the availability and value of these documents is reflected in their inclusion in this review.<sup>48</sup> In these cases, the reviewers and information specialists involved in each project developed strategies to access documents containing relevant data based on their knowledge and understanding of the landscape of the available literature in each specific field.

The variety of document types included in this set of reviews, and especially their reliance on material beyond published academic journal articles, demonstrates the potential for these 'other' documents to contribute relevant data for realist programme theory development. The possible value of grey literature for systematic reviews has long been recognised, often as a means of increasing the exhaustiveness of traditional systematic reviews, especially by helping to ensure the

inclusion of recently published information (e.g., conference abstracts) or to reduce 'publication bias', on the assumption that studies with null or negative results that are less likely to be published in academic journals are more likely to be found elsewhere.<sup>3,50,51</sup> In realist reviews, grey literature may be more likely to contribute directly to the review's findings. The use of a wide range of document types reflects the nature of realist analysis, which may draw not only on the results of empirical studies of the intervention or phenomenon studied, but also on any relevant material that may help to elucidate underlying programme theories or provide data relating to the individual contexts, mechanisms and outcomes that form parts of the analysis. As is reflected in our analysis, this means that realist reviews may draw on data from very diverse materials, including editorials and commentaries, policy documents, reports, theses, social media posts and potentially many others, as well as from published research.<sup>22,24,36</sup>

Inclusiveness in document types naturally extends to study design and disciplinary origin of included research papers, and this apparent in our set of five reviews. Guidance for the conduct of realist reviews explicitly rejects any 'hierarchy of evidence' that might guide inclusion decisions in other review designs, and especially where such a hierarchy privileges the results of randomised controlled trials over other forms of evidence.<sup>52</sup> Instead, realist reviews are guided by the pragmatic principle outlined above, that is, that any study that can contribute relevant data to further the development or refinement of realist programme theory is welcome.<sup>36</sup>

We were unable to explore the specific contributions of each included document in this study. However we are aware that, in practice, different study designs may make different contributions to the analysis within a realist review. For example, documents reporting the results of quantitative study designs may be more likely to provide data relating to patterns of outcomes, while qualitative data may provide more detail on contexts and mechanisms. It is also worth bearing in mind that data may have been extracted from different parts of these included studies: the introductory and discussion sections of research papers and the researchers' own interpretations of their findings may all contribute to programme theory development. Notably, all five of the reviews in this study included some data drawn from other types of evidence syntheses, such as quantitative and qualitative systematic reviews and narrative reviews. The variety of quantitative, qualitative and mixed study designs that contributed data to all five of our reviews demonstrates the value of this inclusive approach in these projects, and that the search strategies employed did not aim to include or exclude any particular study designs.

Although we did not specifically examine study quality or the assessment of 'rigour' of the studies included in this set of reviews, it is worth bearing in mind that many realist reviewers recognise the potential for 'methodologically weak' studies to contribute useful data. The needs of the realist analysis and development of programme theory are always the overriding principle governing the selection of documents.<sup>37</sup>

The diversity of material gathered in the reviews is mirrored by their use of a range of searching methods. As noted above, all five reviews relied on strategies beyond bibliographic database searching to identify relevant documents. The use of multiple data gathering methods, and especially the reliance on iterative techniques like snowballing or citation chaining corresponds with the range of techniques that have been demonstrated to be useful in other reviews of complex evidence.<sup>18–21</sup>

While bearing in mind the importance of 'supplementary' or 'complementary' techniques, it is clear that systematic searching of bibliographic databases during the 'main' searching phase was a key component of the overall approach in each review. The extent of this searching varied across the reviews, both in terms of the exhaustiveness of the search strategies employed, and in terms of the number of databases searched. We did not conduct detailed examination of the searches themselves for this study, but note that there was wide variation in the ratio of documents included against documents identified (from 1.2% to 9.2%), implying potentially different approaches in relation to the sensitivity and specificity of search strategies, and therefore to the workload involved in screening. These differences may reflect both differences in the resources available to devote to searching in each review, and the personal preferences of the review teams involved. The approach taken may also reflect the availability of literature and the nature of the topic area, especially the ease with which it can be described in search terms. The relative fruitfulness of bibliographic database searching in each review also reflects the landscape of the available literature in each specific topic area: prior knowledge of this landscape, informed by earlier scoping or background searches, can help to inform the most efficient approach for a particular project.

