Part 2: Case study and data source selection

Bath SDR

QuIP

Part 2. Case and data source selection

The purpose of this section is to consider:

- What criteria to use to select a case study project or activity for a QuIP study from a ‘portfolio’ of wider activities
- What kind of data to collect
- How to select or to sample sources of data from a larger number of possible sources, and how many sources to use.

Once you have finished Parts 1-3, you may wish to work on a draft case selection strategy for your project using the resources available in this module. Although we won't be using this in the live session, it may prompt some questions for the session and will help when it comes to writing up your assignment.

The journal article referenced in the audio clip (**Case and evidence selection for robust generalisation in impact evaluation**) can be accessed here.

In addition, for a short and simple introduction to thinking differently about sampling when doing qualitative work, see this great article by Kizzy Gandy: Gandy, K. (2024). How many interviews or focus groups are enough? Evaluation Journal of Australasia, 24(3), 211-223. https://doi.org/10.1177/1035719X241266964.

If you don't feel very confident about a non-probability approach to sampling, reading this will help you to prepare for the exercises ahead in this section, and in the live session on your first day.

This quote from the article is an important principle to bear in mind when thinking about sampling for your evaluation:

"It is implicitly assumed in thematic saturation studies that saturation is always important. This is true for generating generalisable knowledge, but evaluators tend to prioritise program-specific performance insights. Generalisable knowledge is associated with a positivist paradigm whereas evaluators typically move between positivist, postmodern, and constructivist paradigms across different key evaluation questions and draw on multiple sources of data. Therefore, qualitative data is typically used by evaluators to develop a depth of understanding rather than breadth, and sometimes qualitative sample sizes as low as one can be justified."[4]

Two nested choices

Start with an organisation that manages a portfolio of activities over time (e.g. projects, contracts, branches).

Choice 1. Case study selection and scoping	Choose one (or more) of these for a QuIP study. Define its boundaries, and identify the populations of relevant stakeholders to consult (secondary or written sources may also be considered for coding and analysis)
Choice 2. Data source selection	Select a feasible number of respondents within each population. This may two or more stages of stratified sampling.

Four core principles and questions

1. **Saturation**

How to maximise the potential to learn about the fullest possible range of drivers of change affecting the portfolio, case study and selected populations? (Exploratory studies)

2. Bayesian updating

Have to maximise the potential to test prior thinking about what matters most? (Confirmatory studies)

3. Equating marginal benefits and costs

Up to what point does the extra evidence obtained from adding a case or data source exceed the cost of obtaining it?

4. Heterogeneity

What are the most important sources of variation in impact we want to understand (e.g. over time, by location, by gender-age, arising from exposure to a variety of actions).

Criteria for selection

The key criterion is to support robust and credible generalisation.

The principles for doing this are NOT the same as those for estimating the average value of a known variable across a population (e.g. average crop yields across a district).

Random selection is NOT necessarily best, indeed is unlikely to be.

Doing all data at the same time is NOT necessarily optimal.

There is no scientifically optimal sample size, but transparent and reasoned case selection is important to rebuff criticisms of ‘cherry-picking’.

Data source selection: an illustrative strategy

The original QuIP standard sample comprised 24 face-to-face interviews and 4 focus group discussions across a selected rural population. This is roughly what two people can cover in a single field trip of 5-7 days.

Step 1. Sample purposively **between clusters**

Step 2. Sample **within clusters** through a mix of further segmentation and random quota sampling.

Step 3. Review if this is sufficient (against saturation and Bayesian updating criteria). If not then consider increasing the sample size (and budget) to collect more diverse experiences.

Example 1. Two stage selection using contextual data: (a) cross-cluster selection

Project X has two packages (X1,X2) aiming to raise crop incomes (Y) in five districts with different cropping potential (Z).

The selection task is to choose two districts from the five. 60 different intended beneficiaries receive each package in each district.

District	Index of potential (Z)
A	5
B	10
C	15
D	20
E	25

The project's theory of change indicates that Y should be easier to achieve in districts with higher potential (measured by Z).

Which two districts would you choose? Why?

Possible answers

First option - choose A and E.

- This ensures data is collected across the full range project potential, from best to worst. Causal drivers in the intermediate range of districts (B to D) may differ in degree more than form and can be interpolated.

or

Second option - choose B and D.

- This ensures data is collected across the range of contexts but avoiding extremes (A and E) which may be exceptional.

There is no correct answer, but these two answers seem most balanced on the basis of the data supplied.

Example 1: (b) within-cluster selection

Assume B & D were selected, and we now have to choose 12 respondents in district D.

This is how intended beneficiaries are distributed within it, by area and package.

How would you select two quota samples of 6, one for each package?

Area	Z	X1	X2
1	16	30	
2	18	10	10
3	20	10	10
4	22	10	10
5	24	-	30

[A quota sample comprises a target number of respondents of a given type drawn opportunistically].

