Optum

Evaluation toolkit

October 2023

Get started

Welcome

This toolkit has been created to help you to plan and design evaluation of a service or intervention. It was not designed to support activities like identification of improvement opportunities or service/intervention planning, design and implementation. The toolkit might be particularly useful for people involved in service evaluation including community and voluntary sector staff, frontline health and care professionals, analysts, commissioners and managers.

The toolkit is designed to be interactive and is made up of sections that describe aspects of evaluation and questions you might need to think about to complete your evaluation.

You can also use the tabs at the top of each page to navigate to the section you want.

The next page shows how these sections should work in order when you are planning and implementing an evaluation.

Guide Use the guide button to take you to additional information about a subject


Use the back arrow to return to where you were before


Use the forward arrow to navigate forward to the next page


Use the home button to return to this page

Glossary Terms are explained in the Glossary and there is a Resources tab above for further information and useful tools.

Why evaluate

Helping you to think about why evaluation is important and what questions you are trying to answer when you evaluate

Measures

More details on outcomes, measures and resources to help you collect information that will help you answer your evaluation questions

Things to think about

Questions to ask yourself before you begin to design an evaluation and resources to help you think through how to answer each

Design

How to go about designing the best possible evaluation to answer your questions with resources to help you in the design and planning stage

Analysis

Considerations and resources to support your thinking about analysis

Conclusions

Drawing conclusions from your evaluation

Resources

Further resources to support your evaluation

Navigating the toolkit

Before you start (ideally)

Start with Why

Why are you evaluating?
What do you hope to show?

Measures

What data do you have
and what do you need?

Things to think about

What other considerations go into
evaluation design?

Design

Best possible design of
your evaluation

When you have some data

Analysis

Bringing together your data

Drawing conclusions

So what?

Start with Why

Why are you evaluating?
What do you hope to show?

Why is evaluation important for your programme?

Start with Why: Consider these three things to be clear about why evaluation is important to your programme

Why evaluate? What are the questions I'm trying to answer through evaluation?

It is important to demonstrate that your service or intervention is having an impact. You may be looking to scale a pilot and need evidence that this would be worthwhile and to seek funding. Or perhaps you need to demonstrate to a commissioner that your service can continue to run.

Some of the evaluation questions you may be looking to answer could be:

- What does my programme achieve compared to "do nothing" or standard care?
- Who are the people that will most benefit from my programme or service?
- How could evaluation of my service contribute to the existing evidence base in this area?

What outcomes am I trying to achieve through my programme?

Have you created a logic model for your service which will support your evaluation? Guide

What is the value of my programme or service?

What are the risks of my programme or service?

What does my programme look to achieve in terms of improving outcomes for people experiencing health inequalities?

What are the priorities for my programme and are they agreed?

Have you used your logic model to define outcomes and outputs and determined metrics for each?

Can you rank these according to importance to your stakeholders and measurability? Guide

Do your stakeholders agree with your ranking/prioritisation?

Have you set expectations amongst stakeholders for when you will realise the intended outcomes?

When to start your evaluation

It is best practice and will make your life much easier to plan and build the foundation of your evaluation **before** you start your intervention.

Your plan should cover activities that help you to:

- Ensure key stakeholders are on side, aware of the process and realistic expectations are set
- Ensure data collection mechanisms are in place to allow you to record activity and agreed outcomes as they happen.

This is important to make sure that:

- Data is recorded regarding when interventions happen
- Outcomes data is trackable over time
- Improvements in health can be linked or attributed to the intervention
- As much information as possible is available about the intervention you have designed that will inform decisions about its future.

> “A well-designed intervention will include provision for evaluation from the outset. Building in data collection, feedback and measurement of outcomes and impact from the start can ensure that the evaluation is integrated into the design and execution of the intervention. This positions the evaluation as an integral part of the process and delivery rather than an interference competing for tight resources.”
>
> — Health Foundation
> health.org.uk/sites/default/files/EvaluationWhatToConsider.pdf

Logic model framework for evaluation

A logic model helps you start to think through the logical theory of change i.e., which outputs lead to which outcomes?

Further information on logic models can be found here: Using Logic Models in Evaluation-Jul16.pdf (strategyunitwm.nhs.uk)

Directions: Start with the problem statement (1), then agree the goal/aim as the long-term outcomes (2), work backwards to populate the model (3), work forwards using an 'if-then' logic to test the model (4). Gap-analysis. Iterative development. Review and update as necessary.

Considerations: Common understanding and agreement on (1&2), define outcome timescales, multiple stakeholders perspectives on (1-4), SMART, success/failure and intended/unintended consequences.

Rationale for intervention (1)

This is the justification for the programme/intervention. For example, what is the nature and scale of the problem being addressed? What will happen if nothing is done?

Inputs	Outputs			Outcomes – Impact		
Inputs	**Participants**	**Activities**	**Outputs**	**Short-term**	**Medium-term**	**long-term (2)**
What we need/invest in?	Who we reach/involve?	What we do?	What we create?	Results — In terms of learning	Results — In terms of changing action	Results — In terms of change to the conditions
The resources available or required to deliver the work.	The stakeholders involved in the programme /intervention and those with an interest in the long-term outcomes.	The activities undertaken as part of the programme /intervention.	Maybe referred to as process measures. Directly related to/ resulting from activity. Delivering to the intended audience at the intended 'dose' of action.	Changes observed in awareness, attitudes, knowledge, skills, functioning, outlook, etc.	Changes observed in behaviour, practice, condition, and decision-making, etc. These follow on from changes observed in the short-term.	The final desired impact of the programme/intervention. At a population or system-level impacts. Relates closely to the original rationale. Relative weights on multi-faceted aims.

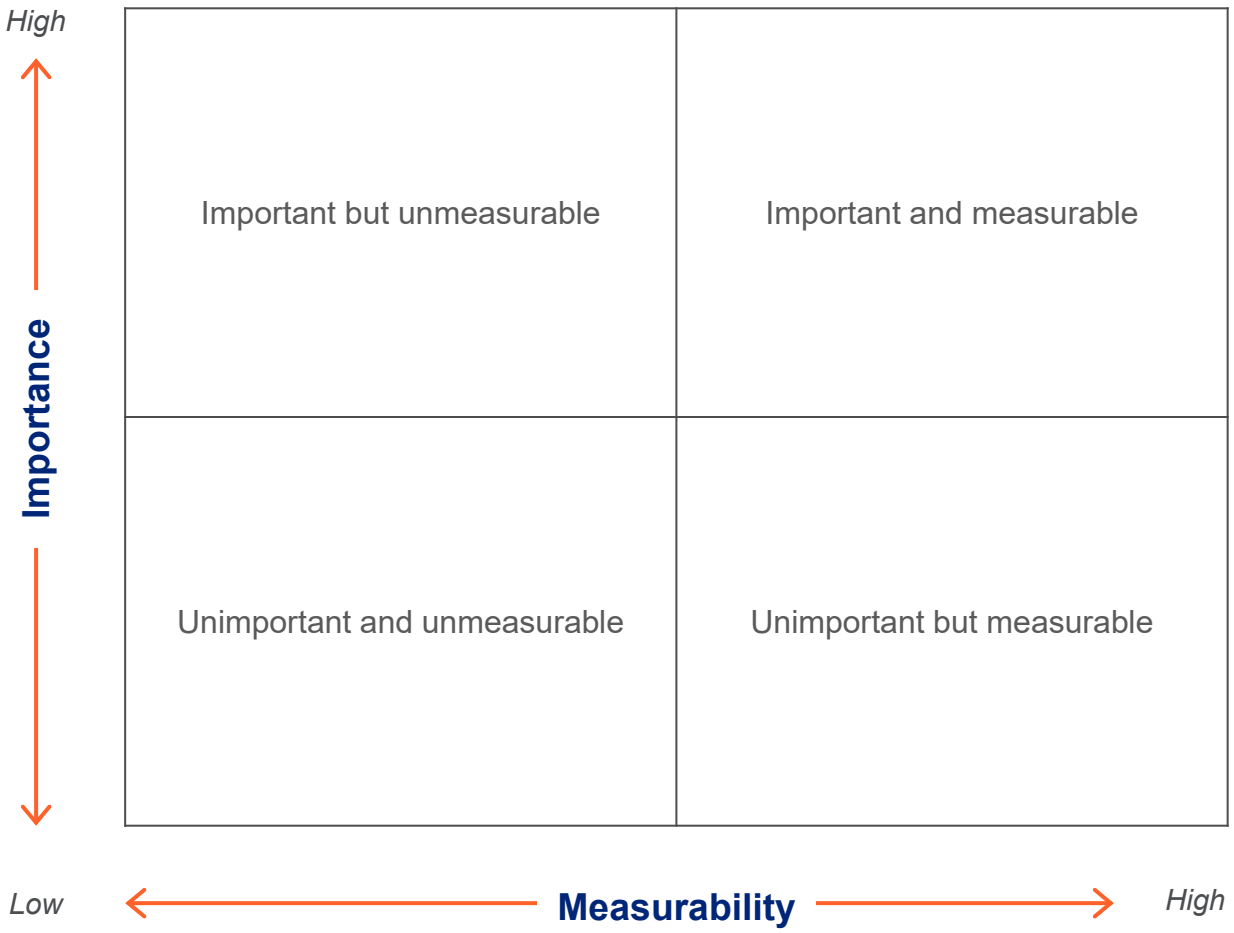
Populate the model (3)

'if-then' logic test (4)

Assumptions

External Factors

Ranking outcomes according to importance and measurability

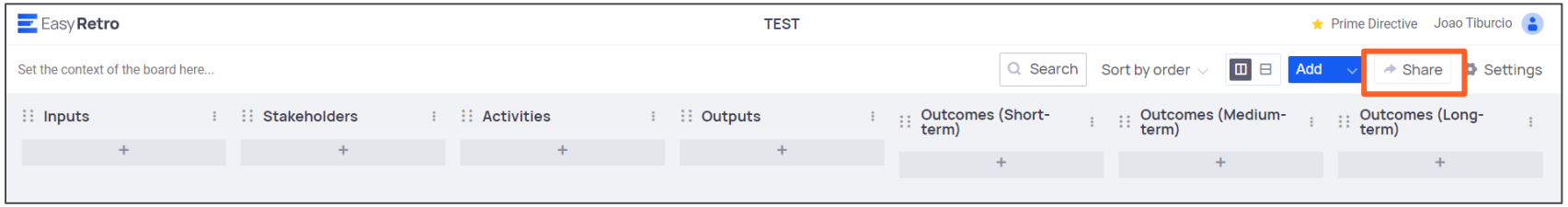


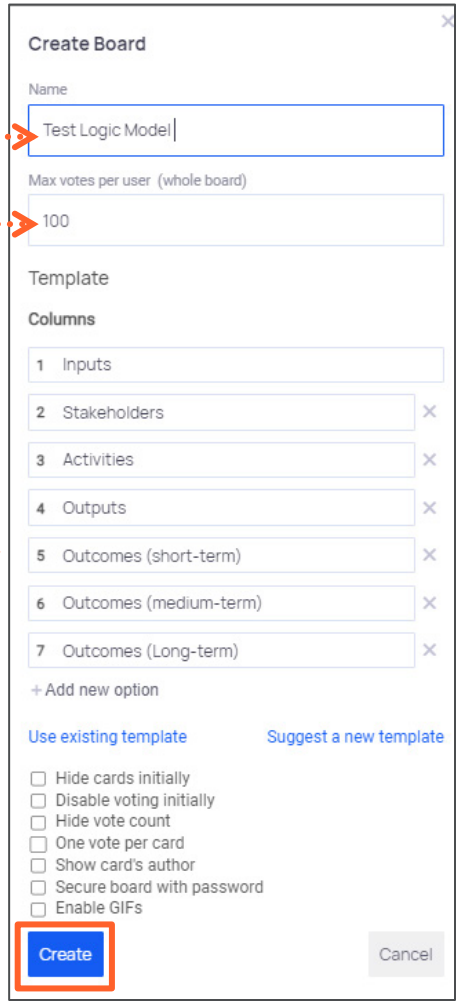
- You can engage with a wider group to help you prioritise the outcomes in your logic model
- Ask your stakeholders to help you by voting on the outcomes in your logic model (tools like EasyRetro are interactive and allow participants to vote with a thumbs up) – see the attached guide to the right on setting up a logic model board in EasyRetro Guide
- Ask your stakeholders to score each outcome on a scale of 1-5 for importance and measurability, where 5 is most important or most measurable, then place these on a 2x2 matrix
- Outcomes with the most no. of votes and those that appear in the top right-hand corner of the matrix (i.e., considered both important and measurable) should be prioritised
- Outcomes that have fewer thumbs up or appear in the bottom left-hand corner (i.e., considered both unimportant and unmeasurable) should be given lowest priority in evaluation
- Outcomes that fall somewhere in between are pertinent for discussion with your stakeholders

Setting up a logic model

Set up your logic model for free using Easy Retro, this tool allows you to share the model for free for other team members to collaborate.

