



Evaluating the impacts of your volunteering service

A guide for managers of hospital volunteering services

May 2021

Introduction: evaluating your volunteer service

Why evaluate?

Enlightenment: Robust evaluation can tell us not only **whether a service worked, but also why and how**. Evaluation can tell you the merit, worth or value of something (Scriven, 1991)

Improvement; Evaluations can help you **improve** your service, learning from what is going well and could be better.

Persuasion: Evaluations can help give you the evidence needed to **maintain** or **grow** the service and investment in it. Your evaluation findings can also inform decision making in your wider organisation. Working with decision-makers from the outset, you can ensure your evaluation timeline aligns with their decision-making cycles. This will increase the likelihood of your findings being applied to make change.

Engagement: Evaluations can also be an opportunity for patient and public representatives' perspectives to be heard and taken into account to change services.

Who is this guide for?

This guide is for managers of volunteering services in hospitals who want to know if their service is working as intended and to identify ways in which it can be improved.

How to get started?

This evaluation guide can help you get started, whether you are planning a new evaluation or adapting an existing approach to investigate the impacts of your service.

Our aim is to give you the knowledge and support to:

- 1) describe how and why your service works
- 2) plan an evaluation
- 3) collect and analyse data
- 4) reflect and report on what your service is achieving.

Introduction: should I monitor or evaluate my service?

First of all, there is a lot of diversity of terminology in monitoring and evaluation. Some people refer to monitoring and evaluation as simply ‘programme evaluation’, but it can be helpful to distinguish them.

Monitoring

Continuous supervision of an activity to check whether plans and procedures are being followed (Ovretveit, 2014)

- Tends to be an on-going collection of information that begins when you start your service and ends when your service pauses / closes
- The data collected is used primarily for programme management

vs.

Evaluation

A comparative assessment of the value of something, using systematically collected and analysed data, in order to decide how to act (Ovretveit, 2014)

- Encourages you to set a goal for your service (e.g. improve staff satisfaction by 10%) and compare your service against that goal (or other similar services) to measure its impact
- Needs at least two data points for comparison (e.g. looking at what the volunteer did or what the service provided now compared to last month, or what people say now about the service compared another service at this time).
- Involves using the data you collect to make a judgement of the service’s merit or worth (and is therefore more analytical than monitoring)
- Is done less frequently than monitoring



Image source:
<https://www.measureevaluation.org/resources/publications/ms-07-20-en>

Introduction:
monitoring vs.
evaluation 2

Background:
Volunteering
Innovators
Programme 4

Before you
begin: issues
to consider 7

Four steps of
evaluation 8

Pit stop:
questions to
ask yourself 9

How to use
this guide 10

Background: how this guide was developed

This guide was informed by an activity that examined which characteristics made the 12 Trusts involved in Helpforce's Volunteer Innovators Programme (VIP) good candidates for an 'impact' or 'outcomes' evaluation.*

The volunteer services were diverse, spanning for example discharge from hospital, nutrition and hydration, and peer debriefing in mental health settings. Each service sought to improve outcomes across four categories: patients, volunteers, staff and the wider system. See the example below and on the following two slides.

VIP service description

- Volunteers were trained to support patients to stay mobile during their hospital stay
- Volunteers provided basic physiotherapy to patients who were waiting to return home



Expected outcomes



Patients: Improve patients' experience of care and emotional wellbeing, and an improved confidence around their mobility



Volunteers: Increase volunteer employability in healthcare, and improve their wellbeing and experience of volunteering



Staff: Save staff time and increase their confidence in the quality of care they provide, improve their perception of volunteers



The system: Faster discharge from hospital, reduce (re)admissions to hospital, or reduce complaints received by the hospital

Before we get into the detail, we should be clear that there are numerous ways of evaluating a volunteering service. This guide focuses on conducting a **robust outcomes evaluation** to measure how much of a difference your service is making for patients, staff, volunteers and the system in the short and longer term.

