

1 A Review of Evaluation Guidance Relevant to Public Health Interventions.

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

Sarah Denford¹ S.Denford@Exeter.ac.uk

Charles Abraham¹ C.Abraham@Exeter.ac.uk

Margaret Callaghan¹ M.Callaghan@Exeter.ac.uk

Peter Aighton¹ Peter.Aighton@avonandsomerset.police.uk

Frank DeVocht² Frank.devocht@bristol.ac.uk

Steven Ariss³ S.Ariss@Sheffield.ac.uk

1 Psychology Applied to Health group, Institute of Health Research, University of Exeter Medical School, Exeter, UK

2 School of Social and Community Medicine, University of Bristol, Bristol, UK

3 University of Sheffield, Sheffield, UK

Corresponding author:

Sarah Denford

2.09 College house,

St Lukes Campus,

University of Exeter.

EX1 2LU

S.Denford@Exeter.ac.uk

01392 725906

29

30 **Abstract**

31 *Objective: To identify accessible high-quality, evaluation guidance, available to researchers and practitioners*
32 *and to catalogue, summarise and categorise the content of a subset of accessible, quality guides to evaluation.*

33 *Methods: We systematically reviewed the literature for documents providing support for evaluation of complex*
34 *health interventions.*

35 *Results: Our initial search identified 402 unique documents that were screened to identify those that were (1)*
36 *developed by or for a national or international organization (2) freely available to all (3) published during or*
37 *after 2000 (4) specific to public health. This yielded 98 documents from 43 organisations. Of these, 48 were*
38 *reviewed in detail. This generated a detailed catalogue of quality evaluation guidance.*

39 *Conclusions: A wide range of guidance on evaluation of public health initiatives is available. Time and*
40 *knowledge constraints may mean that busy practitioners find it challenging to access the most, up-to-date,*
41 *relevant and useful guidance. This review presents links to and reviews of 48 quality guides to evaluation as*
42 *well as categorising their content. This facilitates quick and each access to multiple selected sources of specific*
43 *guidance.*

44

45

46

47 **Background**

48 Evaluation is foundational to identification, implementation and dissemination of effective
49 and cost effective interventions. There have been many calls to ensure and improve
50 evaluation of interventions and initiatives designed to improve public health. For example,
51 the US Centre for Disease Control and Prevention note that in order to improve the health
52 of the public, *“we must devote our skill - and our will – to evaluating the effects of public
53 health actions”* (Centres for Disease Control and Prevention, 1999). In the UK, the 2011
54 House of Lords report on Behaviour change recommend that *“a lot more could and should,
55 be done to improve the evaluation of interventions* (House of Lords Select Committee, 2011)
56 and Public Health England have called for improved evaluation; stating that *“whilst
57 interventions are being commissioned by a variety of organisations, data informing the
58 relative ‘success’ of the interventions, in terms of the intended health outcomes, was patchy
59 and inconsistent”* (Public Health England, 2016). Evaluation is essential to understanding
60 what does and does not work and robust demonstration of effectiveness may be crucial to
61 securing future funding.

62 Evaluation of public health interventions may be complex and expensive, in part, because
63 such interventions are themselves complex (Craig et al., 2013). Public health interventions
64 may, for example, attempt to engage and multiple individuals, organisations, and / or
65 communities and target changes in knowledge, attitudes, and practices at multiple levels
66 over long periods of time. They also operate in complex demographic and socio-economic
67 contexts (Craig et al., 2013). These “real life” complexities render evaluation challenging.
68 Nonetheless, when financial resources are scarce, it is crucial that evidence of effectiveness
69 directs selection of interventions. Repeatedly funding ineffective programs not only
70 squanders valuable resources but has opportunity costs resulting from failure to implement
71 potentially beneficial interventions. Without good evaluation, it is impossible to distinguish
72 between programs that are having a substantial health impact, those that need to be
73 adapted for different populations, those that need to be withdrawn, and those that may be
74 harmful (UK National Institute for Health and Care Excellence, NIHR, 2010). Despite this,
75 programs are often implemented with poor, incomplete or no evaluation.