1. Navigate to http://www.easyretro.io/
2. Sign up for a free account (you get 3 free boards)
3. Click ‘Add Board’
4. Choose ‘Custom Template’ for type of board
5. Name the Board (e.g., Test Logic Model)
6. Change Max votes per user to whatever number you want
7. Add columns (in order); Inputs, Stakeholders, Activities, Outputs, Outcomes (Short-term), Outcomes (Medium-term), Outcomes (Long-term)
8. Click ‘Create’
9. The output is an empty Board with the 7 Headers in place
10. Share the board for collaboration - use the “Share” button to copy the URL or crop the QR code



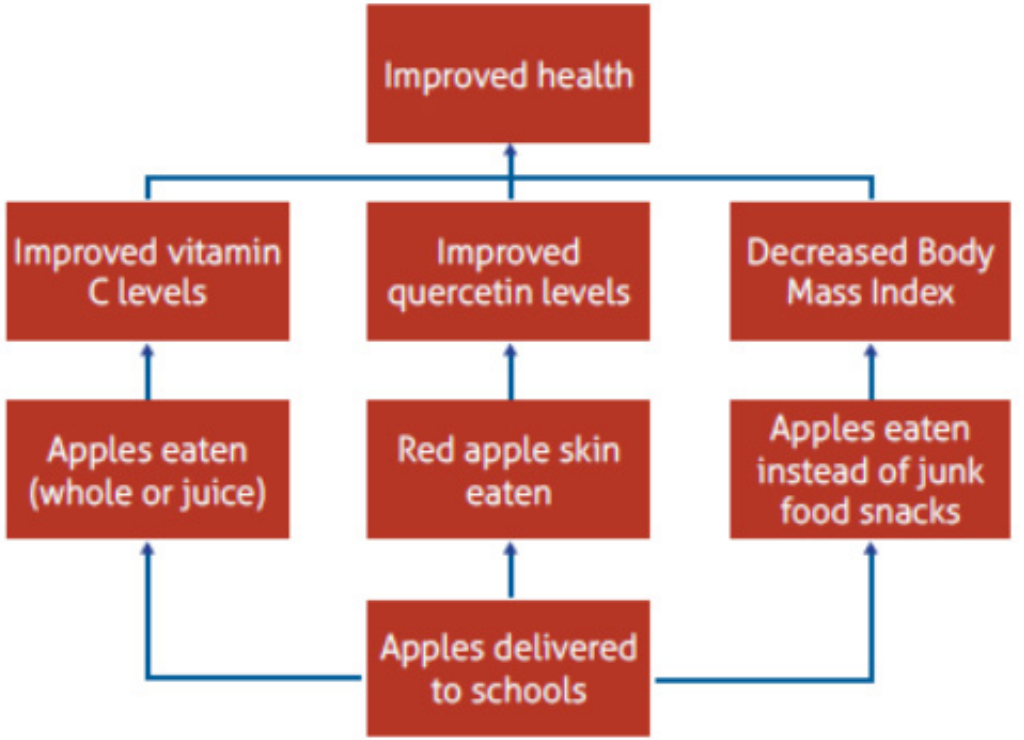


NOTE: The QR code can be shared in a meeting with your stakeholders to allow them to contribute live to the LOGIC model during the discussion and can be accessed by mobile phone or tablet

Linking output and outcomes using an outcomes hierarchy

An outcomes hierarchy can be used to show **cause and effect**, or **temporal sequence**, in an easy-to-understand way. It helps to highlight the '**golden thread**' from output to outcome.

The hierarchy should start with an output or outcome which is **fully controllable by the intervention**; should end with a **higher-level outcome**; and should have a mid-level output or outcome in between that helps **evidence a causal link**.[1]



An outcomes hierarchy for 'an apple a day' program showing possible different causal paths[2]

1. http://www.parkerduignan.com/ot/untitled_text_page.html
2. https://www.betterevaluation.org/en/evaluation-options/outcomes_chain

Start with Why – Checklist

- [ ] Have you used a logic model or similar process to think through the process/output and outcome measures of interest for your evaluation?

- [ ] Have you thought through how the outputs lead to outcomes in a logical flow? Guide

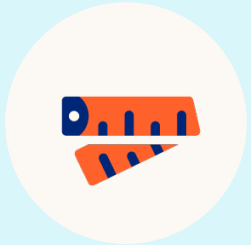
---

- [ ] Have you checked the outcomes ladder guide? Guide

- [ ] Have you voted and ranked your most important and measurable outcome measures? Guide

---

- [ ] Do you now have an agreed, prioritised and achievable list of outcome measures?



Measures

What data do you have and what do you need?

Work through the measurement checklist, find out what data is already available and decide what data you need to collect

Step 1: Measurement Checklist

- ✓ Can I create and agree an evaluation framework before my programme or intervention starts?
- ✓ Am I clear about the outcomes I want to measure?
- ✓ Will my outcomes enable me to demonstrate I have met the aims and objectives of the programme or intervention?
- ✓ Do my outcomes align to at least one of the quintuple aims (better health, better experiences of care, cost-effective care, better staff experiences or improved health inequalities)?
- ✓ Over what period of time do I anticipate to see change in outcomes or outputs ? Over a short period, (e.g., less than six months), intermediate period (6-12 months) or longer-term (over a year or more)?
- ✓ Can I start my measurement before my programme or intervention begins so I have a baseline?
- ✓ Can the people who go through my programme or receive my intervention be identified and tracked?
- ✓ Can I take measurements repeatedly during my programme or intervention and in follow up? Guide

Step 2: Talk to a local analyst – find out what data is already available for each of your outcomes

Contact for access or support via SWL's Health Insights Tool — isl.support@nhs.net

Example patient data:

- Patient characteristics: age, sex, ethnicity, deprivation
- Patient activity and clinical data: acute hospital, community health, mental health, IAPT, 111, referrals, waiting list

Other available secondary data sources Guide

Step 3: Are there readily available, routinely collected, measures that closely align to the outcomes you want individuals to achieve?

All the data I need is available for my outcomes Guide

Next

I need some more data from things that can be quantified (e.g., activity, costs, diagnostic tests)

Next

I need some more data for things that can be described (e.g., experiences, thoughts, feelings, reflections)

Next

Measuring outputs and outcomes

Consider

1. What outcomes do you want individuals to achieve through a specific action?
2. Are there readily available, routinely collected, measures that closely align to the outcomes you want individuals to achieve? If yes, see *secondary data* options below.
3. If there are no readily available measures, see *primary data* options below.

Guide

Primary data (collected first-hand)	Secondary data (collected already for other purposes)
Questionnaires/surveys	Administrative, e.g., GP consultation data, acute data (via Secondary Uses Service (SUS) or otherwise)
Interviews (telephone or face to face)	Patient satisfaction surveys
Focus groups	Historical data allowing for baselining (pre-intervention measurement)

Advantages and disadvantages of primary and secondary data

	Primary data	Secondary data
Advantages	• Can be tailored to your outcomes • Provides richer data about how an intervention is working • Can provide insights on the intervention relatively quickly • Data quality can sometimes be more easily controlled • Validity can sometimes be more easily controlled Guide *Using surveys to collect quantitative data – principles and considerations* Guide *When to collect qualitative data*	• More easily available • Often quite comprehensive information on intervention participants • Wide range of measures available • Timely, established, and specific
Disadvantages	• Time-consuming • Resource-intensive • Requires infrastructure	• Measures may miss outcomes you want to evaluate • Difficult to dig into **why** an intervention works • Data quality can be variable • Insights can take time to arise from available measures

Access the Evaluation Toolkit here

Secondary data sources

Publicly available secondary data sources

Office for Health Improvement and Disparities: Fingertips: Public health profiles - OHID (phe.org.uk)

Office for National Statistics: Home - Office for National Statistics (ons.gov.uk)

NOMIS: Nomis - Official Census and Labour Market Statistics (nomisweb.co.uk)

NHS England: Statistics » Statistical work areas (england.nhs.uk)

NHS England: Statistical publications, open data and data products - NHS Digital

UK Data Service: UK Data Service

Locally available secondary data sources

Contact for access or support via SWL's Health Insights Tool: isl.support@nhs.net

Measurement picklists

"Trusted measures: analytical resources for integrated care", Nuffield Trust 2018 - https://www.nuffieldtrust.org.uk/resource/trusted-measures-analytical-resources-for-integrated-care

"Pic 'n' mix: an introduction to choosing and using indicators", King's Fund 2012 - How to commission for improving health outcomes: an introduction to choosing and using indicators (kingsfund.org.uk)

Quantitative measures

Focus on what and how many (numeric), include administrative sources such as Hospital Episode Statistics/Secondary Uses Service, GP practice data, surveys, questionnaires and structured interviews.