* We used an adapted version of Davies (2013) evaluability assessment, see: Davies, R., 2013. *Planning Evaluability Assessments: A Synthesis of the Literature with Recommendations*. Report of a Study Commissioned by the Department for International Development. Available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/248656/wp40-planning-eval-assessments.pdf

Introduction:
monitoring vs.
evaluation 2

Background:
Volunteering
Innovators
Programme 4

Before you
begin: issues
to consider 7

Four steps of
evaluation 8

Pit stop:
questions to
ask yourself 9

How to use
this guide 10

Background: interventions included in Helpforce's VIP

VIP: NHS England/NHS Improvement provided funding to 10 NHS Hospital Trusts (a further two received funding from the Royal Voluntary Service) to develop, test, measure, and spread a range of high-impact volunteering innovations. Some innovations were new, while others built on existing volunteering initiatives. The innovations are listed below:

- Active Response Volunteers who help staff with a wide variety of tasks including the collection of patient medication, to help save staff time and potentially speed up discharges. (Barts Health NHS Trust)
- Volunteers providing peer to peer emotional support to patients after being restrained in a mental health setting, to help make care more personalised. (Camden and Islington NHS Foundation Trust)
- Volunteers providing support with end of life care in the community, to enable more people to die at home if they choose. (Dorset HealthCare University NHS Foundation Trust)
- Volunteers providing support to patients before, during and after operations, to help reduce anxiety and improve the patient experience. (Moorfields Eye Hospital NHS Foundation Trust)
- Settle in Service with volunteer transport, to improve the patient experience at discharge and potentially reduce the risk of readmissions. (Norfolk and Norwich University Hospitals NHS Foundation Trust)
- Volunteer support with patient discharge, as well as 'Hospital to Home' and transport services, to improve the patient experience at discharge and potentially improve patient flow and reduce the risk of readmissions. (North Tees and Hartlepool Hospitals NHS Foundation Trust)
- Volunteer support with end of life care in hospital, to help staff deliver person-centred care. (Royal Liverpool and Broadgreen University Hospitals NHS Trust)
- Volunteer assistance for patients at mealtimes, to reduce the risk of malnutrition and dehydration. (Salford Royal NHS Foundation Trust)
- Volunteers providing patient engagement and support using a specially developed model for companionship, to reduce anxiety and improve mood. (Salisbury NHS Foundation Trust)
- Volunteers providing motivation, encouragement and support, in moving and getting dressed into day clothes, for patients waiting to return home, to reduce the risk of deconditioning whilst in hospital. (Sandwell and West Birmingham Hospitals NHS Trust)
- Mobility, nutrition and hydration volunteers, to reduce the risk of deconditioning, malnutrition and dehydration whilst in hospital. (St James's University Hospital in Leeds - funded by RVS)
- Mobility volunteers provided on-ward resistance-based training to reduce loneliness and improve physical health. (University Hospitals of Morecambe Bay NHS Foundation Trust – funded by RVS)

Background: examples of two VIP sites evaluation plans

Introduction:
monitoring vs.
evaluation 2

Background:
Volunteering
Innovators
Programme 4

Before you
begin: issues
to consider 7

Four steps of
evaluation 8

Pit stop:
questions to
ask yourself 9

How to use
this guide 10

	Site A	Site B
VIP service description	Activity Support and Mobility Volunteers in three wards (the Acute Medical Unit, Older Persons Assessment Unit, and Discharge Unit) to prevent patient mobility from deteriorating after long stays in hospital.	Debrief Volunteers to support patients who had been restrained to reduce the number of times patients were restrained more than once, and improve patient experience.
Purpose of evaluation and audience	To demonstrate to the Trust's board the benefits of the service for patients and the hospital by: increasing number of discharges to usual residence; reducing numbers of falls and pressure ulcers in patients; reducing length of stay; and reducing physiotherapist referrals.	To demonstrate to the Trust's board the benefits for patients and hospital and to embed debriefing into usual practice by: improving patient experience of care, increasing uptake of debriefs and reducing repeat restraints, and improving staff and volunteer wellbeing.
Size of project	Nearly 60 volunteers across three wards by January 2020.	This was undertaken on a very small scale: 11 volunteers across 10 acute wards.
Working group	Supported by a working group and by senior officials in the Trust (e.g. Director of Communications). The group benefitted from clinical input, for example from physiotherapists and also involved representatives from Research and Development.	Supported by a working group that met monthly and formed part of a broader Trust-wide initiative to improve responses to violence. The project also had established networks with partners such as Voiceability and Mind.
Data collection and analysis	Data included, but was not limited to: number of referrals to physiotherapy team, ED readmissions, staff stressed, staff time saved, patient interactions per month (walked/ exercised, already dressed, and changed into day clothes). The team looked at trends over time and the relationships between outcomes and volunteer activities.	Number of care plans incorporating patient-centred care, patient satisfaction (own survey), number of patient repeat restraints, debrief take-up rate. For debrief take-up rate, comparison wards were used. In the wards with volunteers, 37% of all incidents were followed by a debrief whereas in the control wards with no volunteers only 14% of incidents were followed by a debrief (June-Dec 19).
Service achievements and outcomes	580 patient interactions that involved walking or exercise and around 175 patients were encouraged to put on clothes (April 2019-March 2020), and number of rereferrals to therapy from the discharge ward decreased. 5 of 7 physiotherapists who responded to an early survey reported volunteers saved them time (5-60 minutes in an average day). Contributed to wider evidence around the 'End P.J. Paralysis' campaign. The volunteering team aim to extend the service to additional wards.	More patients had a debrief following a restraint and patients, volunteers and staff reported an improved experience of care and care delivery. Contributes to achieving NICE Quality Standard (QS154). The volunteering team have continued the expansion of the role to more wards across the Trust.