Unsurprisingly, as the five reviews all focused on primary health care settings, there were similarities in the sources searched and these usually involved major health-focused bibliographic databases (e.g., MEDLINE, Embase, CINAHL) and additional databases with broader, social science coverage (e.g., Scopus, components of Web of Science, ASSIA). Interestingly, most reviews also searched using Google and/or Google Scholar which was a source of a significant minority of documents across the reviews. Other research has affirmed the value of Google and

Google Scholar as useful additional sources for identifying grey literature in particular.<sup>53,54</sup> The relative fruitfulness of these web-based resources may reflect the need for grey literature in the included set of reviews, but may also be explained by ease of searching or the ability of the search engines' algorithms to present highly relevant results on the first few pages.

It is notable that where documents were identified and included via searching in the reviews, the majority were sourced via a relatively small number of sources searched (see Table 4). The dominant contribution of MEDLINE in each review reflects the fact that it was the first source searched in the main searching phase of each review. The relative fruitfulness of these resources suggests there may be diminishing returns to extensive database searching, raising important questions about the efficiency of comprehensive searching across multiple bibliographic databases. This approach is familiar from searching for traditional systematic reviews, where there is an emphasis on exhaustiveness. It may be less appropriate for realist reviews where there is a greater emphasis on theoretical saturation, and where additional purposive searching may be conducted at later stages to help fill gaps and identify further data for programme theory development as required. However, it is possible that more limited approaches to searching in these reviews would have failed to identify documents that contained crucial data for programme theory development. Review teams may also consider approaches to searching that cover more ground, aiming to identify different types of documents and material from multiple disciplines, which may be more likely to yield important disconfirming data, or new insights and theoretical perspectives.

Finally, we note that the relative complexity of searching processes in realist reviews necessitates excellent record keeping within review teams. In particular, conducting multiple searches across several stages, a greater reliance on informal and iterative methods such as citation chaining, and the use of web search engines like Google Scholar all pose challenges for transparency in reporting. Setting up thorough processes to record the origin of selected documents at the outset of realist review projects is crucial, and all members of the review team who are involved in searching or gathering data should adhere to these. We were fortunate in this study that the review teams for each project had kept detailed records of their activities, including, for example, adding information relating to retrieval to document records in reference libraries, and keeping diary-style notes on informal searches. In making this recommendation, we echo calls from other information specialists and review methodologists for increased transparency in reporting of searching processes.<sup>1,53</sup>

### 4.3 | Strengths and limitations of this study

This study confirms and adds detail to the findings of two existing reviews that have identified variation in the searching processes used in realist reviews.<sup>1,38</sup> It also extends this work by providing an in-depth examination of the types of documents identified and included in a set of five completed realist reviews, identifying significant diversity in the type of material that can contribute data for realist programme theory development. Our thorough examination of this set of reviews was strengthened by our relationships with the review teams involved, who provided access to their reference libraries and associated documentation. All of the data presented here were extracted in duplicate by CD and NR and missing data and disagreements were resolved by discussing with the lead reviewers from each project. We have produced a series of important implications for reviewers and information specialists based on our examination of the reviews in this series.

There are some important limitations to the present study. We were directly involved all five included reviews, and so were able to influence the conduct and recording of the searching activities undertaken. We have made efforts in this paper to transparently report the different approaches adopted in each project, but we acknowledge that our involvement is likely to have resulted in similarities in approach. An extension of the present study to evaluate and compare approaches to gathering data across a larger set of reviews, produced by other research groups, would provide a broader view of practice.