Type of measure	Example	Pros	Cons
Counts	The number of people attending a chair-based exercise class per month	Simple to understand	Makes comparison over time or between populations more difficult, especially if your denominator (e.g., population size or number of hospital beds) varies
Rates	The number of people A&E attendances per 1,000 population	Allow comparison over time and between population groups	Need to be calculated You need to know the total population size to convert back to a count
Ratios	Standardised mortality ratio: observed mortality: expected mortality	They are easy to interpret and to calculate and illustrate the relative quantity of one item in relation to another of the same units	The units must be the same on both sides of the ratio i.e., number of available hospital beds to the number of occupied hospital beds
Percentages	The percentage of people aged 25-34 who were admitted to hospital in the last 12 months	Allow comparison over time and between population groups	Need to be calculated and the numerator and denominator are not obvious
Averages	The number of GP appointments per patient per year	Allow comparison over time and between population groups	Need to be calculated
Process measures	These are also known as outputs and capture activities or things that are created through a service or intervention e.g., the number of care plans that were created	Help to identify that an intervention is live and to determine which patients are in receipt of the intervention. The type of measurement typically falls into one of the five above.	Do not represent outcomes in and of themselves
Outcome measures	The intended outcomes of the intervention should align to the goals and objectives of your programme e.g., reduction in A&E attendances for diabetic patients (the process measure, HBA1C levels, is distinct from the outcome measure for this intervention)	Indicates whether the intervention has delivered an intended outcome. The type of measurement typically falls into one of the five above.	The challenge is attributing an outcome measure as a true outcome of your intervention i.e., it was caused by your intervention and would not have occurred without the intervention

Using surveys to collect quantitative data – principles and consideration

- Decide on how the survey will be administered e.g., online, by email, on mobile/device, on paper, interview led in person or over telephone
- If digitally administered, think about the software you will use
- Be clear on the purpose of the survey
- Begin the survey with more straightforward questions that ease the respondent in
- Ensure that every question relates to the overarching purpose
- Design clear, concise and neutral questions
- Make sure that all questions are direct and not convoluted
- Try to avoid asking for two things in the same question
- Try and use response scales where possible as these capture the intensity and direction of the response
- Response scales should be balanced with equal number of response that are positive or negative e.g., excellent, very good, poor, very poor
- Avoid yes/no and other binary type questions that limit the amount of data you collect
- Keep the survey as short as possible
- Consider pre-testing the survey before administering to your target population

Further reading

'Writing an effective questionnaire'
– NHS England
bitesize-guide-writing-an-effective-questionnaire.pdf (england.nhs.uk)

'Guide to evaluation design, principles and practice'
– Midlands Decision Support Unit
Guide to Questionnaires for Service Evaluation - Evaluation Works (nhsevaluationtoolkit.net)

Using surveys to collect quantitative data

- Aligning process/output and outcome measures to questions
- For example, through a logic model process, a long-term outcome of an intervention is identified — "Improvement in patients' overall wellbeing scores". If there is no existing data to measure, monitor and evaluate according to patient wellbeing, then there is the option of collecting data quantitatively via survey
- There are many available wellbeing scales; one of the more simply and widely used is the ONS4: Personal Well-being ONS4 measures - Evaluating wellbeing (whatworkswellbeing.org) as seen below:
    - For each of these questions I'd like you to give an answer on a scale of 0 to 10, where 0 is "not at all" and 10 is "completely"
        1. Overall, how satisfied are you with your life nowadays?
        2. Overall, to what extent do you feel that the things you do in your life are worthwhile?
        3. Overall, how happy did you feel yesterday?
        4. On a scale where 0 is "not at all anxious" and 10 is "completely anxious", overall, how anxious did you feel yesterday?

Other wellbeing measures:

What Works Centre for Wellbeing - Wellbeing Measures Bank - Evaluating wellbeing (whatworkswellbeing.org)

When to collect qualitative data

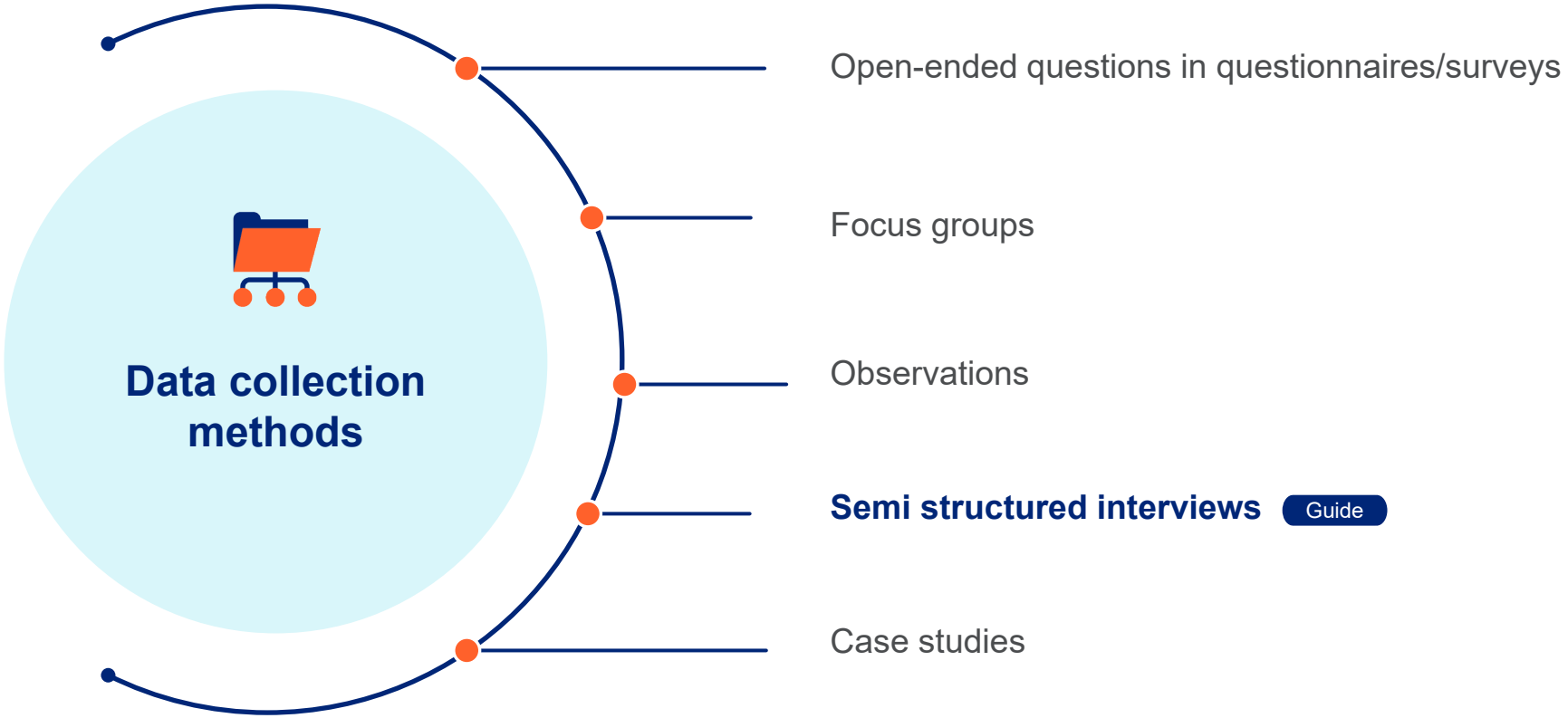
Qualitative methods focus on *why* and *how* (narrative). These might include focus groups, semi-structured or unstructured interviews and observational study, all which can be used to give depth on how and why an intervention is working and provide context and depth of insight that can be used to improve an intervention.

Qualitative methods help to:

- Answer descriptive and exploratory questions
- Explore different experiences related to diversity
- Understand complex issues
- Develop theories and test hypotheses
- Understand the impact your intervention may have on participants
- Determine the outcomes you may want to measure over a longer period
- Give a voice to those who are less often heard.

Types of qualitative data collection methods

The following section focuses on semi-structured interviews, the most common method for collecting qualitative data for service evaluation.
References and resources on other forms of qualitative data collection can be found here.



A common approach: semi-structured interviewing – process steps and principles

- Can be undertaken by telephone, face to face, or virtually
- Are typically 45-75 mins long but this depends on the topic matter and the participant

The 2 'Ps' for successful semi-structured interviewing

Process

- Use community engagement and data to inform intervention design
- Design semi-structured interview questions or topic guide in line with what you are trying to learn or achieve from the interview
- Undertake interviews
- Transcription
- Thematic coding
- Generate a thematic map – themes, subthemes and relationships between them
- Theory of change/logic model and causal mechanisms

Principles

- Use open-ended questions to get descriptive answers
- Specific ‘Why’ and prompt questions
- Questions should help answer your evaluation questions
- Use language that participants can easily understand
- Keep questions as short as possible, ask clarifying and follow-up questions
- Do not phrase questions as negative, keep neutral
- Always ask important questions first
- Be flexible with order
- Know your guide backwards
- Try for natural conversation and use active listening

Example interview questions for understanding physical activity behaviour change

- What did you do as part of the programme?
- How did you incorporate this into your day-to-day life?
- How are things now regarding your physical activity habits?

For those who describe increased/regular physical activity engagement:

- What has helped you to make these changes?
- Have there been any challenges to you staying active?
- Why have you made these changes?
- Has anyone else helped you to stay active?
- Did you have to make changes to other parts of your life to incorporate more physical activity?
- How does your activity impact on you? Prompt — how does it make you feel?
- Does it affect any other areas of your life?
- In what way is physical activity important to your everyday life?

For those who describe an initial increase but then a decline in physical activity levels:

- Why do you think your activity levels reduced?
- Would you like to be doing anything differently?
- If yes – what are the main things that prevent you from doing so?
- How do your physical activity habits compare to times before the programme?

For those who describe no change in physical activity levels:

- Would you like to be doing anything differently?
- If yes – what are the main things that prevent you from being more active?

Grimmett el al, 2020; "Exploring maintenance of physical activity behaviour change among people living with and beyond gastrointestinal cancer: a cross-sectional qualitative study and typology," ***BMJ Open Vol.10 (10): e037136***

Example interview questions for healthy eating behaviour change

Defining Healthy Eating

We often hear a lot about healthy eating. Can you share your thoughts on what you describe a healthy food choice to be?

How important is it for you to make healthy food choices?

Perception of own food choice

1. Every day we make food choices. How would you describe the food choices you are making currently?
2. Can you describe from your experiences, times in your life where your food choices have changed?

Drivers of food choice

3. Thinking broadly, what do you think influences people to make healthy food choices?
4. Thinking about your own food choices and experiences, if you were to choose the top three things that influence you to make healthy food choices, what would they be?

Barriers of food choice

5. Again, thinking more broadly, what do you think interferes with people making healthy food choices?
6. Can you share your thoughts on what you think interferes with you making a healthy food choice?
7. Do you have any thoughts or suggestions as to how these [barriers] can be addressed?

Mete et al, 2019; "What is healthy eating? A qualitative exploration," **Public Health Nutrition Vol. 22 (13): 2408-2418**

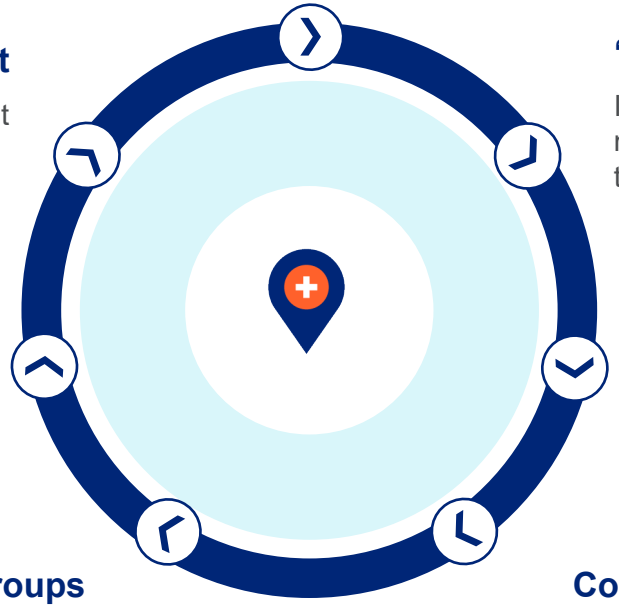
Methods for qualitative data collection in relation to community engagement

Information gathering

(Internet searches, site visits to identify local service providers and events where the community frequent) – attend events and get to know the local area, communities and stakeholders

Measure the impact

Of your community engagement



‘Go-along interviews’

Identified individuals show you around their neighbourhoods, tell you about local people and introduce their lives

Community working groups

Organised via the community champion

Snowball sampling

Community representatives help to identify and engage individuals within the community – including those who are not in touch with formal services

Community focus groups

Organised via the community champion

Community champions (or ‘influencers’)

Well-connected individuals who may believe in the rationale for intervention and can help recruit

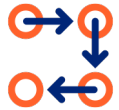
Qualitative approaches inform quantitative approaches and vice versa

Inform hypotheses

Qualitative data can inform your hypotheses that you then test through quantitative methods

Inform evaluation questions

Qualitative data can inform the evaluation questions that you answer through quantitative and qualitative methods



Quant informing qual

Quantitative methods may influence the community/population of interest for qualitative approaches during evaluation

Quantification

Quantitative methods can also help quantify how prevalent a particular view, issue or typology is that you identified during the qualitative phase

Mixed methods

Quantitative methods and qualitative methods (described as mixed methods when brought together) provide a more holistic and robust view of intervention efficacy – what has changed **as a result of** the intervention, **how** and crucially **why?**

Understanding why

Only if you understand the why, can you confidently invest in scaling an intervention – and maximise the effective elements, whilst minimising the ineffective elements

[Combine qualitative and quantitative data - Rainbow Framework (betterevaluation.org)]

Final measures checklist

- ☐ Do you have a plan for using existing data sources to track and monitor your selected measures?