Introduction:
monitoring vs.
evaluation 2

Background:
Volunteering
Innovators
Programme 4

Before you
begin: issues
to consider 7

Four steps of
evaluation 8

Pit stop:
questions to
ask yourself 9

How to use
this guide 10

Before you begin: four evaluation issues to consider

1 Purpose

What is the purpose of the evaluation?

For example, will it be used to prove the value of your service to Trust executives?

Or will it be used to improve your service?

2 Audience

Who are the key audiences of your evaluation?

What level of evidence do they need?

Which outcomes and/or impacts are they interested in?

3 Resources

What resources (time, money, expertise) will be needed and are these resources available?

Do you have a working / steering group?

4 Goals

What are the goals for the service?

What would you like the service to achieve for patients, staff, volunteers, the hospital?

Note: The evaluation work should be underpinned by a working group made of up key stakeholders (e.g. patients who use or could potentially use the service, a mix of frontline staff, service managers, and where possible, data analysts, transformation / quality improvement team members, data analysts, and Trust senior leaders / executives). The group should develop processes/tools to help capture and store data in an appropriate way.

This guide follows four steps of evaluation

Introduction:
monitoring vs.
evaluation 2

Background:
Volunteering
Innovators
Programme 4

Before you
begin: issues
to consider 7

Four steps of
evaluation 8

Pit stop:
questions to
ask yourself 9

How to use
this guide 10

Reflect on results and report your service achievements

Regularly report your findings and consolidate all of your learning into a summary with recommendations

Issues to consider:

- On-going reporting (e.g. PDSA) vs summative reporting

Collect and analyse data

Collect stories and numbers. Compare and combine these into evidence. Share emerging findings.

Issues to consider:

- Baselines, comparison groups
- Data collection and analysis

Describe how and why your service works

Write down what will change as a result of your service, how and why (in a theory of change)

Issues to consider:

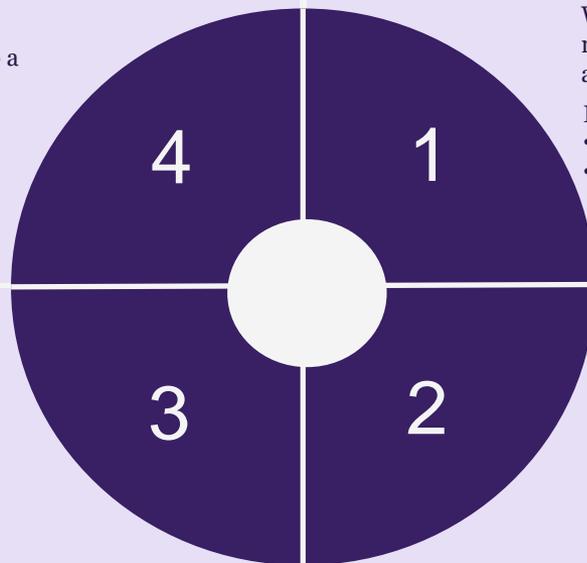
- Relevance
- Clarity and complexity

Plan the evaluation with your working group

Build a working group. Think carefully about why you are doing an evaluation – which decisions are you trying to influence?

Issues to consider:

- Stakeholder agreement
- Plausibility and sustainability
- Link between indicators and outcomes
- Context
- Ethics and governance



In this guide you will be taken through four steps covering a total of 10 points to consider and checklists that will encourage you to ask yourself questions to ensure you've understood the concept. We also compare 'good practice' and 'practice that could be improved' and give you tips to overcome common challenges.

Introduction:
monitoring vs.
evaluation 2

Background:
Volunteering
Innovators
Programme 4

Before you
begin: issues
to consider 7

Four steps of
evaluation 8

Pit stop:
questions to
ask yourself 9

How to use
this guide 10

Questions to help prompt your evaluation planning

Relevance

- What is the problem your service is trying to solve?
- Is your service relevant to the problem identified?

