

AUTHOR GUIDANCE

EPOC Qualitative Evidence Syntheses guidance on when to sample and how to develop a purposive sampling frame

Aim of this guidance: To introduce review authors to one possible method for purposively sampling included studies for a qualitative evidence synthesis.

For some qualitative evidence synthesis questions, there are a large number of primary qualitative studies available. However, in contrast to reviews of effectiveness, the inclusion of a large number of primary studies with a high volume of data in a qualitative synthesis can threaten the quality of the synthesis. There are a number of reasons for this: firstly, qualitative processes of analysis require detailed engagement with text and large volumes of data make this difficult to achieve. Secondly, qualitative evidence syntheses aim for greater variation in concepts, to help ensure conceptual generalizability, whereas effectiveness reviews aim to be exhaustive in order to achieve statistical generalizability. Sampling may help to achieve variation while also ensuring that the analysis is not overwhelmed by a very large volume of primary data.

Purposively sampling from the primary qualitative studies identified as eligible for inclusion in a QES is one way to reduce the amount of data contributing to the analysis. The objective of this guidance is to provide practical guidance for EPOC QES authors on how to approach the issue of sampling for qualitative evidence syntheses.

When can sampling be considered for studies included in a qualitative evidence synthesis?

In a QES, the threshold at which the number of primary qualitative studies contributing data to the QES becomes too large can be affected by:

- the amount of relevant data in the included studies (data richness)
- the study design (for example mixed methods studies or surveys with open ended questions typically provide less data than in-depth qualitative studies)
- how closely and completely the objective of the included primary studies matches the review objective. Where many of the included primary studies only address part of the review objective, more studies may need to be included to obtain data that address the full scope of the review objective.

What constitutes a sufficient number of primary studies, and a sufficient amount of data, for analysis will vary across reviews. A judgement therefore needs to be made by each review team on whether there is sufficient data, but not so much data that the analysis process will be difficult.

The following steps may be useful in making a judgement on whether sampling should be used:

1. Identify from the search outputs all of the primary qualitative studies that meet the inclusion criteria for the review
2. Familiarize yourself with the relevant data in the included studies, including how rich these data are

Suggested citation: Cochrane Effective Practice and Organisation of Care (EPOC). [Resource title]. EPOC Resources for review authors, 2017. epoc.cochrane.org/resources/epoc-resources-review-authors (accessed DD Month YYYY)

3. Carry out a simple mapping of the included studies based on the key elements of your review question and important contextual considerations for the synthesis. This information can be summarized in a table so that the key descriptive information for all studies can be easily viewed. The mapping could include:
 - a. Geographic setting/s of the studies
 - b. Population and or participants
 - c. Health issue/s addressed by the studies
 - d. Intervention/s addressed by the studies, if applicable
 - e. Study type/design (for example hypothetical study, pilot study, evaluation)
 - f. Health care settings of the studies
 - g. Data collection methods
 - h. Temporal characteristics (how old is the data?) (This may be important if there have been changes to laws in the area of study or introduction of new technologies)
 - i. Any other considerations that may be relevant to the synthesis question, such as policy or political issues in the study settings, social climate (for example if a practice is socially acceptable such as abortion), legislative issues (such as whether a particular practice is legal)
 - j. Data richness (See worked examples)
4. Decide within the review team if using all of the included studies in the analysis would lead to more data that can be reasonably managed in the analysis
5. If this seems likely, then consider your sampling options as described below. Please note that the decision to sample can be revisited later in the synthesis process, if the earlier judgement made regarding the number of studies and amount of data is no longer viewed as appropriate.

How to purposively sample articles for a qualitative evidence synthesis using a sampling frame

There are a variety of ways authors can sample from primary studies for qualitative evidence synthesis. (See appendix). In this author guidance, we present a way of sampling that builds on a few of the examples in the table. We have chosen this example as it is relatively straight forward to apply, mirrors what would be done in primary qualitative research to sample participants and has been used in several Cochrane qualitative evidence syntheses to date (Odendaal 2015, Ames, Glenton et al. 2017, Ames, Glenton et al. 2019).