- ☐ Do you have a plan for collecting new quantitative or qualitative data if required?

- ☐ Have you assigned roles and responsibilities for measurement as part of your evaluation action plan?

Things to think about

What other considerations go into evaluation design?

What are the evaluation questions you are trying to answer?

Looking at an example:

'Felix PCN are looking at 230 patients with obesity. This cohort of patients will be given lifestyle advice via a Health Coach, with the aim they are empowered to improve their health by changing their lifestyle habits, e.g., increased physical activity levels, eating healthier foods'…

Questions we need to ask:

Programme effectiveness

- Is the programme effective in supporting people to increase their levels of physical activity?
- Is the programme effective in supporting people to eat more healthily?
- How do we know the project/intervention made a difference? Can we use outcomes to help determine further intervention design and implementation?
- What does the data tell us about the effectiveness of the programme?
- What are the critical success elements of the programme?

Participants

- Why do participants change their behaviour?
- Why do some participants not change their behaviour or sustain positive behaviour change?
- Are factors that are intrinsic to the individual relevant in whether the individual changes and sustains more positive health behaviour?

Other things to think about now you're ready to begin designing your evaluation

Things to think about checklist

- [ ] Who are the key stakeholders and audience for the evaluation findings and are their expectations realistic?
- [ ] Do you have a sense of how long it will take to achieve the intended outcomes of the intervention?
- [ ] What are the decision points and what needs to be known before any decisions can be made and what is the governance process?
- [ ] What is the requirement for evidence, are you able to make a judgment about the balance of practicality vs robustness of the evaluation?
- [ ] Don't reinvent the wheel – are you aware of how others have approached evaluation of similar interventions?
- [ ] Is it possible to begin the evaluation before intervention implementation?
- [ ] What are your resources and who will undertake the evaluation? Will it be internal or commissioned externally?
- [ ] Will there be a mechanism to identify intervention participants in the data analysis phase?
- [ ] How will evaluation findings be communicated to key stakeholders and audiences during and after the intervention?
- [ ] Have you considered whether you will need an ethics review? Do I need NHS Ethics approval? (www.hra-decisiontools.org.uk)
- [ ] Have you assessed the risks of evaluation and added them to a risk register?

Design

Best possible design of your evaluation

Evaluation design

The evaluation design describes how and when you collect and analyse data and whether you compare outcomes in the intervention group to a control group. Deciding which evaluation design is most appropriate for your intervention should be based on why you are evaluating and what quality of evidence you need.

Ask yourself and your stakeholders:

Has the intervention not started? Next

Has the intervention started but you haven't taken baseline measurements? Next

Could you retrospectively identify a matched control group to compare outcomes against? Next

Is it possible to randomise participants to intervention and a different intervention or control group before the intervention has started? Next

Have you checked the Hierarchy of Evidence to assess the quality of evidence your evaluation may produce? Next

Pre-experimental design

(before, during and after evaluation focusing only on an intervention group)

Glossary

What is it?

A prospective evaluation where data is collected and analysed before the intervention is implemented (baseline). Further data is collected during the intervention and to track outcomes once the intervention has ended. There is no control group, and all measurement focuses on intervention participants.

How do you do it?

- Agree your plan for data collection from intervention participants including baselining, tracking and follow up
- Begin collecting data before the intervention starts
- Track measures during and in follow up to the intervention
- Analyse changes in the measures (using paired t-tests or repeated measures analyses of variance, like one-way repeated measure ANOVA)

Guide

A practical example

Tracking wellbeing scores three times for a wellbeing coach intervention; once before intervention, once at the intervention midpoint and once three months after the intervention has finished.

Advantages	Disadvantages
• It is a relatively straightforward design • You only need to track the outcomes of one group – the intervention group • There are no ethical concerns with denying access to care for a control group	• The lack of a comparison/control group means this design lacks internal validity and attributing outcomes to the intervention is difficult • It is difficult to control for regression to the mean • It assumes the experience of the intervention is the same for each participant and does not account for any changes to the intervention

Paired Samples T-test in R - Easy Guides - Wiki – STHDA
11: Introduction to Repeated Measures - Statistics LibreTexts

Retrospective evaluation without a comparison group

Glossary

What is it?

A backwards look at whether the intended outcomes of the intervention were achieved. This requires knowledge of the intended outcomes and the ability to identify the intervention group. It also requires the use of secondary data (data that has already been collected about the participants) to determine impacts.

How do you do it?

- Identify all your intervention participants
- Find the secondary data you can use for measures that align to the outcomes you're interested in to show change over time
- Your secondary data source needs to enable you to observe those measures for your participants pre and post intervention
- Collect data in retrospect, asking intervention participants to recall information

A practical example

A falls prevention program has already begun and the metrics of interest are data points that can be found in existing secondary data sources such as GP records and hospital data

Advantages	Disadvantages
• It is possible to evaluate your intervention even if it has already started • It takes less time to complete and you do not have to wait for the results	• It is not possible to prove your intervention caused the outcomes because you cannot tell whether participants would have experienced those outcomes without your intervention • You cannot control for extraneous or confounding factors that may also impact the outcomes you are interested in • Your participants may experience recall bias when you are collecting information about events, thoughts, feelings, emotions in the past • You're likely to be reliant on data collected for other purposes (secondary data) that may not align accurately to what you are trying to measure

Paired Samples T-test in R - Easy Guides - Wiki – STHDA
11: Introduction to Repeated Measures - Statistics LibreTexts

Quasi-Experimental Design

(intervention group compared with a matched control group)

Glossary

What is it?

An evaluation that compares outcomes in your participants with a suitable control group that you choose (before, during or after your intervention).

How do you do it?

- Identify your participants and collect baseline data (if applicable)
- Track, monitor and feedback outcomes data during and after your intervention, in line with your plan and for both your intervention and matched control if you have one at the start
- Match your control group according to age, sex, ethnicity, deprivation and comorbidities
- Think through other variables or things you haven't measured that might influence the effect you're trying to evaluate (confounders) that you want to try and control for in multivariable regression analysis
- Use the logic model or theory of change to inform the evaluation analysis
- Undertake analysis to test for significance in outcomes between intervention and control group

A practical example

A health coaching intervention to encourage individuals to improve their diet and take physical exercise is offered to a cohort and outcomes for that cohort are tracked and compared to a similar control group chosen as the intervention is carried out.

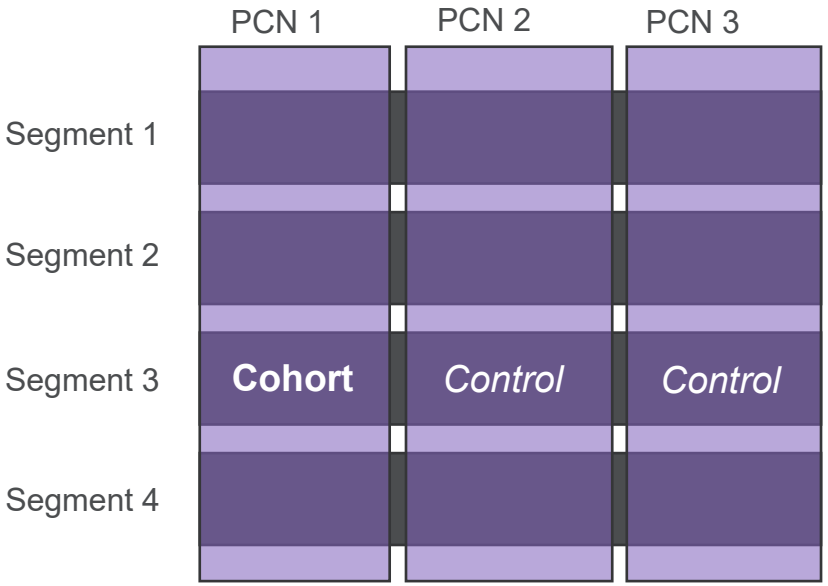
Advantages	Disadvantages
• A control group can be easily selected by matching the characteristics of individuals and choosing a different geography • There are no ethical concerns if a matched control group is selected retrospectively • You can adjust for the effects of confounding variables using multivariable regression	• May still be difficult to prove causality between intervention and outcome without random allocation to a control group • The effect of confounders may be more likely in one of the groups because allocation is not random • Evaluators may introduce bias in selecting the control group or might miss some of the potential confounding factors in selecting their control

Guidance on undertake multivariable regression can be found here - Introduction to Multivariate Regression Analysis (https://www.mygreatlearning.com/blog/introduction-to-multivariate-regression/)

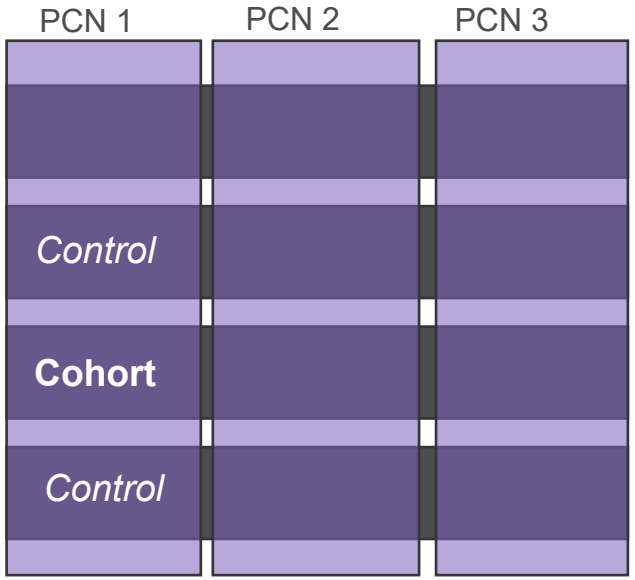
Choosing control groups

The simplest and best control groups are often people with the same criteria from neighbouring geographies who have similar demographics and experience similar health challenges. When you look at the data, check to make sure there are no individuals included in both intervention and control groups.