Possible answers

First option

- Choose one at random for every ten respondents in each village receiving each package (three in Areas 1 and 5, two in Areas 2, 3 and 4). This provides a more representative distribution of respondents, but entails having to visit all five areas.

or

Second option

- Choose three in Areas 1 and 5, and six in Area 3 (three receiving each package). This reduces logistical costs, and permits deeper analysis of who got which package in one place (Area 3) which also has potential typical of the district.

Neither answer is necessarily better, given the information supplied

Bath SDR

QuIP

Example 2. With outcome data

We now also know how average crop income in each district changed (from comparisons between a baseline and endline survey). To simplify, assume there is now only one standard package of activities X.

District	Index of potential (Z)	Mean change in income (ΔY%)
A	16	+10
B	18	+5
C	20	0
D	22	-5
E	24	-10

If you can interview 36 interviews across three districts then which would you select? Why?

Does your choice change if the study is exploratory (no theory of change) or confirmatory (testing a theory of change)?

Possible answers

First option - choose A, C and E.

This ensures data is collected across the full range of contexts (project potential) from best to worst, as well as outcomes (also referred to as positive and negative deviance). Adding C helps to confirm that intermediate districts fall between the extremes (minimising how far we have to interpolate).

Or

Second option - choose A, E and D.

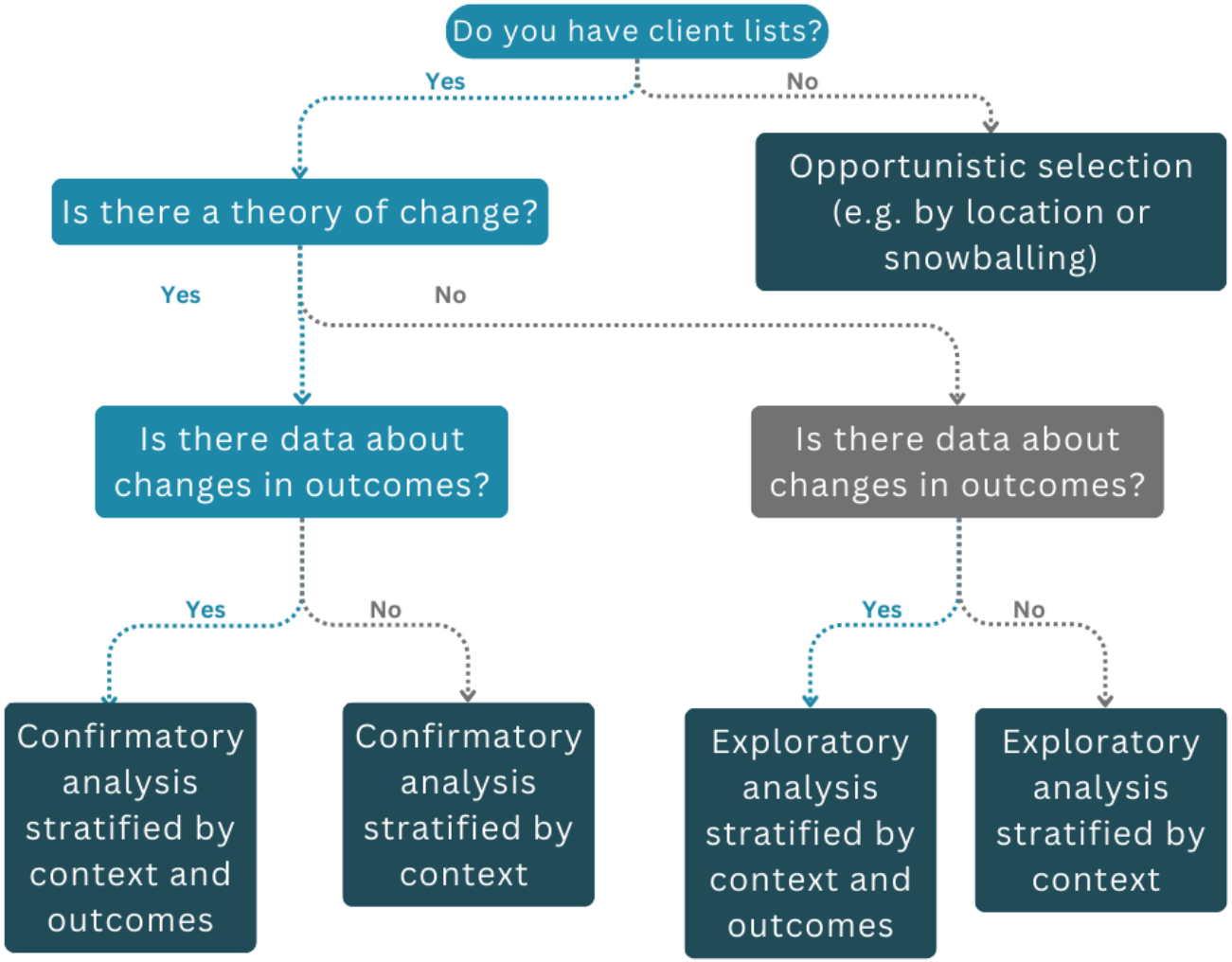
D is anomalous because the outcome is worse than would be expected on the basis of a theory of change that yield changes should are positively correlated with the index of potential (Z).