76 Evaluability assessment facilitates decisions by practitioners, commissioners and
77 researchers about what programs most need to be evaluated. Ogilvie et al. (2011) identify

78 five key questions that should be answered before investing in an evaluation. These focus
79 on the stage of development or intervention implementation, whether or not the results of
80 the evaluation are likely to lead to changes in policy or practice, how widespread or
81 important effects of an intervention are likely to be (i.e., is it likely to have a large effect on a
82 large number of people), and how will findings of the evaluation contribute to existing
83 evidence? Such guidance enables evaluation priorities to be identified when funds for
84 evaluations are limited.

85 When evaluation is undertaken it is vital that it is conducted in a manner that will produce
86 robust answers to the questions addressed. This requires expertise and, not all practitioners
87 may have had adequate training to enable them to undertake evaluations without support.
88 Consequently, public health practitioners and commissioners may feel under-skilled to
89 conduct evaluations (Denford et al., 2016). This highlights the need for high-quality, useable,
90 feely available and practical evaluation guidance on how to evaluate public health
91 programs.

92 Appeals for better evaluation practice led to a proliferation of guidance and advice to
93 support evaluation of public health programs (Ogilvie et al., 2011; Craig et al., 2013). For
94 example, the US Centres for Disease Control and Prevention (CDC) developed a framework
95 to guide public health professionals perform evaluations. The framework includes a series of
96 steps; from engaging stakeholders through to dissemination of findings. Many of these
97 activities are part of routine practice but others are not. This framework includes a series of
98 standards that are intended to ensure that evaluations are well designed, rigorous, and
99 suitable for purpose. Different guidance is provided by the revised UK Medical Research
100 Council (MRC) framework (Craig et al., 2013), aimed mainly at academic researchers. This
101 provides an overview of the phases and processes involved in the development,
102 implementation and evaluation of complex interventions and, in contrast, to an earlier MRC
103 framework (MRC 2000), this revised guide provides advice on how to evaluate highly
104 complex programs using a variety of methods – not just controlled clinical trials. These are
105 just two of many guides to public health evaluation (see too, for example, Centres for
106 Disease Control and Prevention, 1999; Better Evaluation, 2002; Moore et al., 2015; Cavil,
107 Roberts, & Rutter 2012).

108 Such guidance has the potential to facilitate the quality of evaluations and increase the
109 number of programs with strong evidence for their effectiveness which, in turn would allow
110 withdrawal of programs that are ineffective or lack evidence of effectiveness. However, to
111 be useful such guidance has to be accessible to practitioners and commissioners; they need
112 to know which guides to evaluation to use for what purpose. Unfortunately, the many
113 different guides available may be a barrier to identification of relevant guidance. Novice
114 evaluators who turn to the internet for guidance are faced with many choices and have no
115 map describing the content of available evaluation and so may find it difficult to know what
116 guidance to follow. Some guides provide generic advice on evaluation generally, others on
117 particular types of evaluation (e.g., process or economic evaluation). Some documents are
118 written for academics, or policy makers, or funders, or experienced evaluators. Other guides
119 are topic specific; such as the UK Public Health England frameworks for obesity prevention.
120 This can be overwhelming for novice evaluators who have no easy way to select appropriate
121 high-quality guidance for particular evaluation projects (Denford et al). This lack of guidance
122 on how to access and use guidance on public health guidance is clear from conversations
123 with practitioners. For example, in a qualitative study of UK public health practitioners’
124 views of evaluation on practitioner commented, *“If I were to begin evaluation tomorrow and
125 I did a search on evaluation I’d probably... come up with about 500 hits - but actually they
126 aren’t all equal and some are more appropriate than others and understanding which is the
127 best one to use would be difficult for me. So some advice about quality and/or types of tools
128 or particular types of evaluation would be helpful”*

129 ***The Present Study***

130 We aimed to review this literature in order to assess the extent of available guides to
131 evaluation relevant to public health interventions and to identify the content of such
132 guidance. We also planned to provide a catalogue of high-quality, readily-accessible guides
133 that would help practitioners navigate this literature. We had two specific aims.

134 ***Specific Aims***

- 135 1. To identify accessible high-quality, evaluation guidance, available to public health
136 researchers and practitioners
- 137 2. To summarise and categorise the content of a subset of accessible, quality guides to
138 evaluation.

139 **Methods**

140 ***Search strategy***

141 Evaluation guidance documents were identified using five strategies: (1) searching
142 electronic databases (2) hand searching of identified guides and journals (3) searching
143 internet resources (4) citation searching and (5) contacting key authors and professionals in
144 the field.