Step 1

Clarity & complexity

- Is the service clearly described?
- Are you clear on the service's goals?
- Have you identified short and long-term outcomes?
- Does the target population have boundaries?

Step 1

Stakeholder agreement

- Does your working group have wide representation?
- To what extent does the working group agree on evaluation plans?

Step 2

Plausibility & sustainability

- Is there a clear timeline?
- Is it likely that the project objectives could be achieved in that timeline?
- What are the service's plans for sustainability?

Step 2

Indicators & outcomes

- Could the indicators plausibly reflect the expected outcomes?
- Are the expected outcomes realistic?

Step 2

Context

For the service and its evaluation:

- Have internal / external enablers and barriers been identified?
- How will you monitor these?

Step 2

Ethics & governance

- What approvals will you need for your evaluation?

Step 2

Baselines & comparisons

- What baseline data will you use?
- Who will form your comparison group?

Step 3

Data collection & Analysis

- How consistent will data collection be? How frequent?
- Will the sample be big enough for quantitative analysis – or would qualitative approaches be better?

Step 3

Reporting

- Will you report findings throughout (for service change) or at the end?
- In what ways can you share your findings with varied audiences?

Step 4

How to use this guide

This guide covers ten issues to consider when doing an evaluation. Each issue is described on a single slide or a few slides – depending on the amount of information needed to explain the issue.

In this section of the slide, the issue is introduced and examples are provided.

Sometimes the explanation is brief and summarised in a paragraph or two (e.g. 'Is your service relevant to the problem identified?'). Other times, the explanation is more complex and more detail and diagrams are needed and spans a few slides (e.g. 'Is your service logic clear?'). Each issue compares good practice vs. practice that could be improved, as well as top tips and a check-in question before you move on.

Examples of good practice

- In this section of the slide, we provide examples of good practice based on what was visible in the VIP programme and also in other NHS organisations that have carried out local evaluation. This list is not exhaustive but gives you a good sense of what to aim for.

Practice that could be improved

- In this section of the slide, we provide examples of poor practice. The list suggests what you should avoid doing.

Questions to ask yourself



- In this section of the slide we summarise questions that you should ask yourself to see if you have fully considered and addressed the issue in your evaluation.

Top tips



- In this section of the slide we provide examples of challenges experienced by the VIP sites and other NHS organisations we have evaluated, as well as top tips to overcome these challenges.

Introduction:
monitoring vs.
evaluation 2

Background:
Volunteering
Innovators
Programme 4

Before you
begin: issues
to consider 7

Four steps of
evaluation 8

Pit stop:
questions to
ask yourself 9

How to use
this guide 10

The four steps of evaluation

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
	Ethics & Governance	25
STEP 3	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

Step 1: Describe how and why your programme works

After this step, you should be able to answer:

- What is the problem your service is trying to solve?
- Is your service relevant to the problem identified?
- Is the service clearly described?
- Are you clear on the service's goals?
- Have you identified short and long-term outcomes?
- Does the target population have boundaries?



STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
STEP 3	Ethics & Governance	25
	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

Is your service relevant to the problem identified? (1/1)

Where possible, it is advised that you gather proof that the people who you designed your service for (your ‘target population’) – either staff or patients or both – *actually* need your proposed service. This proves your service is **relevant** to its stakeholders. This can be demonstrated through, for example, existing studies and literature, national or local policy commitments, or patient- or clinician-led campaigns.

It is also important to clarify who you think will benefit from your service.

For example, the Trusts that were involved in VIP were clear that there was a problem that needed solving by leveraging volunteer support. At one VIP site, new NICE guidelines meant that debriefs following a mental health episode would need to be introduced to continue delivering an excellent service. At another VIP site, mealtimes were proving difficult for patients who did not have any family or carer support. A pilot ‘Dining Companions’ programme was tested and proved effective, which provided evidence for a wider roll out.

Examples of good practice	Practice that could be improved
<ul style="list-style-type: none"> Gaining access to or gathering your own supporting evidence of the need for the service (e.g. previous services’ reports, surveys that demonstrate need, etc.) and capturing lessons learnt. Writing down your service’s rationale (i.e. the problem you are trying to solve), and adding it to your theories of change. This will help put your service in context. 	<ul style="list-style-type: none"> Having no or little evidence that describes the need for your service. Not having a clear problem or combination of problems.

Questions to ask yourself

- What is the problem that the service is trying to address?
- Has specific work / analysis been done to determine the need for the service?