People in the same segment, but different PCNs or geographies



People in similar segments (or subsegments)



Potential challenges and opportunities:

- Not viable with whole system intervention
- Relies on having data for non-participating PCNs or geographies

- Potential variance between control and intervention groups in terms of demographics and health risk
- Important to compare the make-up of control and intervention groups closely to ensure they are comparable

Experimental Design

(randomised in intervention group and control group selection)

Glossary

What is it?

Known as a randomised controlled trial, experimental evaluation involves random allocation of individuals to the intervention or the control group.

How do you do it?

- Find people to enrol in your study who meet your eligibility criteria
- Randomise this group to intervention or control (no intervention), or to two different interventions if you want to compare two options and it is not possible to have a control group
- Track and monitor outcomes between the two groups
- Analyse and test if outcomes are statistically significantly different between the two groups

A practical example

Random assignment of individuals either to StopAdvisor (interactive website for smoking cessation) or an information-only website about smoking cessation.

Advantages	Disadvantages
• Random assignment of individuals to the intervention vs control group makes it more likely confounding variables affect each group equally • The effects of selection bias are reduced through random allocation • You can compare two different interventions by having a well-matched group of people randomly assigned to each	• There are significant ethical concerns with assigning individuals to a control group who do not receive an intervention that may benefit them – ethics approval will be needed to proceed • Seeking ethical approval can be time consuming • Evaluation requires more significant investment of resources compared to other designs • Evaluation requires a larger number of participants in the groups being compared

https://www.gov.uk/guidance/randomised-controlled-trial-comparative-studies

The hierarchy of evidence

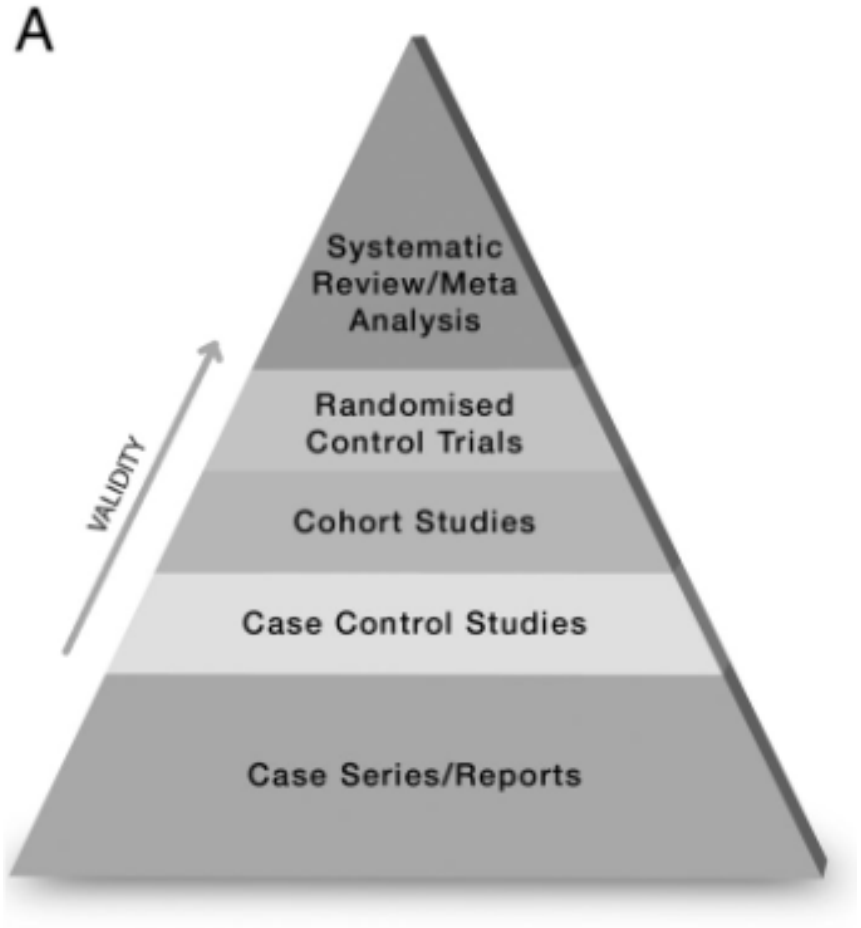
A hierarchy of evidence ranks the relative strength of conclusions from experimental research or evaluation.

The design of a study (e.g., retrospective evaluation) and the measurements (e.g., quality of life) affect the strength of the evidence.

The best possible evidence for an intervention is generally meta-analyses of randomised controlled trials (e.g., Cochrane Collaboration reviews). There is a high degree of confidence that results from these represent the true effect of an intervention.

Studies including control groups are better able to attribute outcomes to an intervention avoiding confounding factors (e.g., alcohol as a confounder to the relationship between smoking and lung cancer).
These types of evaluation create moderate levels of confidence in the estimated effectiveness of an intervention.

It is likely you do not need to design for the highest possible levels of confidence and can aim to design cohort, case control or retrospective evaluations that provide sufficient evidence for your service evaluation.



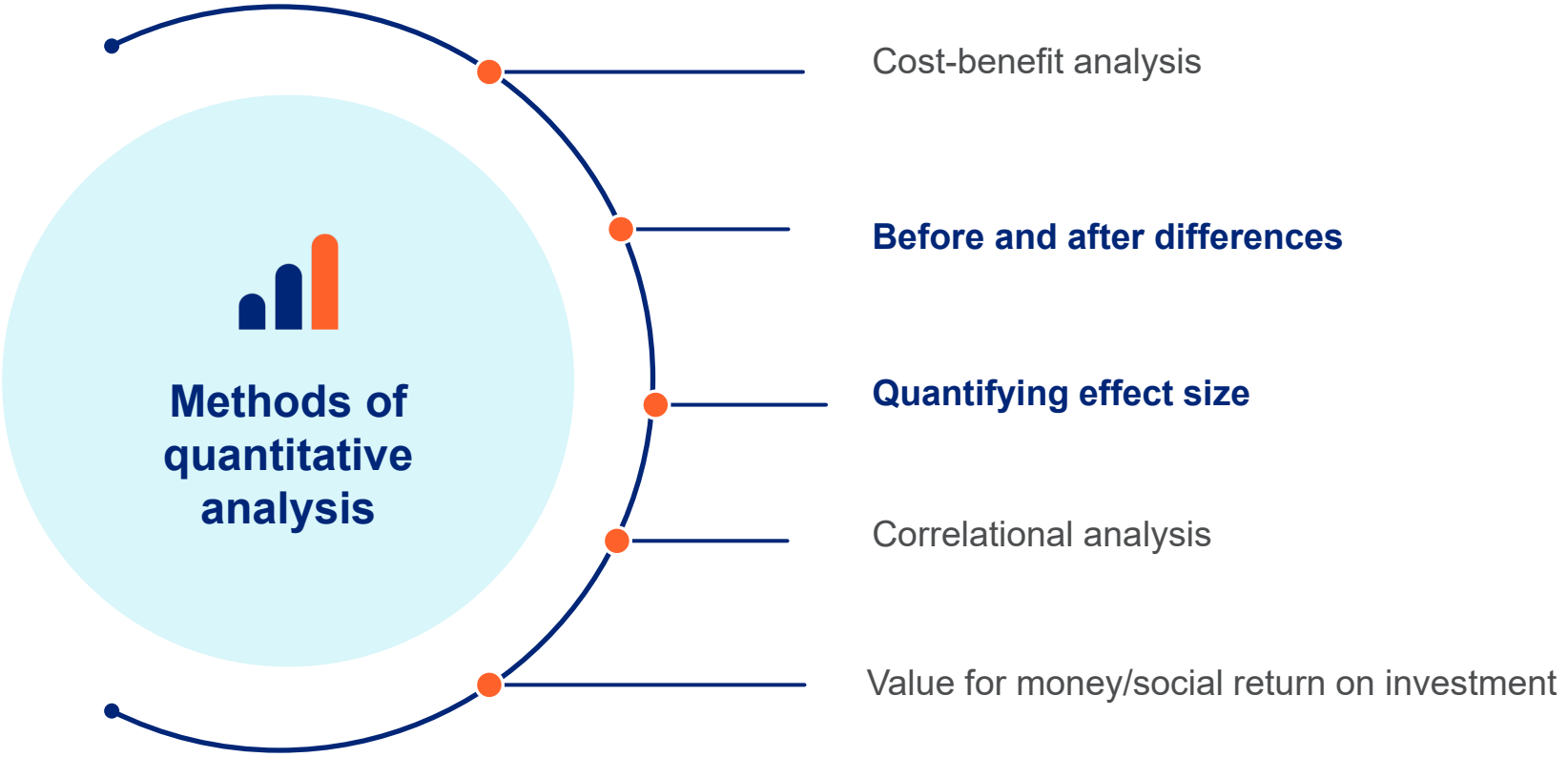
Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ (December 1995). "Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group". JAMA. 274 (22): 1800–1804.

Analysis

Bringing together your data

Analysing quantitative data

- Robust answers can be derived about whether an intervention has had an effect, provided there is sufficient sample in size and representativeness with an accurate comparator
- Sample size calculation is complex but here are 3 easy rules of thumb:
  1. Minimum sample should be 100.
  2. A good sample size is usually 10% if it does not exceed 1000.
  3. A sample of 1,000 is nearly always statistically robust
- Using robust samples is useful when wanting to understand trends in a larger population
- Can be performed on secondary data without the need for further data collection (though often benefits from more collection of data).



There are many other forms of quantitative data analysis that support intervention evaluation. References and resources can be found here.

Before and after differences

As described in the pre-experimental and retrospective design slides, there is a need to examine changes in measures over at least two points in time – usually before and after intervention. This can be done simply by comparing the average in a measure for the intervention group.

Example – the average wellbeing score on a 1 to 5 scale in the intervention group was 3.2 before intervention and 4.1 after intervention. So, we can say that the wellbeing score increased by 0.9 points on the scale.

However

- This form of straightforward comparison does not tell you whether the difference is down to anything other than chance.
- That can be tested using a paired t-test. The paired t-test is one of the more rigorous methods for comparing the difference in measures pre and post intervention.