145 EMBASE, MEDLINE, MEDLINE-in-process, Health management information consortium
146 (HMIC), Social Policy and Practice (SSP), Web of Science, and PsycINFO were searched using
147 the search strategy presented in Supplementary Document 1. Reference lists of identified
148 papers, and key journals were also scrutinised.

149 It was anticipated that a substantial proportion of guidance documents would not take the
150 form of academic papers and so would not be identified using a traditional literature search.

151 To identify as many non-academic documents as possible, we searched the four main
152 internet search engines (Google, BING, Yahoo, WebCrawler) using a modification of the
153 search strategy (Supplement 1). The first 30 pages that were retrieved from each database
154 using each term were screened. We then searched a series of health and evaluation
155 websites (Supplement 1) using the term “evaluation.” Websites were identified through
156 discussions with academics and public health professionals. Finally, key authors and experts
157 in the field were asked to suggest documents/ websites/ policies that we missed.

158 The search was conducted by two authors (author one and three) with guidance and
159 support from author 2. We included websites, books, journal articles, policy
160 recommendations, educational resources, tools and frameworks that provide support to
161 public health practitioners undertaking evaluations of public health interventions.

162 We excluded all documents and articles in which the aim was to evaluate a specific
163 intervention (as opposed to offering advice on how to undertake evaluations). Papers
164 reporting on the development of questionnaires or assessment scales for specific
165 interventions were also excluded.

166 ***Selecting a Subset of Evaluation Guidance Documents***

167 Our search identified 402 guidance documents that can be used to support the evaluation of
168 public health programs. This list included books, reports, webpages, and academic articles.
169 These documents had a range of aims including: (i) elucidating the principles that
170 practitioners should follow when conducting evaluations, (ii) specifying standards of good
171 evaluation practice, (iii) offering advice on how and when to undertake evaluations, (iv)
172 offering instruction on particular type of evaluation e.g., economic or process evaluations,
173 (v) providing online support and advice to practitioners, including tools to support
174 evaluation.

175 From this list of 402 documents, we selected a subset to review and catalogue in detail.
176 Through discussion the authors agreed four selection criteria. First, that documents were
177 free and readily available to public health practitioners. Second, given the changing nature
178 of public health practice, we focused on documents written in or after the year 2000. Third,
179 to provide a quality indicator we selected documents sourced or created by national or
180 international organisations. Fourth, and finally, we assessed the relevance of each
181 document to public health.

182 This resulted in a reduced list of 98 documents produced by 25 organisations. We then
183 reviewed the documents produced by each of these organisations and selected the most
184 comprehensive or recent evaluation guidance, resulting in 25 guides. A further 23 guides
185 were added because the authors agreed that documents provided by the US Centers for
186 Disease Control and Prevention, the UK Medical Research Council, the UK National Health
187 Service, and Public Health England were complementary to the initial selection of 25 and
188 judged to be equally valuable to practitioners. So, for these four organisations, more than
189 one guide to evaluation was retained for detailed examination. The final list of these 48
190 evaluation guidance documents is provided in Supplementary Document 2.

191 ***Summarising the Content of Selected Guides to Evaluation***

192 Each of the 48 guides were read by two authors. A content template was developed through
193 discussion and short one-page summaries of each of the 48 guides were produced
194 independently by each reviewer. Each summary provided information on the target
195 audience of the guide, its main aim, a short overview of the guide, and strengths and
196 limitations. Links to the resource and associated resources were also included. The two
197 reviews of each guide were then combined, retaining content from each review and

198 resolving any discrepancies through discussion. The final (integrated) review was then sent
199 to the original author of the document for verification. Authors were asked if they (i)
200 considered the summary to be an accurate and good reflection of their document (ii) if
201 anything was missing and (iii) if they thought this would be useful to practitioners.

202 ***Content Categorisation of Evaluation Guidance Documents***

203 Once summaries had been completed, two authors independently coded the content of the
204 48 guides. Initially, a selection of 5 documents were read by each reviewer and their content
205 listed. Merging of these two content lists through discussion resulted in a set of 44 content
206 categories. The two authors then jointly coded a further 5 documents, and discussed and
207 refined the list accordingly. The final list contained 37 content categories generated by the
208 initial two-stage coding of 10 guides; which was then used to categorise the content of the
209 remaining 38 guides.