Top tips

- Speak to your stakeholders, particularly seeking clinical input/expertise, possibly using short interviews or even focus groups, to gain their views on who might benefit from the service – however, informal discussions can be helpful when interviews or focus groups aren’t feasible
- Conducting a scoping search can help to identify any previous studies or literature that might be relevant.

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
	Ethics & Governance	25
STEP 3	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

Is your service logic clear? (1/5)

The first step in an evaluation is describing your volunteering service and justifying how and why it works.

Inputs

the service's resources (those available and those missing), which can include units of time, staff, money, equipment, know-how, ideas, etc.

For example: volunteers are trained to identify and support patients with mealtime support needs

Activities

what volunteers need to do with/for the target patient group to ensure that the service is successful

For example: volunteers assist patients unable to eat and drink independently during all mealtimes on a daily basis)

Outputs

what will result from volunteers carrying out the described activities

For example: number of volunteers trained as a proportion of all volunteers, number of patients assisted as proportion of all patients in need in a given week, number of mealtimes where assistance was provided as a proportion of all mealtimes in a given month

Outcomes

the changes, benefits, learning or other effects that happen as a result of services

For example: decrease in the number of patients who were at risk of dehydration and malnutrition

Create lists in each of the boxes above and once your lists are ready, start linking items: 'If [input A + activity A], then [outputs A + outcomes A]' will result. Continue with input B, C, etc. Sometimes it is easier to work backward from your outcomes.

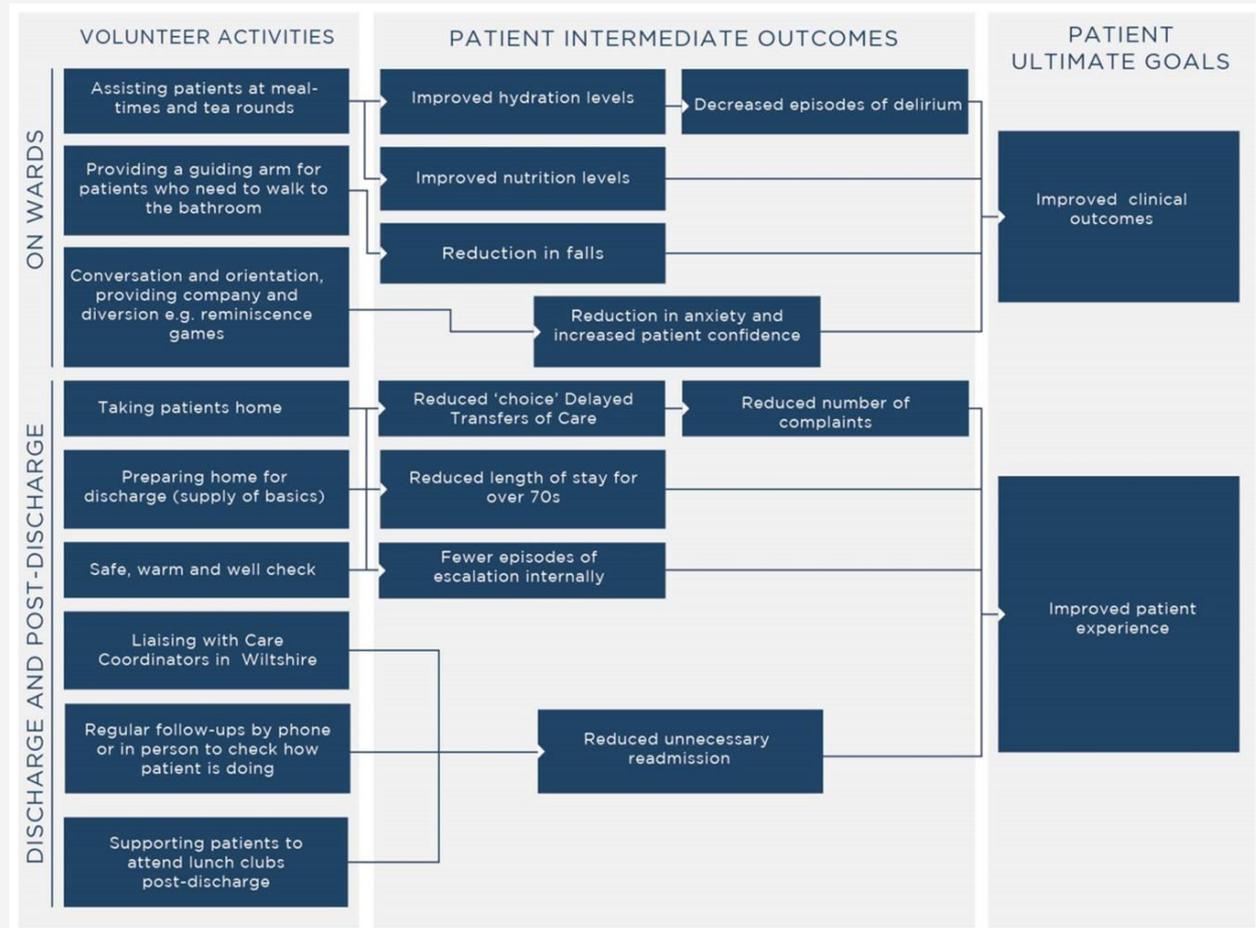
These 'if, then' statements outline your assumptions about how and why your service works. Each statement is called a 'theory of change'.