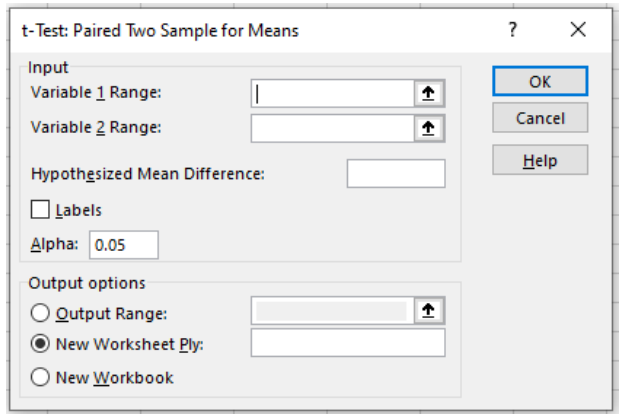
A paired t-test assesses whether the difference is statistically different (i.e., not just due to chance)

It can be used to compare the average (mean) difference between pairs of measurements

Paired Samples T-test in R - Easy Guides - Wiki – STHDA
11: Introduction to Repeated Measures - Statistics LibreTexts

How to undertake a paired t-test in Excel

- If you don't have access to the Data Analysis under the Data tab in Excel, you will need to download the Analysis Toolpak which is free - https://www.statology.org/how-to-load-the-analysis-toolpak-in-excel/
- Select Data Analysis and then the t-Test: Paired Two Sample for Means
- Input your first set of measures under Variable 1 Range and second set under Variable 2 Range
- Hypothesised mean difference is your null hypothesis which would typically be that there is no difference between the means and you are looking to disprove this – therefore the value would be '0'
- Then choose where you would like to output the results under Output Range
- In this example, we have 100 intervention participants with a wellbeing score between 1 and 5



- The outputs will look like this

t-Test: Paired Two Sample for Means

	Variable 1	*Variable 2*
Mean	2.94	2.99
Variance	2.137778	2.090808
Observations	100	100
Pearson Correlation	0.052269	
Hypothesized Mean Difference	0	
df	99	
t Stat	-0.24976	
P(T<=t) one-tail	0.401644	
t Critical one-tail	1.660391	
P(T<=t) two-tail	0.803288	
t Critical two-tail	1.984217	

- You can see the mean score in each sample (2.94 before intervention and 2.99 after intervention), the variance (a measure of spread) and number of observations
- **The important statistics here are the P two-tail and t Critical two-tail. The P two-tail shows whether the difference in means is significant – a value of less than 0.05 indicates that it is. In this example, the difference is not significant and therefore we can say that the change in the wellbeing score is as likely to be down to chance as it is the direct result of the intervention and therefore the positive change cannot be attributed to the intervention itself**
- The Pearson Correlation is a measure between -1 and +1 that indicates the strength and direction of the correlation between the means
- There is more information here about how to undertake this test in Excel — How to Conduct a Paired Samples t-Test in Excel - Statology
- There are other approaches for comparing differences between three or more observations as mentioned in the pre-experimental design slide. One such example is the ANOVA one-way repeated measure. There is more information on how to conduct such an analysis here - 11: Introduction to Repeated Measures - Statistics LibreTexts

Quantifying effect size between your intervention and control group in quasi-experimental design

- Effect size can be used to measure the standardised size of the difference between the intervention and control group for any given measure of interest
- The P-value alone only indicates whether any change in measures between the two groups is statistically significant (i.e., is down to more than just chance) and does not indicate the magnitude of the effect
- When your data is normally distributed – you can use Cohen's D to test the magnitude of the difference in measures between the two groups. This is a good statistic to use alongside a paired t-test to demonstrate a statistically significant difference in outcomes you are trying to impact through intervention.

0.2 = Small effect size | 0.5 = Medium effect size | 0.8 = Large effect size

Considerations

- The size of the sample (intervention group) — is the cohort big enough?
- The precision of the measures (the construct validity) will also affect how well what you are measuring reflects the concepts

Quantifying effect size between your intervention and control group in quasi-experimental design – an example

A&E attendance rates in the intervention group compared with a control group:

$$\text{Cohen's d} = (x_1 - x_2) / \sqrt{(s_1^2 + s_2^2) / 2}$$

x_1 , x_2: mean of control group and intervention group, respectively
s_1, s_2: variance of control group and intervention group, respectively

A&E attendance rates per 1,000 individuals in the intervention group compared with a control group:

$$\text{Cohen's d} = (3.4 - 0.9) / \sqrt{(7.2^2 + 2.4^2) / 2}$$

= 0.465 indicating a medium effect size for the intervention

= Mean difference is significant with a reduction in 2.4 A&E attendances per 1,000 individuals

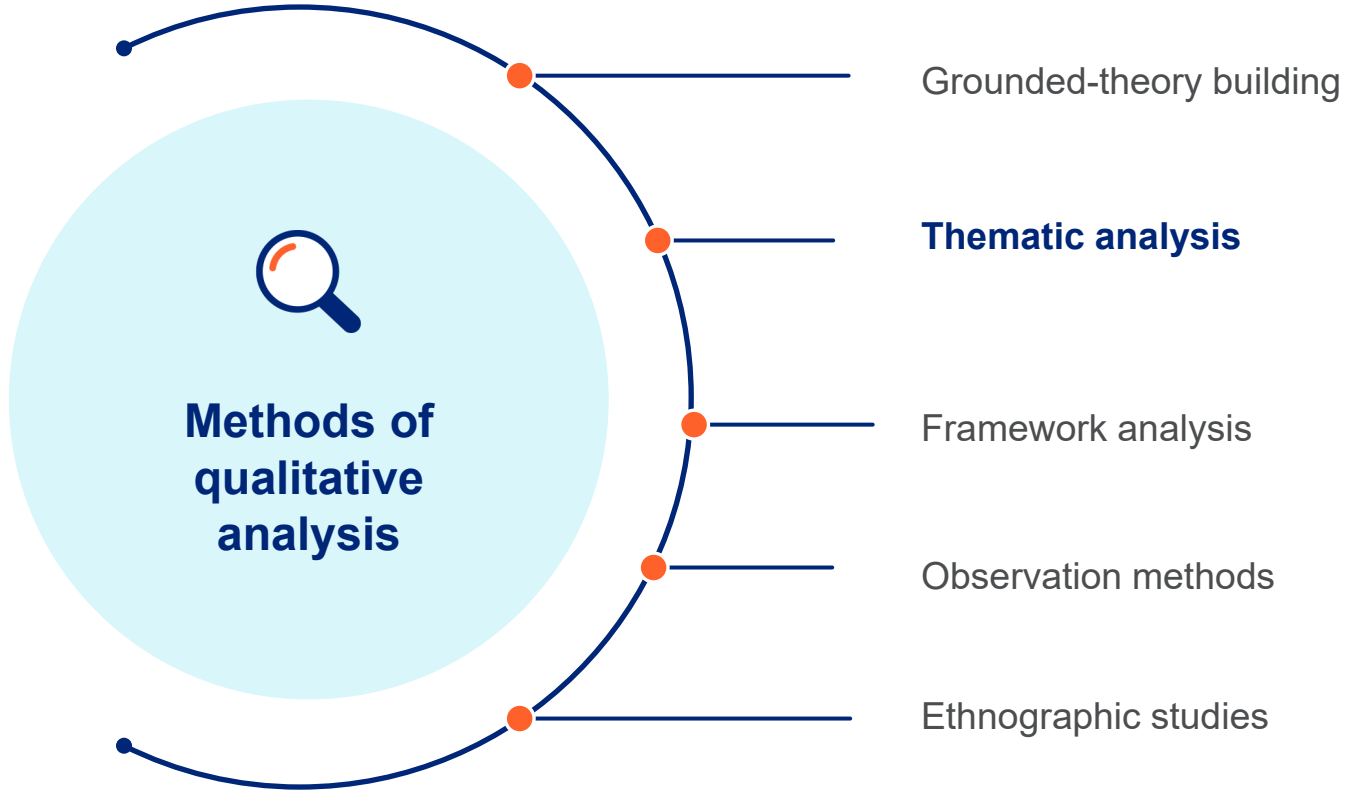
Analysing qualitative data

Qualitative data is non-numerical and unstructured. It could include text, such as open-ended responses to survey questions or user interviews, audio, photos and video, which all need to be transcribed into text files ready for coding and analysis. Coding is the process of labelling your data. Guide

You can then identify themes in the data. Guide

Codes and subcodes can then be analysed to understand the relationship between them. Guide

Further information on other forms of qualitative analysis can be found here.



Coding and theme generation

Step-by-step approach

Step 1: Follow your plan as to whether you want to collect qualitative data before, during or after your intervention

Step 2: Decide whether you will use deductive or inductive coding. Deductive coding - you create a list of predefined codes, and then assign them to the qualitative data. Inductive coding is creating codes based on the data itself and labelling them as you go – see below

Step 3: Read through the data to give yourself an overview. Now assign a set of codes to statements and sections of text

Step 4: Repeat step 3, adding new codes and revising the code description as often as necessary

Step 5: Once all the text and transcribed data has been coded, go through everything again, check for inconsistencies and make sure nothing has been overlooked

Overarching theme	Main code	Sub-code	Illustrative example
Healthy food choices are important, but not a daily priority 'Intellectually, I know I need to make healthy food choices all the time to lose the weight and to be healthy … In the moment when there's emotion, exhaustion, I find it very, very hard … So that's something I really struggle to disconnect what I know I should do, and what I want to do in the moment.' (P9)*	Food choices are important	Improved daily functioning	'[Be]cause I've gone from like eating reasonably healthy to not doing that at all and gaining a lot of weight and then I noticed it on my mental health and how it impacted me and then when I've gone from that stage back to eating healthy and exercising regularly and I've noticed I'm a lot happier.' (P10)
		Well-being	'I don't function as well if I'm not eating what I call healthy'. (P14)
		Improved physical activity performance	'It gave me the confidence to try new things, to meet new people, to get a new job, to lift up and go to uni … without the changes to food I'd still be in my rut.' (P17)
		Family	'Yeah, I think it's really important for me apart from being healthy and fit anyway, I have young kids, I need to be a positive role model for them, I don't want them to grow up you know, with poor food choices.' (P2)
	External barriers re-prioritised positioning of healthy food choices	Perceived lack of time	'It's very important but you've got to factor in time, at the moment time for me is impossible, I don't even know what day it is.' (P3)
		Daily stressors	'It's not always easy because it depends where you are. You just have to make do sometimes.' (P21)
		Poor energy	'Yeah, I suppose it's just you don't have the energy to think about it, trying a new recipe.' (P1)
Healthy eating information is known but can be difficult to apply into everyday life 'Because I do know what, I believe I know what healthy eating looks like, I just sometimes find it hard to translate it hard into practice and as, as you know that can be influenced by many things.' (P11)	Difficulty in translation	Confusion/uncertainty and fear (wrong choice)	'I find it confusing because although I want to lose weight, I want it to be something that I will actually enjoy.' (P12)
		Conflicting information	'The Internet is hopeless because everyone's got a different opinion, so it's very hard to find a similar answer to anything.' (P3)
		Choice	'I have heard it all, but when push comes to shove, if I've got given a choice of there have a salad [because] everyone is going to Macca's, or have a cheeseburger, I think I would go for the cheeseburger.' (P22)
Popular diets are used in attempts to improve healthy eating 'So, you know, I'm trying to eat healthy, I'm trying to get some exercise … I mean at the same time I was tossing up you know, do you do Opti-slim? Or do you just do the supplements?' (P2)	Solution	History	'I've tried a lot of different diets I suppose over the years and different ways of eating, like I've done juice cleansers and I've done an alkaline diet.' (P14)
		Success of others	'My mother was doing it and was working for her, so I just went okay, I'll give it a go.' (P2)
		Initial weight loss	'It was effective, I lost weight, but I felt awful on the inside and the weight came flying back on as soon as I started eating things that weren't you know so restricted, yeah I think I gained almost double what I lost.' (P2)
		Quick fix/easy	'I think people again, it comes back to their time and you know the whole idea that someone else can do it for you, like all I need to do is sign up for this thing and then they're going to tell me what to eat.' (P16)
	Unsustainable	Trust	'It was crazy, and I had payed these people and trusted them to put me on something that would help me lose weight and I assumed it to be healthy, in fact it was completely ridiculous.' (P9)
		Compliance	'I think they're too restrictive. They make it hard to live a normal life where people go, oh yeah, we're going to go out for lunch you know, they just don't build in, the usual expectations of life.' (P2)
Social media inspires and connect people with healthy eating 'I've done bits and pieces of research for myself like different blogs and things like that I enjoy following, they have more of a whole food ingredient for recipes, no chemicals or additives or that kind of stuff, I find it's more, it's just more realistic kind of recipes.' (P5)	Social media	Facebook/Instagram/blogs	'I think, you see, as women we see it all the time, healthy food eating, what to do, especially with so much social media especially with Facebook and Instagram people taking photos on Snapchat and all that.' (P19)
	Connection	Journey/experience	'Lots of stuff in the media, time of the year, you get new year's resolution in January … that come on the media and people are inspired.' (P6)
	Practical advice	Food alternatives	'I think that at the moment they are all looking to find alternatives to unhealthy food options, so I think that's good.' (P10)
		Food Ideas	'I think that it [social media] influences people and their health choices … promote good food choices like, I never tried that stuff for breakfast, I'll give that a go.' (P19)