210 **Results**

211 ***Overview of 98 Guides to Evaluation***

212 While we identified 402 guides to evaluation only 98 were relevant to public health, free
213 and readily available and produced since 2000 by a national or international organisation.
214 These 98 varied in terms of purpose, topic or condition, and audience. The large majority
215 were general overviews of evaluation focusing on principles of evaluation and how to assess
216 evidence to support evaluation. A variety of evaluations were considered including, trials,
217 naturally-occurring experiments, process evaluations and economic evaluations and various
218 advice and instructions were provided on how to plan prepare for and conduct such
219 evaluations. The guides were mainly generic but some focused specifically on evaluating
220 international development, obesity, asthma, sexual health, mental health and physical
221 activity. One focused on children and families, one on healthy eating and one on drugs and
222 one on violence on women and girls. However, even condition-specific guides provided
223 instruction and support that relevant to evaluations of other types of interventions. For
224 example, some guides, focused on evaluating interventions relevant to particular health
225 problem (asthma, smoking and obesity) to illustrate more general lessons. Documents
226 ranged from targeting those with no knowledge of evaluation, to sophisticated guides for
227 those experienced in evaluation practice with target audiences including novices, evaluation

228 experts, program managers, health care professionals, government officials and
229 academics/researchers.

230 ***Response from authors***

231 Of the 48 documents we reviewed, copies of our review were sent to the lead authors of
232 each guide. If we were unable to find contact details for the main author, or there was no
233 named author, we contacted the chair of a research group or a general enquiries email. In
234 some cases contacts passed us on to someone else. We sent reminders up to three times,
235 in some cases trying alternative addresses. Fourteen authors did not respond. Two
236 organisations (representing a total of seven documents) sent standard replies stating that
237 they did not respond to such requests. Twenty seven authors replied to state that they were
238 happy with the summary or to suggest minor changes

239 ***Content Categorisation***

240 Merging categories identified by two researchers independently coding the content of two
241 sets of five guides resulted in a list of 37 content categories that were used to describe the
242 content of 48 selected guides. These categories were grouped into 1) Background to
243 evaluation 2) Pre-evaluation preparatory work 3) The Evaluation process 4) Types of
244 evaluation and 5) Additional support. Tables 1-5 list content categories within each of these
245 five groupings. The tables also lists each of the 48 documents that contained content
246 corresponding to each of the 37 categories. Thus these tables can be used by readers to find
247 guides (among the 48) with particular content. These tables can be used in conjunction with
248 the 48 brief summaries of these guides which are provided in Supplementary Document 3.
249 Each of the 48 guides was been given a brief title that is listed below and used both in
250 Tables 1-5 and in Supplementary Document 3.

251 ***1. Background to evaluation***

252 Seven content categories were grouped as “background to evaluation” (see Table 1). These
253 were: Evaluation overview; assessing the evidence; evidence based practice; evaluability;
254 common evaluation challenges; policy and evaluation; and using theory in evaluation.

255 Eighteen guides provided an overview of evaluation. This included explanation of the nature
256 of evaluation and how it differs from other types of research; why evaluation is needed and

257 what it can tell us; and the benefits of conducting evaluation. These documents usually
258 targeted practitioners who were new to evaluation.

259 Eight guides informed readers how to identify and assess the quality and relevance of
260 existing research and evaluations. This frequently included links to quality assessment scales
261 such as CONSORT and the Equator network. Four of these eight guides also discuss evidence
262 based practice (EBP); what EBP is, why it is important, and how to conduct evaluations
263 within an evidence based framework. The importance of choosing “best available” methods,
264 even if they are not optimum, are highlighted.

265 Theory is a critical part of intervention development. Theory driven evaluation aims to
266 examine hypothesised causal processes. Nine guides discuss the use of theory in program
267 development and evaluation and two UK Medical Research Council documents (MRC1
268 Framework and MRC2 Process evaluation- see below) were found to be particularly useful in
269 this regard.

270 Challenges inherent in evaluation of public health initiatives are numerous. It is not always
271 possible or practical to conduct a high quality evaluation, and five documents present
272 sections on methodological and practical challenges and ways of overcoming such
273 challenges. Just one document (Evaluability Assessment) focused on evaluability; which is
274 defined as “*the extent to which an activity or project can be evaluated in a reliable and*
275 *credible fashion.*” The authors review the literature on evaluability, and provide an overview
276 of what it is, its purpose, what it includes, and how it can be completed.