You will likely have multiple theories of change for a single service because you want to have an impact on more than just one group of people (e.g. patients AND staff), or because your service has multiple components or steps.

Writing out your theories of change clearly ensures that everyone involved understands what exactly is being evaluated and why you think it works. See [an example](#) of the questions our team used to prompt a VIP service to develop their theories of change in [slide 17](#).

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
STEP 3	Ethics & Governance	25
	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

Is your service logic clear? (2/5)



Drawing your theory of change

There are numerous ways to draw your theory of change.

This example focuses on the connections between the activities and outcomes.

Find additional examples in this [Gov.uk guide](https://www.gov.uk/guidance/drawing-your-theory-of-change).

Source: <https://www.nesta.org.uk/toolkit/helping-hospitals/>

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
STEP 3	Ethics & Governance	25
	Baselines & Comparisons	27
STEP 4	Data collection & Analysis	29
	Reporting	35

Is your service logic clear? (3/5)

In order to create a clear theory of change it is important to be clear about your ‘target population(s)’.

Your target populations should have clear boundaries so you can distinguish who should be included in your evaluation and who should be excluded.

Boundaries could be based on certain criteria such as age, gender, health condition, job role, etc. You should be able to justify your rationale for inclusion and exclusion.

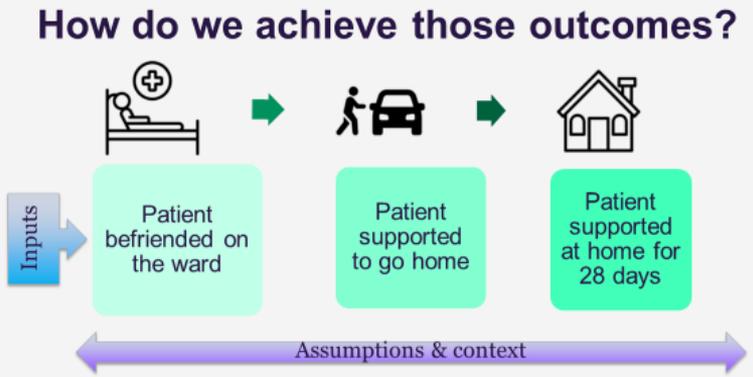
Examples of inclusion/exclusion criteria from VIP services	
Mobility intervention	Included: Long-stay patients (any age), frail older adults at risk of falls
	Excluded: Patients with significant mobility problems were not approached by the Activity, Support and Mobility Volunteers as they required support beyond the volunteers’ remit
Hand-holding intervention	Included: In a hand-holding intervention in an eye hospital, patients who could receive an operation under local rather than general anaesthesia (e.g. glaucoma or cataract patients)
	Excluded: flexible, but typically patients not requiring day surgery, and those who refused the service
Restraint debrief intervention	Included: All mental health patients who experienced the use of restraints over the intervention period and had refused a debrief by a healthcare professional, or did not have ability to consent
	Excluded: Patients who refused the service

Being clear about your inclusion/exclusion criteria will help you choose your data sources and indicators later.

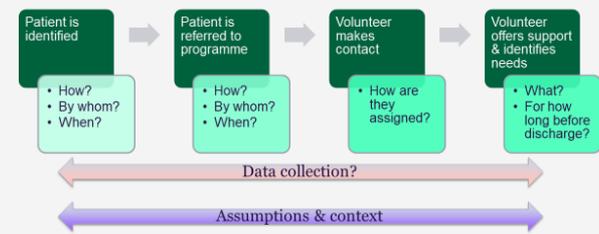
However, it is important to note that if the group under investigation is too narrow in definition the size of the population becomes too small for a robust evaluation.