The data presented above focus at the level of the documents included in each review, but this does not provide information about the individual pieces of data extracted from these documents that were used to test or refine realist programme theories. Due to large gaps in the data relating to the date on which each document was retrieved, we were unable to present information on the types of documents retrieved at each stage of searching in each review. A prospective study, introducing processes for collecting such data from the beginning of a review project, would be able to present more detail on these areas. With more resources, the present study could have been extended in several interesting directions. For example, we were unable to collect data relating to the time spent on searching and other data gathering methods during the reviews and so are unable to present any findings relating to the efficiency or costs of these processes. Future research might also consider the development of search strategies for use in realist reviews in more detail; interviewing reviewers and information specialists could provide a better understanding of how searching is undertaken within realist review projects.

## 5 | CONCLUSIONS

There were important similarities and differences in the approaches used to gather data for inclusion in the five realist reviews examined in this study. The similar overall approach, involving multiple searches conducted in stages during the review project, reflects existing guidance on the conduct of realist reviews. Within this, a diversity of techniques was employed, reflecting the individual review projects' resources and, importantly, the landscape of the existing literature in each specific topic area. Review teams and information specialists should bear in mind the need for diversity and flexibility in approaches to data gathering for realist reviews. Searches conducted in multiple stages over the life of a review project may necessitate ongoing collaboration and responsive strategies to seek documents likely to contain data that can be used in programme theory development. Where resources are limited, careful consideration of the strategies that are likely to be most fruitful in identifying relevant documents will be required. Each individual review project is unique; the approach taken to gather data should reflect each project's individual needs, and take into account the nature of the existing literature on the topic in question. The potential complexity of searching processes employed across multiple stages in a review necessitates excellent record keeping and accurate presentation of how data were gathered over the course of a review to ensure transparency in reporting.

### ACKNOWLEDGMENTS

We are extremely grateful to the lead reviewers of the realist review projects included in this study, who generously contributed their files and notes with us, and were happy to answer our questions about their review methods and included documents. While this study was undertaken, C.D. was supported by an NIHR Systematic Reviews Fellowship (NIHR-RM-SR-2017-08-018).

### CONFLICT OF INTEREST

The authors declare no conflicts of interest.

### AUTHOR CONTRIBUTIONS

CD conceived the study and CD and NR developed the project. Both authors contributed to data collection, extraction and analysis, and writing the manuscript.

### DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.



## ORCID

Claire Duddy  <https://orcid.org/0000-0002-7083-6589>

Nia Roberts  <https://orcid.org/0000-0002-1142-6440>

## REFERENCES

- Booth A, Briscoe S, Wright JM. The “realist search”: a systematic scoping review of current practice and reporting. *Res Synth Methods*. 2020;11(1):14-35.
- Wong G. The rise and rise of realist reviews? In *Centre for Evidence-Based Medicine Blog*; 2019.
- Lefebvre C, Glanville J, Briscoe S, et al. Chapter 4: searching for and selecting studies. In: JPT H, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 6.1 (Updated September 2020)*, Cochrane; 2020.
- Centre for Reviews and Dissemination. *Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care*. CRD, University of York; 2009.
- Eden J, Levit L, Berg A, Morton S. In: Eden J, Levit L, Berg A, Morton S, eds. *Standards for Finding and Assessing Individual Studies*. National Academies Press (US); 2011.
- Sterne JAC, Egger M, Moher D. Chapter 10: addressing reporting biases. In: JPT H, Green S, eds. *The Cochrane Collaboration*; 2011.
- Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. *Cochrane Database Syst Rev*. 2009;1:MR000006.
- Hopewell S, Clarke MJ, Stewart L, Tierney J. Time to publication for results of clinical trials. *Cochrane Database Syst Rev*. 2007;2:MR000011.
- Horsley T, Dingwall O, Sampson M. Checking reference lists to find additional studies for systematic reviews. *Cochrane Database Syst Rev*. 2011;8:MR000026.
- Dwan K, Gamble C, Williamson PR, Kirkham JJ. Systematic review of the empirical evidence of study publication bias and outcome reporting bias—an updated review. *PLoS One*. 2013;8(7):e66844.
- Dwan K, Altman DG, Cresswell L, Blundell M, Gamble CL, Williamson PR. Comparison of protocols and registry entries to published reports for randomised controlled trials. *Cochrane Database Syst Rev*. 2011;1:MR000031.
- Morrison A, Polisena J, Husereau D, et al. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. *Int J Technol Assess Health Care*. 2012;28(02):138-144.
- Halladay CW, Trikalinos TA, Schmid IT, Schmid CH, Dahabreh IJ. Using data sources beyond PubMed has a modest impact on the results of systematic reviews of therapeutic interventions. *J Clin Epidemiol*. 2015;68(9):1076-1084.
- Schoones JW, Vissers T. Re: “using data sources beyond PubMed has a modest impact on the results of systematic reviews of therapeutic interventions”. *J Clin Epidemiol*. 2016;78:127-127.
- Booth A. Over 85% of included studies in systematic reviews are on MEDLINE. *J Clin Epidemiol*. 2016;79:165-166.
- Savoie I, Helmer D, Green CJ, Kazanjian A. Beyond Medline: reducing bias through extended systematic review search. *Int J Technol Assess Health Care*. 2003;19(1):168-178.
- Doree C, Elke H, Mathsien M, Waffenschmidt S. Value of using different search approaches. *Health Technology Assessment International (HTAi) Vortal*; 2017.
- Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *BMJ*. 2005;331(7524):1064-1065.
- Booth A, Harris J, Croot E, Springett J, Campbell F, Wilkins E. Towards a methodology for cluster searching to provide conceptual and contextual “richness” for systematic reviews of complex interventions: case study (CLUSTER). *BMC Med Res Methodol*. 2013;13(118).
- Booth A. Searching for qualitative research for inclusion in systematic reviews: a structured methodological review. *Syst Rev*. 2016;5(74).
- Campbell M, Egan M, Lorenc T, et al. Considering methodological options for reviews of theory: illustrated by a review of theories linking income and health. *Syst Rev*. 2014;3(1):114-114.
- Wong G, Westhorp G, Pawson R, Greenhalgh T. Realist synthesis: RAMESES training materials; 2013.
- Pawson R. *Realist synthesis: New Protocols for Systematic Review*. SAGE; 2006:73-104.
- Booth A, Wright J, Briscoe S. In: Emmel N, Greenhalgh J, Manzano A, Monaghan M, Dalkin S, eds. *Scoping and Searching to Support Realist Approaches*. SAGE; 2018:148-165.
- Hanneke R, Young SK. Information sources for obesity prevention policy research: a review of systematic reviews. *Syst Rev*. 2017;6(1):156-156.
- Armstrong R, Waters E. Systematic reviews of health promotion and public health interventions (Presentation); 2007.
- von Hippel E, Franke N, Prügl R. Pyramiding: efficient search for rare subjects. *Res Pol*. 2009;38(9):1397-1406.
- Delaney A, Tamás PA. Searching for evidence or approval? A commentary on database search in systematic reviews and alternative information retrieval methodologies. *Res Synth Method*. 2017;9:124-131.
- Howes F, Doyle J, Jackson N, Waters E. Evidence-based public health: the importance of finding ‘difficult to locate’ public health and health promotion intervention studies for systematic reviews. *J Public Health*. 2004;26(1):101-104.
- Armstrong R, Jackson N, Doyle J, Waters E, Howes F. It's in your hands: the value of handsearching in conducting systematic reviews of public health interventions. *J Public Health*. 2005;27(4):388-391.
- Grayson L, Gomersall A. A difficult business: finding the evidence for social science reviews; 2003.
- Ogilvie D, Hamilton V, Egan M, Petticrew M. Systematic reviews of health effects of social interventions: 1. Finding the evidence: how far should you go? *J Epidemiol Community Health*. 2005;59(9):804-808.
- Ogilvie D, Egan M, Hamilton V, Petticrew M, Ogilvie D. Systematic reviews of health effects of social interventions: 2. Best available evidence: how low should you go? *J Epidemiol Community Health*. 2005;59:886-892.
- NIHR School of Primary Care Research. Evidence Synthesis Working Group; 2018. <https://www.spcr.nihr.ac.uk/eswg/evidence-synthesis>
- Heneghan C. Improving the evidence base for primary care using evidence synthesis. In *NIHR School of Primary Care Research Blog*; 2018.