277 Four guides (including the UK Government Magenta book and The Green book) focus on the
278 evaluation of policy. These books are intended to be used by policy makers working in or
279 with the UK government in order to support evidence for policy making.

280 2. *Pre-evaluation work*

281 Nine content categories were used to identify information on the steps prior to launching an
282 evaluation (see Table 2). This included: completing a needs assessment; developing a logic
283 model; planning the evaluation; developing a protocol; budgeting; developing contracts and
284 establishing communications; pilot testing; obtaining ethical approval; and involving
285 stakeholders.

286 Thirteen guides provide information and resources to support needs assessments; including
287 how to collect data about a population or community to inform the intervention, how to
288 identify issues and problems, how to assess whether or not issues and problems are shared
289 by the target population, and how to assess populations or communities at the start of an
290 intervention or program. A needs assessment may feed into the development of a logic
291 model. These are diagrammatic representations of the program, describing delivery
292 mechanisms, intervention components, mechanisms of impact, and intended outcomes.
293 Logic models were discussed in twenty guides. A logic model is part of the evaluation plan; a
294 written document specifying the direction the evaluation should take based on priorities,
295 resources, time, and skills needed to complete the evaluation. These guides recommend
296 that all stakeholders should be involved in the development of such a plan to ensure that
297 the process is clear, and to establish consensus on the purpose and procedures of the
298 evaluation. Twenty one guides discuss the process of developing an evaluation plan;
299 including one document by the CDC which focuses exclusively on planning (CDC2 Evaluation
300 Plan). Only three documents specifically discussed the processes involved in the
301 development of an evaluation protocol (Better Evaluation, MRC2 Process Evaluation and
302 APCRC). Of particular note is the UK NHS document (APCRC) in which a protocol template is
303 provided.

304 Once developed, guides recommend that evaluation methods, materials, and procedures
305 should be piloted testing for feasibility. It is important to know, for example, if it is possible
306 to recruit participants, if the data collection tools are suitable, if the outcomes measured are
307 appropriate. Five guides discuss pilot testing. Evaluation can be costly and when designing
308 an evaluation, the questions asked and methods chosen must reflect the funds available.
309 Fifteen guides provide information and support on how to budget for an evaluation. A range
310 of ethical issues need also be considered when planning an evaluation of public health
311 initiatives. For example, in some cases, it may not be ethical to withhold an intervention

312 from a group of people. In such cases, randomised controlled trials would not be suitable.
313 There are also ethical issues surrounding informed consent and data collection, and issues
314 surrounding health inequalities. Ten guides discuss such ethical issues.

315 Seventeen guides discuss the importance of stakeholder involvement, or strategies for
316 involving stakeholders in the evaluation process. Many of these documents also provide
317 advice about facilitating healthy communications between stakeholders, or developing
318 evaluation contracts so that each party has a clearly specified role.

319 *3. Evaluation processes and procedures*

320 Eight content categories relate to the processes and procedures of completing an evaluation
321 (see Table 3). Eighteen guides include an overview of the processes involved in evaluation.
322 This frequently took the form of a checklist of activities involved in evaluation. In some
323 instances, the checklist was structured in terms of essential and desirable features. Other
324 evaluation process content categories include defining an evaluation question; specifying
325 outcomes; describing the intervention; choosing research design and methods; collecting
326 data; managing, analysing and interpreting data; and learning and reporting.

327 A total of nineteen guides provide information on developing a research question. Such
328 documents highlight the need to choose a question that can be answered within the
329 confines of the time, resources, and skill sets that are available; and note the importance of
330 designing a question which is important and useful to all key stakeholders, and can feasibly
331 be answered. Defining the research question is intricately linked to understanding the
332 program, and understanding the outcome. Specifically, how to describe the goal of the
333 program, any activities, and what is and is not part of the program, and how to choose,
334 define, and develop outcomes and outcome measures. Nineteen guides include a section
335 describing how programs should be described, and twenty four guides discuss factors
336 including what makes a good indicator, and suggest a number of considerations when
337 selecting indicators. In some cases outcome indicators or measures are suggested, or
338 objective or validated measures provided.