STEP 1	Relevance	13
	Clarity & Complexity	14
	Working group agreement	20
	Plausibility & Sustainability	22
STEP 2	Indicators & Outcomes	23
	Context	24
	Ethics & Governance	25
STEP 3	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

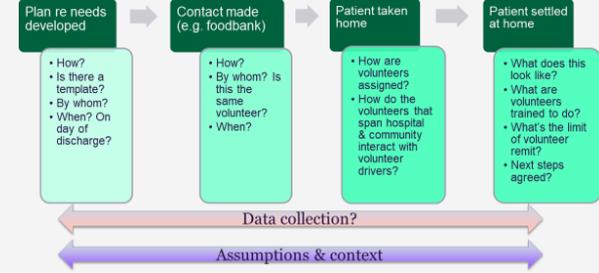
Is your service logic clear? (4/5)



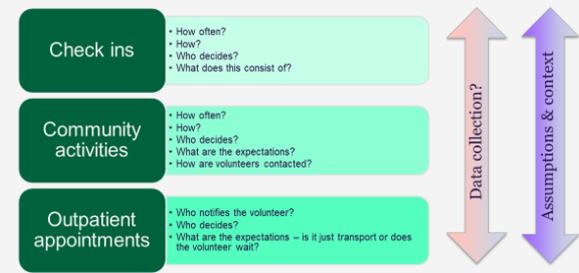
1. Befriending on the ward



2. Support to go home



3. Support at home



STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
	Ethics & Governance	25
STEP 3	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

Is your service logic clear? (5/5)

Examples of good practice Practice that could be improved

- | | |
|---|--|
| <ul style="list-style-type: none"> • Involving your (sufficiently representative) working group in developing your theories of change. • Drawing on local knowledge, clinical expertise, experience from similar services or academic theories to build the evidence behind your theories of change. • Developing multiple theories of change (flow diagrams) for each part of your service. • Creating a longer written description of your theories of change. This should involve, for example, descriptions of the role of the volunteers, staff, etc. carrying out activities, and descriptions of the local context and time. • Having clear boundaries around the groups of people who you hope will benefit from your service (i.e. clear criteria for the target group) | <ul style="list-style-type: none"> • Working alone or not consulting your service's stakeholders and available evidence when developing your theories of change. • Being superficial: having only one theory of change when your service is more complex (and would benefit from many). • Not being sufficiently detailed: loosely listing all inputs in a box, then all actions in a separate box, etc. without making the links between the specific components in each box. This creates 'black boxes' that make it difficult to unpick the cause and effect between components of your service. • Overlooking lessons learnt: When you don't consider evidence from other similar services, or the lessons learnt from the predecessors to your service, it can create doubt that your service is going to work. |
|---|--|

Questions to ask yourself



- Is the service clearly described?
- Are you clear about what the service aims to achieve?
- Are the short and long-term outcomes clearly identified?
- Does your target population have clear boundaries?
- If your service has multiple components, can each part be distinguished into its own 'chain' of inputs, activities, outputs and outcomes?

Top tips



- You may need to rewrite / redraw your theories of change numerous times before you get it right. This is frustrating but completely normal.
- Consult your working group to help and include a group of people who will be part of the service (e.g. patients, staff, volunteers).
- Keep your theories of change updated whenever your service changes. This will enable you to share it at a moment's notice if needed for presentations, meetings or business cases.

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
	Ethics & Governance	25
STEP 3	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

Step 2: Describe how and why your programme works

After this step, you should be able to answer:

- Does your working group have wide representation?
- To what extent does the working group agree on evaluation plans?
- Is there a clear timeline?
- Is it likely that the project objectives could be achieved in that timeline?
- What are the service's plans for sustainability?
- Could the indicators plausibly reflect the expected outcomes?
- Are the expected outcomes realistic?
- Have internal / external enablers and barriers been identified?
- How will you monitor these?
- What approvals will you need for your evaluation?



Do you have agreement within the working group? (1/2)

To give you the support you will need during the evaluation it can be a good idea to invite the evaluation's stakeholders to join a working group. Stakeholders include anyone who could be affected by the evaluation (i.e. anyone with a 'stake' in the evaluation). This will include the patients, staff, volunteers involved in your service and local decision-makers who can help support your service within and outside of the Trust.

The evaluation working group can help make decisions about the evaluation, act as a sounding board for adaptations, help you gather data, and ensure that action results from the evaluation.

Working groups benefitted most at VIP sites when there was input from Trust senior management (e.g. Director of Communications), or input from other Trust working groups (e.g. Salford working group on hydration and nutrition).

(1) The service

What are the lessons learnt from previous projects? How can you build on evidence? Are there plans for service sustainability in place beyond any initial trials/pilots?