Now that you have identified that you will purposively sample from the included studies, we will present a step-by-step guide to **one way** of sampling. This has the aim of achieving the broadest possible variation within the included studies while still providing rich and relevant data for your synthesis. We recommend following these steps:

1. Review the map of included studies that you created above. Which of these elements are key to answering your synthesis question? For example:
 - a. Which geographic and health care settings need to be included in the synthesis? Are there specific settings that need to be represented, such as low and middle-income countries or tertiary hospitals? If so, how well represented are these settings in the included studies?
 - b. Is there a certain population or group of participants that needs to be represented in your synthesis?
 - c. Are there multiple health issues or interventions that you need to address or include such as different vaccines or family planning methods?

Suggested citation: Cochrane Effective Practice and Organisation of Care (EPOC). [Resource title]. EPOC Resources for review authors, 2017. epoc.cochrane.org/resources/epoc-resources-review-authors (accessed DD Month YYYY)

2. Decide on which are the key elements that will enable you to capture sufficient rich data to answer your review objectives. These elements then become the base for your sampling framework
3. Decide in which order you will apply the sampling framework to the included primary qualitative studies.
4. Pilot the sampling framework on 10 studies to see if this results in the inclusion of the most relevant studies with rich data that answer your review objectives
5. Apply the sampling frame to all of the included studies

Please see below for two worked examples of how this type of purposive sampling was applied in two different qualitative evidence syntheses.

Suggested citation: Cochrane Effective Practice and Organisation of Care (EPOC). [Resource title]. EPOC Resources for review authors, 2017. epoc.cochrane.org/resources/epoc-resources-review-authors (accessed DD Month YYYY)

Writing up your sampling strategy

It is important that you provide a clear and transparent description of your (planned) sampling process in both the protocol and the qualitative evidence synthesis. (You can find more information in the EPOC QES template). As authors, you will need to reflect on any possible limitations of your sampling and describe these in the discussion section of your synthesis.

You should include both the 'sampled' and 'included but not sampled' studies in your characteristics of included studies table. In this table, clearly indicate which studies were sampled and which were not. Also, indicate the reason for not sampling for each study in this group. For studies that were not sampled, the amount of descriptive information included in the table can be less – as a review author team, you should decide what information is important to include. For further guidance, please refer to the examples below and the EPOC QES template.

Worked example 1:

This example will describe the sampling procedure from Parents' and informal caregivers' views and experiences of communication about routine childhood vaccination: a synthesis of qualitative evidence (Ames, Glenton et al. 2017).

The specific objectives of the synthesis were to identify, appraise and synthesize qualitative studies exploring: parents' and informal caregivers' views and experiences regarding communication about childhood vaccinations and the manner in which it is communicated; and the influence that vaccination communication has on parents' and informal caregivers' decisions regarding childhood vaccination.

79 studies met the inclusion criteria and 38 were sampled for inclusion in the data synthesis.

In order to decrease the number of included studies to a manageable amount for the synthesis, the authors chose the following three step sampling frame (Ames, Glenton et al. 2017):

First, we wanted to ensure a geographic spread and reasonable representation of findings from LMICs, given that the synthesis intended to cover all geographic settings. We therefore sampled in all studies from low- and middle-income country (LMIC) settings, as most studies took place in high-income country (HIC) settings. Second, we created a simple 1 to 5 scale for assessing the richness of data (See table 2). To our knowledge, there is not existing system for assessing data richness so we created one to fit our needs. We sampled in all articles that scored a 4 or higher for data richness. We decided to focus on the richness of the data within the included studies to ensure that we would have enough data to work with for the synthesis. We based this decision on the rationale that rich data can provide clearer insights into the phenomenon of interest.

Table 2: Data richness scale

Score	Measure	Example
1	Very little qualitative data presented that relate to the synthesis objective. Those findings that are presented are fairly descriptive.	For example, a mixed methods study using open ended survey questions or a more detailed qualitative study where only part of the data relates to the synthesis objective
2	Some qualitative data presented that relate to the synthesis objective	For example, a limited number of qualitative findings from a mixed methods or qualitative study

Suggested citation: Cochrane Effective Practice and Organisation of Care (EPOC). [Resource title]. EPOC Resources for review authors, 2017. epoc.cochrane.org/resources/epoc-resources-review-authors (accessed DD Month YYYY)

3	A reasonable amount of qualitative data that relate to the synthesis objective	For example, a typical qualitative research article in a health services journal
4	A good amount and depth of qualitative data that relate to the synthesis objective	For example, a qualitative research article in a social sciences journal with more context and setting descriptions
5	A large amount and depth of qualitative data that relate in depth to the synthesis objective.	For example, from a detailed ethnography or a published qualitative article with the same objectives as the synthesis

Finally, we examined the remaining studies after applying the first two elements and sampled in any studies that closely matched the synthesis question. This was to ensure that the data of highest relevance to the review was included, even if these data were thin and from a setting already represented in the synthesis.