339 Twenty six guides discuss different research designs and methodological approaches that
340 may be used when evaluating a program. This includes describing qualitative and
341 quantitative methods, as well as discussing particular trial designs (e.g., randomised

342 controlled trials). The majority of these documents discussed data collection approaches
343 (i.e., observations, surveys, focus groups, interviews, existing records etc), and how to
344 manage, analyse and interpret data. They also discuss the importance of methodological
345 rigor, cost effectiveness and validity, reliability and credibility. However, the level of detail
346 provided was limited. Often, guides include a brief overview of types of analyses, or the
347 importance of matching the type of analyses to the study design and research question.
348 Statistical advice was not provided – although links to statistics books were frequently
349 included. Twenty two documents discuss how to learn from and disseminate findings once
350 data has been analysed.

351 *4. Types of evaluation*

352 Seven content categories were used to identify guidance on types of evaluation (see Table
353 4). Five documents provide a brief overview of different types of evaluation (i.e., process
354 evaluation, outcome evaluation, economic evaluation). These documents briefly compare
355 and contrast these forms of evaluation, but do not provide any real detail. Other documents
356 either focus exclusively or in part on specific types of evaluation including; process
357 evaluation; outcome evaluation; economic evaluation; natural experiments; community
358 projects; and fidelity.

359 Fourteen guides discuss process evaluations, and especially, the UK Medical Research
360 Council (UK MRC) document, dedicated entirely to the conduct of process evaluation.
361 Fifteen guides discuss outcome evaluations in terms of what they are and how they differ to
362 other types of evaluations, and eight guides include a discussion of economic evaluation.
363 This includes two documents by Gov.UK and the World Health Organisation (The Green
364 Book and WHO2) which focus exclusively on economic evaluation. Other documents provide
365 a definition and explanation of economic evaluations, and discuss the importance of
366 considering cost effectiveness of programs. A separate UK MRC document provides detailed
367 consideration of evaluating natural experiments, and two documents were developed
368 specifically to aid the evaluation of community projects. These two documents note the
369 challenges associated with community projects, and provide suggestions for overcoming
370 such problems. Eight guides focus on fidelity, specify the importance of assessing fidelity,
371 and provide suggestions regarding how it may be achieved.

372 *5. Additional support*

373 Six content categories were used to identify guidance on supporting the conduct of
374 evaluations (see Table 5). For example, one of the reviewed guides was a website (UKES)
375 offering support and forum for communication for all involved in evaluation. Four guides
376 provide recommendations for achieving high quality evaluations. For example, the National
377 Institute for Health Care and Excellence (NICE) lists a number of recommendations for
378 evaluation. Twenty one guides include tools or tool kits to support the evaluation process.
379 Frequently, tools were included as an appendix, and include tools such as checklists,
380 templates, outcome indicators and surveys. Nine guides provide links to other resources in
381 which further information is detailed. Six documents include information and support to
382 ensure evaluations are of the highest quality. For example, the Centres for Disease Control
383 and Prevention (CDC) framework include a series of standards that evaluations should
384 adhere to. Six documents discussed factors to consider when hiring an external evaluator to
385 complete the evaluation. Finally, eleven guides provide either links to training courses, or
386 online training in evaluation.

387 **Discussion**

388 We have presented findings from a systematic and comprehensive search of documents
389 providing guidance on evaluation of public health initiatives. We identified 402 guides on a
390 range of topics created for a variety of different types of users with different initial
391 expertise. In order to render this literature more accessible we have suggested a series of
392 five criteria that reduced our list to 98 guides. Then by selecting non-overlapping guides
393 from national and international organisations to just 48. We have provided a brief summary
394 of each of these guides and categorised the content of each across 37 categories. We
395 believe that this will make the evaluation guidance literature much more accessible for
396 public health practitioners and commissioners.

397 Despite the abundance of evaluation guidance, many practitioners claim that they do not
398 use guidance documents and do not find them useful (Denford et al., 2016). This may be in
399 part due to an inability to easily understand the purpose, content and target audience of
400 available guides. For example, a simple guide may not be of much use to an experienced
401 evaluator, but could be ideal for a novice. Whilst generic guides may be used to support
402 evaluators in any situation, they may also be lacking the necessary detail to support specific
403 activities such as choosing outcomes. The complexity of public health evaluation makes it

404 impossible to develop a guide that suits all needs. Our project was to begin to map out what
405 guidance is available for whom and to provide a guide to a limited range of easily-accessible,
406 quality-assured guidance to evaluation.