36. Wong G. Data gathering in realist reviews: looking for needles in haystacks. In: Emmel N, Greenhalgh J, Manzano A, Monaghan M, Dalkin S, eds. *Doing Realist Research*. SAGE; 2018:131-146.
37. Pawson R. Digging for nuggets: how 'Bad' research can yield 'Good' evidence. *Int J Soc Res Methodol*. 2006;9(2):1364-5579.
38. Berg RC, Nanavati J. Realist review: current practice and future prospects. *J Res Pract*. 2016;12(1).
39. Abrams R, Wong G, Mahtani KR, et al. Understanding the impact of delegated home visiting services accessed via general practice by community-dwelling patients: a realist review protocol. *BMJ Open*. 2018;8(11):e024876.
40. Dixon S, Griffiths F, Ziebland S, Papoutsi C. English general medical practitioners managing female genital mutilation in light of recent policy and legislative changes: a realist synthesis. *Prospero*. 2018;CRD42018091996.
41. Duddy C, Wong G. Explaining variations in test ordering in primary care: protocol for a realist review. *BMJ Open*. 2018;8(9):e023117.
42. Grünwald L, Duddy C, Moncrieff J, Crellin N, Byng R. Facilitators and barriers of antipsychotic medication management in primary care: a realist review protocol. *Prospero*. 2018;CRD42018107573.
43. Tierney S, Wong G, Roberts NW, et al. The role of care navigators in primary care: a realist review (protocol). *Prospero*. 2018;CRD42018095090.
44. Abrams R, Wong G, Mahtani KR, et al. Delegating home visits in general practice: a realist review on the impact on GP workload and patient care. *Br J Gen Pract*. 2020;70(695):e412-e420.
45. Duddy C, Wong G. Efficiency over thoroughness in laboratory testing decision-making in primary care: findings from a realist review. *BJGP Open*. 2020;5(2):bjgpopen20X101146.
46. Tierney S, Wong G, Roberts N, et al. Supporting social prescribing in primary care by linking people to local assets: a realist review. *BMC Med*. 2020;18(1):49.
47. Grünwald LM, Duddy C, Byng R, Crellin N, Moncrieff J. The role of trust and hope in antipsychotic medication reviews between GPs and service users: a realist review. *BMC Psychiatry*. 2021;21(390).
48. Dixon S, Duddy C, Harrison G, Papoutsi C, Ziebland S, Griffiths F. Conversations about FGM in primary care: a realist review on how, why, and under what circumstances FGM is discussed in general practice consultations. *BMJ Open*. 2021;11:e039809.
49. Tierney S, Abrams R. Getting the most from freedom of information requests as part of a realist review: finding treasure or futile endeavour? In: Research NSoPC, ed. *NIHR School of Primary Care Research Blog*; 2019.
50. Paez A. Gray literature: an important resource in systematic reviews. *J Evid Based Med*. 2017;10(3):233-240.
51. Mahood Q, Van Eerd D, Irvin E. Searching for grey literature for systematic reviews: challenges and benefits. *Res Synth Methods*. 2014;5(3):221-234.
52. Pawson R. Systematic obfuscation: a critical analysis of the meta-analytic approach. *Evidence-Based Policy: A Realist Perspective*. SAGE; 2006:38-72.
53. Haddaway NR, Collins AM, Coughlin D, Kirk S. The role of Google scholar in evidence reviews and its applicability to Grey literature searching. *PLoS One*. 2015;10(9):e0138237.
54. Coleman S, Wright JM, Nixon J, Schoonhoven L, Twiddy M, Greenhalgh J. Searching for Programme theories for a realist evaluation: a case study comparing an academic database search and a simple Google search. *BMC Med Res Methodol*. 2020;20(1):217.

**How to cite this article:** Duddy C, Roberts N. Identifying evidence for five realist reviews in primary health care: A comparison of search methods. *Res Syn Meth*. 2022;13(2):190-203. doi: 10.1002/jrsm.1523