Worked example 2:

Patients' and the public's perceptions and experiences of targeted digital communication via mobile devices for reproductive, maternal, newborn, child and adolescent health: A qualitative evidence synthesis (Ames, Glenton et al. 2019).

The specific objective of the synthesis was to explore patients' and the public's perceptions and experiences of targeted digital communication via mobile device in the areas of reproductive, maternal, new-born, child or adolescent health.

48 studies met the inclusion criteria and 35 were sampled for inclusion in the data synthesis. We divided the studies by type of participant;

- Adolescent and youth populations as potential users of SRH services
- Adult populations as potential users of SRH services
- Pregnant and postpartum women (up to 6 weeks)
- Pregnant and postpartum women (up to 6 weeks) living with HIV
- Parents and other caregivers of children under five years of age

In order to decrease the number of included studies to a manageable amount for the synthesis, the authors chose the following three step sampling frame and applied it to each of the participant groups:

1. We sampled in studies conducted in low and middle-income countries as there were fewer of these studies and the focus of the review was global.
2. We applied the data richness scale and sampled in studies rating 3 or higher. This data richness scale was adapted from the work above. It rates the amount of rich data relevant to the review objective specifically. (See table 3)
3. We looked at the topic of the digital targeted communication interventions represented within each participant group and made sure that there was a broad variation (for example family planning, medication reminders, sexual health promotion).

Suggested citation: Cochrane Effective Practice and Organisation of Care (EPOC). [Resource title]. EPOC Resources for review authors, 2017. epoc.cochrane.org/resources/epoc-resources-review-authors (accessed DD Month YYYY)

Table 3: Adapted data richness scale adjusted to rate richness of the data related to the review objective

Score	Measure	Example
1	Very little qualitative data presented that relate to the synthesis objective. Those data that are presented are fairly descriptive.	For example, a mixed methods study using open ended survey questions or a more detailed qualitative study where only part of the data relates to the synthesis objective
2	Some qualitative data presented that relate to the synthesis objective	For example, a limited number of qualitative findings from a mixed methods or qualitative study
3	A reasonable amount of qualitative data that relate to the synthesis objective	For example, a typical qualitative research article in a journal with a smaller word limit and often using simple thematic analysis
4	A good amount and depth of qualitative data that relate to the synthesis objective	For example, a qualitative research article in a journal with a larger word count that includes more context and setting descriptions and a more in-depth presentation of the findings
5	A large amount and depth of qualitative data that relate in depth to the synthesis objective.	For example, from a detailed ethnography or a published qualitative article with the same objectives as the synthesis

Links to qualitative evidence syntheses that have used purposive sampling

- Ames 2017 (Worked example 1): <https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD011787.pub2/abstract>
- Ames 2019 (Worked example 2): Submitted for final editorial approval

Links to papers discussing how to sample for qualitative evidence synthesis

- Ames 2019 (<https://bmcmmedresmethodol.biomedcentral.com/articles/10.1186/s12874-019-0665-4>)
- Benoot 2016 (<https://www.ncbi.nlm.nih.gov/pubmed/26891718> (see full review at <https://lirias.kuleuven.be/bitstream/123456789/569295/1/benoot+sexual+adjustment+synthesis.pdf>))
- Suri 2011 <http://www.emeraldinsight.com/doi/abs/10.3316/QRJ1102063>
- Cochrane Qualitative and Implementation Methods Group guidance paper 3: <http://discovery.ucl.ac.uk/10047708/1/CQIMG%20Paper%203.pdf>

References

Ames, H. M., et al. (2017). "Parents' and informal caregivers' views and experiences of communication about routine childhood vaccination: a synthesis of qualitative evidence." Cochrane Database of Systematic Reviews. 2: CD011787.