407 ***Strengths and limitations***

408 We have provided researchers and practitioners with a tool to identify and use relevant
409 evaluation guidance documents. We will have undoubtedly missed some guidance in this
410 wide-ranging and desperate literature. However, every attempt was been made to be as
411 inclusive and transparent as possible and we are confident that the sample selected is
412 representative of the literature as a whole. Moreover the list of 48 reviewed and
413 categorised guides are recent, accessible and high-quality. Our categorisation of the content
414 of these guides will allow readers to identify guides that provide relevant, high quality
415 information.

416 To promote accuracy, the content of the guides, and the associated reviews were checked
417 and agreed by two authors and the authors of the original guidance documents whenever
418 possible.

419 Evaluation methods, tools and approaches are continuously developing and progressing.
420 Consequently, our review will need to be updated on a regular basis. This would be
421 relatively easy given the systematic and transparent search strategies.

422 **Conclusion**

423 A wide range of guidance on evaluation of public health initiatives is available. However,
424 time and knowledge constraints may mean that busy practitioners find it challenging to
425 review the range of available guidance and access the most, up-to-date, relevant and useful
426 guidance. This review presents links to and reviews of 48 quality guides to evaluation as well
427 as categorising their content. This facilitates quick and easy access to multiple selected
428 sources of specific guidance.

429 **Declarations**

430 **Ethics approval and consent to participate**

431 Not applicable

432 **Consent for publication**

433 Not Applicable

434 **Availability of data and material**

435 All data is available on request

436 **Competing interests**

437 The authors declare that they have no competing interests

438 **Authors' contributions**

439 The initial idea and project design was developed by CA, SD, FD, and SA. SD, MC, and PA
440 carried out the initial search. Data extraction was completed by SD, PA, SA, FD, and MC.
441 Data analysis was completed by SD and MC. The paper was developed by SD with
442 considerable input from CA, and MC. All authors read and approved the final draft.

443 **Acknowledgements**

444 This study was funded as part of National Institute of Health Research's School for Public Health
445 Research (NIHR SPHR) project with additional support from the NIHR Collaboration for Leadership in
446 Applied Health Research and Care of the South West Peninsula (PenCLAHRC). The views expressed
447 are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of
448 Health. Author order represents extent of contribution. Funding statement
449 The School for Public Health Research and the Collaboration for Leadership in Applied Health
450 Research and Care of the South West Peninsula is funded by the National Institute for Health
451 Research (NIHR). SPHR is a partnership between the Universities of Sheffield, Bristol, Cambridge,
452 Exeter, UCL; The London School for Hygiene and Tropical Medicine; the LiLaC collaboration between
453 the Universities of Liverpool and Lancaster and Fuse; The Centre for Translational Research in Public
454 Health, a collaboration between Newcastle, Durham, Northumbria, Sunderland and Teesside
455 Universities.

456 **Abbreviations**

457 **CDC Centres for Disease Control and Prevention**

458 **EBM Evidence Based Medicine**

459 **MRC Medical Research Council**

460 **NICE National Institute for Health and Care Excellence**

461 **WHO World Health Organisation**

462 Reference List

463

464 Better Evaluation. Better evaluation. (n.d.). *Better Evaluation; sharing information to improve*
465 *evaluation*. Retrieved 2nd November, 2015, from <http://betterevaluation.org/>

466 Centres for Disease Control and Prevention (CDC) (1999). Framework for program evaluation in
467 public health. *Morbidity and Mortality Weekly Report*, 48, RR-11.

468 Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2013). Developing and
469 evaluating complex interventions: the new Medical Research Council guidance.
470 *Int.J.Nurs.Stud.*, 50, 587-592.

471 Ogilvie, D., Cummins, S., Petticrew, M., White, M., Jones, A., & Wheeler, K. (2011). Assessing the
472 evaluability of complex public health interventions: five questions for researchers, funders,
473 and policymakers. *Milbank Q.*, 89, 206-225.

474 Denford, S., Callaghan, M., Abraham, C., Lakshman, R (in preparation) A qualitative study of public
475 health practitioners' views of evaluation guidance. *American Journal of Public Health*.

476 House of Lords Science and Technology Select Committee. Behaviour change. 2nd Report of Session
477 2010-12. Stationery Office, 2011.

478

479

480

481