Mete et al, 2019; "What is healthy eating? A qualitative exploration," ***Public Health Nutrition Vol. 22 (13): 2408-2418***

Example themes related to healthy eating

- Disease processes
- The role of ageing
- Confidence
- Emotions and psychological well-being
- Incorporating physical activity into everyday life
- Self-monitoring/ self-managing
- Support
- Motivation
- Social interaction
- Engagement
- Competing demands
- Activation
- Health behaviour change
- Staying power
- Experience of intervention
- Relationship and trust with health coach/dietician

Thematic map analysis

Step-by-step approach

Step 1: Lay out all the themes identified from the coding exercise

Step 2: Establish whether the data illustrates no relationship, a correlation or a possible causal relationship between the themes, using the data that you have collected (this can be tested through controlled statistical analysis)

Step 3: Map out links between themes

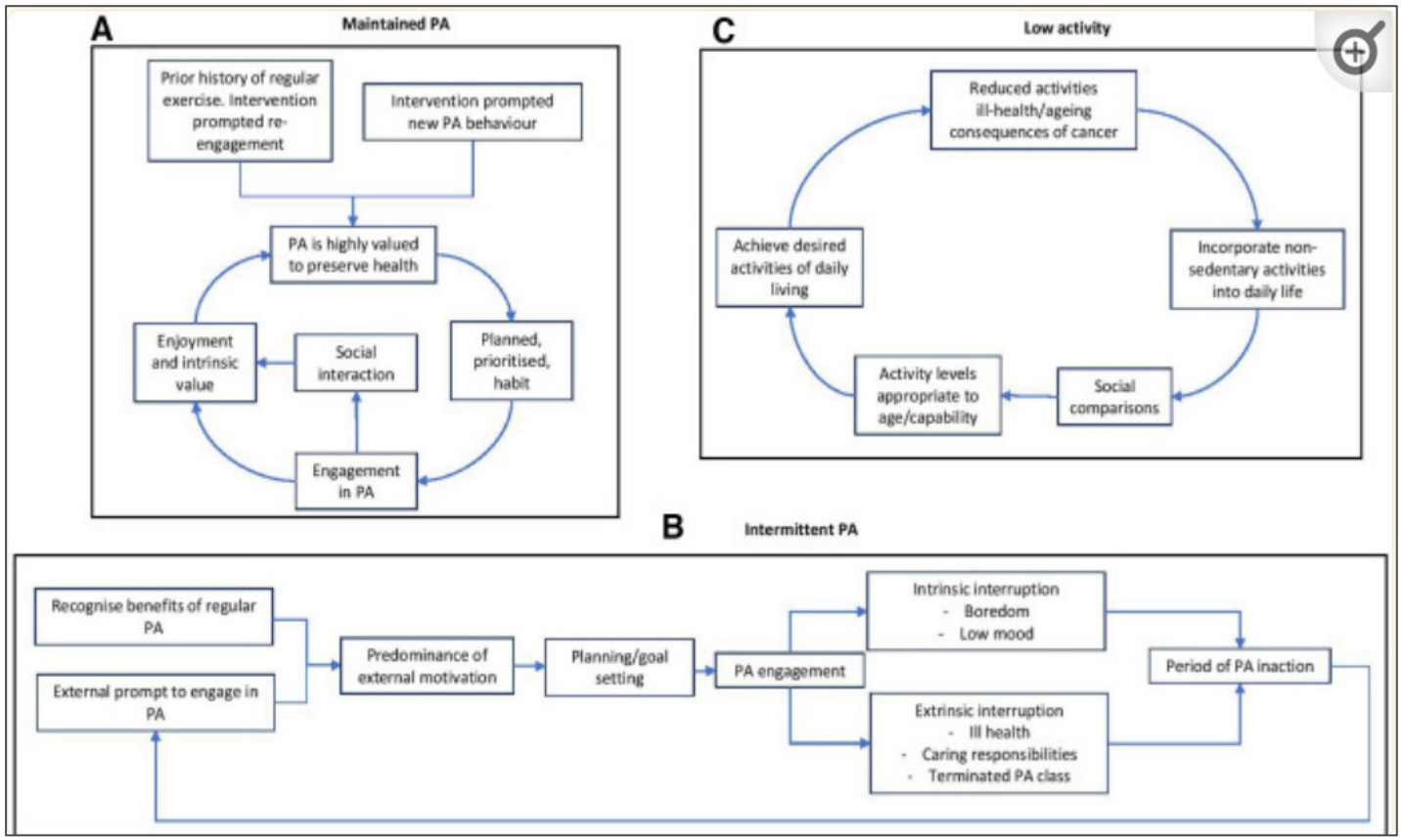
Step 4: Review the map to help you understand why your intervention may or may not have been successful in delivering the intended outcomes

Step 5: The example shows for three different types of interventions;

a) those who maintained physical activity

b) those who were intermittently physically active and

c) those doing no or minimum physical activity, how this type of analysis can help you understand outcomes



Grimmett el al, 2020; "Exploring maintenance of physical activity behaviour change among people living with and beyond gastrointestinal cancer: a cross-sectional qualitative study and typology," ***BMJ Open Vol.10 (10): e037136***

Further considerations checklist

- [ ] Will your evaluation produce valid and reliable information?

- [ ] Will your evaluation provide timely and relevant information for decision-making?

- [ ] Will you able to tell whether your intervention is/is not working and importantly why, not just what and how?

Drawing conclusions

So what?

Scaling the intervention – areas to consider

Intervention efficacy

- Is the absolute intervention effect size between intervention and control groups significant?
- How and why the intervention was successful/failed? Qualitative evaluation?
- Unintended/differential outcomes
- How did the intervention lead to the effects? Did the logic model play out as expected?
- Generalisability of effects are dependent on;
    - Sample size and representativeness
    - Study design
    - The selected control group
    - Participants
    - Setting
    - Processes and mechanisms

Evaluation and monitoring

- Scalability assessment;
    - Potential reach (rate of adoption over eligible population)
    - Viability - costs over perceived value to the health system/population
    - Recruitment of eligible population
    - Timeframes
    - Expectations
    - Sustainability
- What variables are important to measure over time whilst scaling?
- Measuring effectiveness throughout scaling
- Is it possible to measure the efficiency evaluation of intervention delivery?
- Monitoring fidelity through process measures
- Conceptualisation of value to the ICP/ICS

Template report card

What did we do? (effort)		How well did we do it? (effect)	What difference did we make? (impact)	
Engagement inputs (What did we do?)	Engagement outputs — what was the output of what we did (participants and activities)	Outtakes — quality, how well did we do it?	Engagement outcomes — what did we achieve — consider short term, medium term and longer term	Project/programme impact — what was the bigger impact?

Example evaluation plan on a page – Shanklin Village Health & Wellbeing Programme

Rationale for intervention

- Increase awareness of health screening
- Increase access to care
- Improve health and wellbeing
- Reducing loneliness/social connectedness in the community
- Core20Plus5

Cohort for intervention

- Shanklin residents
- Interested in exercise but struggle
- Wheelchair bound
- Underserved community
- CASS has health inequalities
- Number of outpatients
- Number of GP

Logic model/intervention design process exercise

- 6-9 months start a conversation, build trust/relationships – piggy backed on existing community meetings
- Understanding the needs – through community engagement – working through the community connectors e.g., residents association/local council

Top logic model output measures

- Set up health inequalities
- Agreeing health topics
- Number of people attending
- Number of people returning
- Engaging with health services
- People attending the whole session
- More diversity in attendees
- Positive feedback

Top logic model outcome measures

- Chair exercise – strength changes
- Weight measurements
- Diet/food related measurement
- Healthy eating metrics?
- Wellbeing?

Secondary data to be analysed

- GP records – all residents are registered with their GP
- Community pharmacist interaction
- Health coach, social prescriber, mental health practitioner, local councillor
- Baseline activity and prevalence data from Health Insights Tool e.g. on chronic pain, COVID boosters, deprivation, LTCs

Data to be collected

- Chair exercise – strength changes
- Weight measurements
- Diet/food related measurement
- Healthy eating metrics?
- Wellbeing?

Evaluation considerations

- Audience are people living on the Shanklin estate identified as a deprived community
- Single parents, elderly, disabled, BAME, etc.
- Hard to engage with this community- use connectors and estate champions
- Map out roles and responsibilities for data to be collected

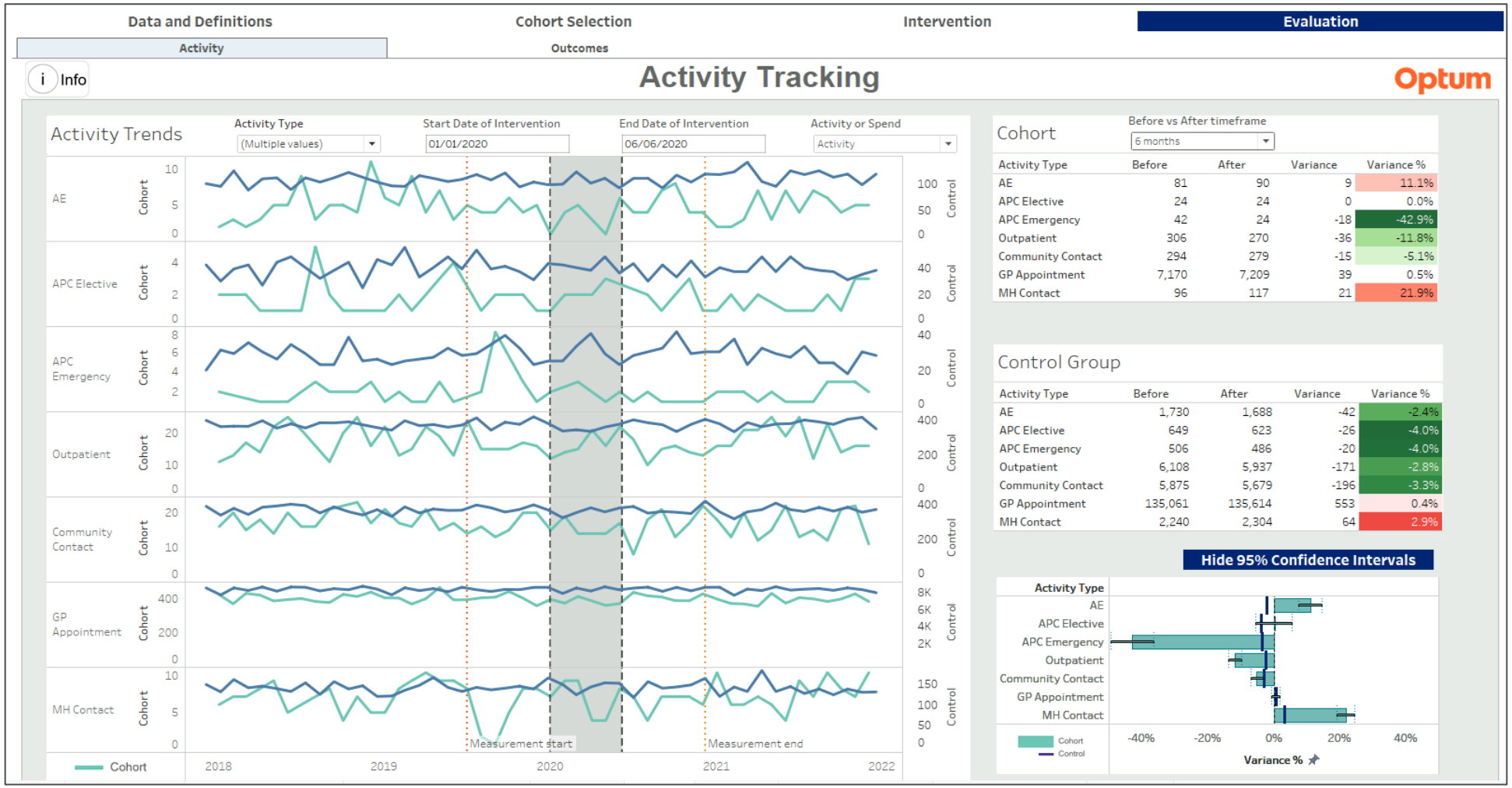
Evaluation questions

- Using community engagement via Community Connectors to understand and frame EQs
- Understanding what is important to the community
- Agree measures with the community