There is no correct answer, but the theory of change strengthens the case for the second option.

Bath SDR

QuIP

Source data selection - a summary



(For simplicity, this diagram assumes a uniform intervention, X)

Bath SDR

QuIP

Cluster source data selection scenarios (for a uniform X)

Context	Suggestions
No outcome data or theory of change	Maximise dispersion of cases across context, or select randomly
Theory of change linking context and outcomes, but no outcome data	Choose *expected* negative and positive cases
Theory of change and outcome data	Select anomalous cases or outliers relative to theory

Bath SDR

QuIP

Source selection options depending on availability of data across a treated population

Option	Treatment data (X)?	Outcome data (Y)?	Contextual data (Z)?	Comment
A	No	No	No	Random selection across full population is the only option
B	Yes	No	No	Select randomly from quota samples across categories of treatment or exposure
C	No	Yes	No	Select purposively to include positive and negative deviants
D1	No	No	Yes	Select purposively to reflect important dimensions of variation across the population (e.g. gender, age)
D2	No	No	Yes	Select purposively to include likely positive and negative deviants according to prior theory.
E	No	Yes	Yes	Select purposively to include anomalous cases poorly explained by prior theory linking Z and Y.

Practical constraints to source data selection

- Weak or missing monitoring data (about X, Y and Z)
- Lack of clarity about relevant theory
- Geographical dispersion of the population
- Arbitrary budget constraints

Selection criteria

To serve exploratory goals: aspire to 'saturation' by maximising case variation by socio-economic characteristics and context

To serve confirmatory goals: aspire to improve confidence in theory by testing against anomalous and/or contrasting cases

Focus group discussions (FGDs)

Why do them too? Changing the social context of questioning (e.g. from other household members who may not be given the opportunity to respond to an individual interview, or bringing people of the same gender or age group together as peers) may change what people feel they can say. It can also be useful in contexts where there is a 'communal' aspect to the programme, e.g. community groups. If you want to find out about community/group dynamics or see how people respond in a group situation, focus groups are a good alternative source of information.

Who should attend? This depends on what sources of variation in respondents' perceptions matter most to you. The practical feasibility of getting people together also matters. Inviting other members of interview respondents' households is a good option, as well as encouraging them to bring along a friend.

Should the questions be changed? Yes, but the more radical the changes, the less scope for joint coding and analysis of the responses.

FGD selection: an example

Continue with Example 1 (district D), and assume it is only possible to invite FGD participants from the same area)

Area	Z	X1	X2
1	16	30	
2	18	10	10
3	20	10	10
4	22	10	10
5	24	-	30

If you could only hold one focus group then were would it be?

What you could hold two focus groups?

What about three, four and six?

Possible answers

1. Area 3 would reflect the most typical context and allow participants to attend who had received Package X1 and X2.
2. Areas 1 and 5 would allow FGDs to focus on areas most intensively targeted for Packages X1 and X2, hence maximising diversity.
3. Combine the above (Areas 1, 3 and 5).
4. Conduct two each in Areas 2 and 4, allowing scope for stratification by another criterion (e.g. by Package or by gender). However the pool of participants is quite small.
5. Allows scope for stratification by another criterion in Areas 1, 3 and 5 (e.g. gender or age).

Case study selection and scoping

What criteria inform the choice of one (or more) activities for a QuIP study across the portfolio of an organisation's activities?

It is formally not so different from cluster/source selection. Key questions to consider:

1. What justifies choosing this particular activity for a study?
2. Which are the key stakeholder populations who need to be consulted?
3. What existing monitoring information do you have about each to inform source data selection?
4. How can the population be usefully stratified?
5. What appetite is there for a study that challenges received wisdom within the organisation?
6. Is the budget big enough to sustain a sufficiently robust and credible study to do this?

Questions for further reflection

(thinking about your own potential QuIP study)

1. *What justifies choosing this particular activity for a study?*
2. *Which are the key stakeholder populations who need to be consulted?*
3. *What existing monitoring information do you have about each to inform source data selection?*
4. *How can the population be usefully stratified?*
5. *What appetite is there for a study that challenges received wisdom within the organization?*
6. *Is the budget big enough to sustain a sufficiently robust and credible study to do this?*

If you are planning to submit an assignment then you may want to come back to the additional resources below to help you draft a case selection strategy for your project:

- Sample selection for QUIP studies - briefing note
- How many interviews are enough? Bath SDR blog post with lots of useful links to references!
- Case and evidence selection for robust generalisation
- How many interviews of focus groups are enough? Kizzy Gandy, 2024
- Example sampling strategy notes

Bath SDR

QuIP