Ames, H. M., et al. (2019). "Clients' perceptions and experiences of targeted digital communication accessible via mobile devices for reproductive, maternal, newborn, child and adolescent health: A qualitative evidence synthesis." Cochrane Database of Systematic Reviews. Forthcoming.
Suggested citation: Cochrane Effective Practice and Organisation of Care (EPOC). [Resource title]. EPOC Resources for review authors, 2017. epoc.cochrane.org/resources/epoc-resources-review-authors (accessed DD Month YYYY)

Odendaal, W. G., Jane; Griffiths, Frances; Tomlinson, Mark; Leon, Natalie; Daniels, Karen (2015). Healthcare workers' perceptions and experiences on using mHealth technologies to deliver primary healthcare services: a qualitative evidence synthesis (Protocol). Cochrane Database of Systematic Reviews. **11**:Art. No.: CD011942.

Suri, H. (2011). "Purposeful sampling in qualitative research synthesis." Qualitative Research Journal **11**(2): 63-75.

Suggested citation: Cochrane Effective Practice and Organisation of Care (EPOC). [Resource title]. EPOC Resources for review authors, 2017. epoc.cochrane.org/resources/epoc-resources-review-authors (accessed DD Month YYYY)

2019 10 03

Appendix 1: Some examples of purposeful sampling methods (Suri 2011)

Type of sampling	Description
Extreme or deviant case sampling	<ul style="list-style-type: none"> • Selecting illuminative cases that exemplify 'extreme' or 'deviant' contexts or examples, for instance: <ul style="list-style-type: none"> – where an innovation in a primary study was perceived notably as a success or failure – where findings of a primary study are very different from those of most studies identified for the synthesis
Maximum variation sampling	<ul style="list-style-type: none"> • Constructed by: <ul style="list-style-type: none"> – identifying key dimensions of variation, and then – finding cases that vary from each other as much as possible along these dimensions • This sampling yields: <ul style="list-style-type: none"> – 'high-quality, detailed descriptions of each case, which are useful for documenting uniqueness, and – important shared patterns that cut across cases and derive their significance from having emerged out of heterogeneity' (Patton, 2002, p. 235)
Snowball or chain sampling	<ul style="list-style-type: none"> • Trying to locate a key work in the field through talking with experts or locating a key article that is often cited • Then follow on with primary studies that have cited the key or landmark study
Theoretical or operational construct sampling	<ul style="list-style-type: none"> • Selecting cases that represent important theoretical or operational constructs about the phenomenon of interest • Set out operational definitions of key theories or constructs related to the phenomenon of interest • Develop boundaries for these by creating specific inclusion and exclusion criteria in relation to selecting primary studies for the synthesis
Criterion sampling	<ul style="list-style-type: none"> • Used by those trying to construct a comprehensive understanding • Studies are sampled based on a predetermined criteria • Specific inclusion and exclusion criteria are clearly stated • Studies are then analysed as a whole
Stratified purposeful sampling	<ul style="list-style-type: none"> • Following on from criterion sampling where each of the criteria would become a sample • Stratified samples are samples within samples where each stratum, or group, is fairly homogenous and are analysed within these groups • Useful for examining variation in a key phenomena of interest
Purposeful random sampling	<ul style="list-style-type: none"> • Randomly select from the list of included studies for inclusion in the analysis • For example, use a random internet based selector, choose every 3rd included study or pull study names from a hat • Provides an unbiased way of selecting studies for inclusion but may not provide studies with rich data
Combination or mixed purposeful sampling	<ul style="list-style-type: none"> • Choosing a combination or mix of sampling strategies to best fit your purpose <ul style="list-style-type: none"> – For some syntheses, it may be useful to use a combination or mix of sampling strategies. For instance, by applying theoretical sampling in a first stage and deviant case sampling in a second stage. This should be guided by the review methods and purpose, and the time available

Suggested citation: Cochrane Effective Practice and Organisation of Care (EPOC). [Resource title]. EPOC Resources for review authors, 2017. epoc.cochrane.org/resources/epoc-resources-review-authors (accessed DD Month YYYY)