Evaluation design

- Baseline questionnaires at the start of the intervention, 3 months and 6 months into the intervention to measure progress
- Health questionnaire at baseline
- Case studies

An example: Our PHM Reporting Suite Evaluation Module



Glossary of terms

Term	Definition
Association	Association is a relationship between two variables. It does not state whether or not they are linked or if there is a causal relationship. Also known as correlational relationships.
Causation	The indication that one event is the result of the occurrence of the other event. Cause and effect.
Confounder	A variable that predicts the outcome of interest and also predicts whether an individual receives one or the other interventions of interest. A variable must meet two conditions to be a confounder: it must be correlated with the independent variable. This may be a causal relationship, but it does not have to be. It must be causally related to the dependent variable. An example would be pollution being causally related to cancer however the intervention was focused on smoking. It wouldn't be possible to determine whether cancer was the result of the smoking intervention or pollution that affected intervention participants.
Control group	The population who did not receive the intervention but are used for comparison benchmarking analysis. Efforts should be made to ensure the control group are representative of the intervention group in size and demographics.
Intervention group	The population who received the intervention (also known as case group, experimental group, treatment group).
Statistical process control	Statistical process control (SPC) is an analytical technique that plots data over time. It helps us understand variation and in so doing guides us to take the most appropriate action. SPC is a good technique to use when implementing change as it enables you to understand whether changes you are making are resulting in improvement.
Regression to the mean	A statistical phenomenon describing how variables much higher or lower than the mean are often much closer to the mean when measured subsequent times. Regression to the mean is due to natural variation or chance. It is essential to account for regression to the mean when evaluating care utilisation of High Intensity Users following an intervention.

References and resources

Why Evaluate

http://www.parkerduignan.com/ot/untitled_text_page.html
https://www.betterevaluation.org/en/evaluation-options/outcomes_chain
NHS England
bitesize-guide-writing-an-effective-questionnaire.pdf (england.nhs.uk)
Guide to evaluation design, principles and practice – Midlands Decision Support Unit
Guide to Questionnaires for Service Evaluation - Evaluation Works (nhsevaluationtoolkit.net)
Using Logic Models in Evaluation- Jul16.pdf (strategyunitwm.nhs.uk)

Measures

Personal Well-being ONS4 measures - Evaluating wellbeing (whatworkswellbeing.org)
Grimmett el al, 2020; "Exploring maintenance of physical activity behaviour change among people living with and beyond gastrointestinal cancer: a cross-sectional qualitative study and typology," ***BMJ Open Vol.10 (10): e037136***
Mete et al, 2019; "What is healthy eating? A qualitative exploration," **Public Health Nutrition Vol. 22 (13): 2408-2418**
Combine qualitative and quantitative data - Rainbow Framework (betterevaluation.org) NHS England: A bite size guide to running focus groups for patient and public engagement (england.nhs.uk)
Office for Health Improvement and Disparities: Focus group study: qualitative studies - GOV.UK (www.gov.uk)
NHS England: bitesize-guide-qualitative-research.pdf (england.nhs.uk)
NHS England: nhsi-clp-1discover-t6-2-analysis-1609020.pdf (england.nhs.uk)
Survey Example
NHSE - A model for measuring quality care
Nuffield Trust - Trusted measures

Things to think about

www.hra-decisiontools.org.uk

Design

https://www.gov.uk/guidance/randomised-controlled-trial-comparative-studies
Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ (December 1995). "Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group". JAMA. 274 (22): 1800–1804.
Frontiers | Levels of Evidence, Quality Assessment, and Risk of Bias: Evaluating the Internal Validity of Primary Research (frontiersin.org)
Quantifying and addressing the prevalence and bias of study designs in the environmental and social sciences | Nature Communications
The Strategy Unit - Guide to evaluation design, principles and practice
Pick 'n' mix: An Introduction to Choosing and Using Indicators

Analysis

Microsoft Word - 10_Sampling and sample size calculation 2009 revised NJF_WB.doc (bdct.nhs.uk)
Value for money | Better Evaluation
NHS England – How to understand and measure impact
Paired Samples T-test in R - Easy Guides - Wiki – STHDA
11: Introduction to Repeated Measures - Statistics LibreTexts
Guidance on undertake multivariable regression can be found here - Introduction to Multivariate Regression Analysis (https://www.mygreatlearning.com/blog/introduction-to-multivariate-regression/)
https://www.statology.org/how-to-load-the-analysis-toolpak-in-excel/
How to Conduct a Paired Samples t-Test in Excel - Statology
11: Introduction to Repeated Measures - Statistics LibreTexts

General Evaluation Resources

Impact and evaluation | NCVO
Treasury Magenta Book
Analyse data - Rainbow Framework (betterevaluation.org)
WISE: Web Interface for Statistics Education | Better Evaluation
Rainbow Framework - Rainbow Framework (betterevaluation.org)
Home - Evaluation Works (nhsevaluationtoolkit.net)

Outcomes progress ladder

Another method of measuring change is to monitor progress towards an overall outcome on a ladder.

To use this type of outcome measurement, you will need to have assigned progress descriptions at each point of the scale.

See the example below from South West London with progress on a scale towards the overall outcome of a child's general health. There are descriptions for progress on a 10-point scale towards the overall goal of healthy children being knowledgeable about their health care needs.

Whilst deciding on outcome measures and progress descriptors, you should also plan for how measurement should be taken and how frequently (e.g., once a month).

		Indicators and Levels of Need on Impact Tool scale of progress									
		1	2	3	4	5	6	7	8	9	10
Domain	**Dimension**										
Child's Development: Health	**General Health (3-19 years old)**	Chronic/serious condition, untreated. Not registered with GP	Being treated but not effectively/unknown cause. Not following care plan. Not taking meds	Missed immunisations; or missed appointments. Contradictory/unsubstansiated illness possbily facricated.	Unhealthy dietary regime; Poor dental hygiene or care. No exercise, significantly over or under weight	Family/child/young person engaged with relevant health professionnals and health needs being assessed	Receiving appropriate health advice and support and care plan developed	Care plan is being followed by child/young person and family with the support of relevant professionals. Under regular medical review	Child's health is improving, child/family have strategies to maintain good health and child's ability to participate in age appropriate activities maximised.	Child's health much improved. No longer requires regular medical review	Child is healthy and knowledgeable about health care needs and acts accordingly. No health needs

London Borough of Sutton – Children's Services, 2016

Contact nadine.wyatt@swlondon.nhs.uk for more details

Workshop	Content of the Session	Learning outcomes. By the end of this session attendees will…	Relevant Evaluation Toolkit section
Workshop 1: The Evaluation Toolkit and introduction to the programme of support	• Finding your way around the Evaluation Toolkit • Demonstrating how the Evaluation Toolkit can support practical evaluation of your health inequalities programmes • How to access and use the Evaluation Toolkit in your day-to-day work • An outline of the local programme of support available	• Be aware of the Evaluation Toolkit and how it can support their local programmes of evaluation • Be aware of the wider programme of support	**Start with Why** Why are you evaluating? What do you hope to show?
Workshop 2: From data, to measures and measurement	• Understanding how to conceptualise outputs and outcomes for your programmes of work • Getting familiar with using a measurement checklist • Understanding different types of measures and how to express them • Learning how to assess measurability (mapping measures to available data) • Balancing value with measurability • Identifying the need to collect new data • A focus on data and approaches to measuring health inequalities	• Have reviewed some real examples from the logic model workshops and expressions of interest • Have explored the best ways to practically measure outcomes for their local programmes • Understand better how to define metrics that measure health inequalities	**Measures** What data do you have and what do you need?
Workshop 3: Data sources for evaluation of health inequalities programmes (BI team attendance critical)	• Local data and intelligence tools and their application in evaluation • An introduction to other health and care data sources (including health and wellbeing status, quality, patient experience, utilisation, cost and organisational process measures) • A focus on data sources across partnerships and for the wider determinants of health	• Increase awareness of the types of data useful for local evaluation • Hear information about the local availability of different types of data and data sources • Get familiar with local resources and teams to support data collection, reports and analysis that can support evaluation of programmes	**Measures** What data do you have and what do you need?

Workshop	Content of the Session	Learning outcomes. By the end of this session attendees will…	Relevant Evaluation Toolkit section
Workshop 4: Qualitative study methods	• Introduction to qualitative data and analysis methods for evaluation • Explore community-based participatory research and its application in participatory evaluation • Understand how and when to apply these methods for intervention evaluation • Gain an understanding of how qualitative methods complement quantitative methods in a mixed approach	• Have a better understanding of how to collect qualitative data through engagement with communities • Improve their knowledge about analysing qualitative data to evaluate programmes of work • Understand more about how to blend quantitative and qualitative methods in practical programme evaluation	**Design** Best possible design of your evaluation
Workshop 5: Quantitative study methods	• Introduce quantitative data and analysis methods for intervention evaluation • Explore the routes from evaluation questions to hypothesis generation and testing • Surveys, questionnaires, longitudinal data series and other formats • Baselining, tracking, monitoring and following up • Statistical analysis to disentangle causality from association	• A better understanding of how to collect quantitative data where necessary • Knowledge of how to analyse quantitative data in programme evaluation • Knowledge of where to go for secondary data analysis to support evaluation	**Design** Best possible design of your evaluation
Workshop 6: Evaluation bias	• An introduction to different forms of bias • Outlining the links between evaluation design and bias • Approaches to minimise bias in study design • Ways to build in comparison and control groups	• Have gained knowledge to support better design of programme evaluation • Gained the ability to identify and mitigate potential biases in programme evaluation	**Analysis** Bringing together your data
Workshop 7: Value for money	• How is value understood and measured? • Methods for quantifying value for money from intervention investment • Identifying cost-saving opportunities • How to prove that your intervention has delivered value for money • Social return on investment methodology – reputable sources and use of financial proxies for outcomes	• Understand the construct and measurement of value and how to demonstrate a return on investment through evaluation	**Drawing conclusions** So what?