(2) Your 'team'

Does everyone in the working group have a role and know what they're doing? Is there anyone missing? Would the working group benefit from evaluation training / coaching? Where can they access support?

(3) Evaluation queries

What are the questions that are most important to answer? What data collection approaches are most suitable? What and who do you want your evaluation findings to influence?

(4) Evaluation planning

What would success look like for your evaluation? What resources and approvals will the evaluation need?

(5) Engagement & communications

How will you communicate with the working group and share emerging findings throughout the evaluation with other stakeholders?

- At your first few meetings, consider covering the 'big' topics in the box on the left
- See the Working Group appendix (Appendix B) for a full list of questions to cover
- Write down what is agreed in a monitoring and evaluation plan (that sets out the evaluation plans) and attach plans to a timeline

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
	Ethics & Governance	25
STEP 3	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
	Ethics & Governance	25
STEP 3	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

Do you have agreement within the working group? (2/2)

Examples of good practice

- Having a team of representatives from each stakeholder group (patients, staff, volunteers, the hospital). Ideally patients and volunteers would have varying demographics, and hospital staff and local decision makers would have varying levels of seniority.
- Assigning activities or roles to team members.
- Holding regular meetings covering each of the topics in the Working Group appendix.
- Writing down what is agreed on each topic in the Working Group appendix, and bringing it together in some form of evaluation document that would ideally involve a project planning timeline.

Practice that could be improved

- Working alone or not consulting your service's stakeholders.
- Meeting and communicating with your working group irregularly, making it difficult for them to keep up to date and give you informed advice.
- Keeping only verbal agreement of plans – leading room for potential disputes or miscommunications

Questions to ask yourself



- Does your working group have wide representation?
- Who are your other evaluation stakeholders (not on the working group)?
- What kind of evidence and evaluation plans will be required? To what extent does the working group agree?

Top tips



- You may have trouble getting a diverse group of stakeholders to form a 'team'.
- You may also experience difficulty in assigning them roles – especially if it involves helping with data collection.
- These challenges could be overcome by seeking out membership from people who have good access to data you might need and / or have the analysis skills to help (e.g. Quality Improvement teams).
- It would also help if you could ensure that the members of your working group are interested and clear on the time commitments required.
- Keeping a record of key decisions in an evaluation plan is helpful for the whole team, and as evaluation plans progress one person will likely need to regularly update this document and add detail as it develops.

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
STEP 3	Ethics & Governance	25
	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

Is your service logic plausible and sustainable? (1/1)

Where there are particular goals or outcomes that you would like the service to achieve, it is important to carefully consider whether they are achievable within your evaluation timeline (and the lifetime of the service too).

Drawing on lessons from previous programmes can be helpful: is there evidence from elsewhere that the service could meet its goals? And sustain achievement?

Examples of good practice	Practice that could be improved
---------------------------	---------------------------------

- | | |
|--|---|
| <ul style="list-style-type: none"> • Having very clear long-term objectives that are measurable and time-bound (e.g. to reduce admission rates for over-65s from 261 to 245 per 1,000 by December 2021) and can be attributable as far as possible solely to the intervention. • Clearly describing any pre-existing services in the context of the problem, new service, resources, or timeline – thus making a clear link between pre-existing and new services. • Keeping decision makers up to date with regular findings can help secure Trust buy-in. • Agreeing upfront with decision makers whether the evaluation will have an influence on a service's sustainability. | <ul style="list-style-type: none"> • A timeline that does not reflect reality: Gaining approvals, enrolling patients to participate in the service and evaluation, analysing your evaluation data are all time-consuming tasks. • Outcomes focused in a single time period (long or short term) are problematic because if too short it is difficult to know if they can be sustained and if too far away the service or evaluation may end before data can be collected to demonstrate the service's impact. |
|--|---|

Questions to ask yourself

- Is there a clear timeline? Is it likely that the project objectives could be achieved in that timeline?
- Is the timing right for an evaluation?
- What are the service's plans for sustainability?
- How will the service's sustainability affect the evaluation?

Top tips

- The development of new interventions takes time, so reaching sample sizes 'large enough' for an evaluation can prove difficult, especially within a short timeframe.
- Speak with quality improvement and data teams to discuss how frequently data is collected and with what delay to help be mindful of what questions and measures will be appropriate to demonstrate the intervention's impact.
- Expect delays at every stage, e.g. identifying and accessing data in a format appropriate for the evaluation, designing new surveys.

Link between indicators and desired outcomes (1/1)

Outcomes evaluations typically focus on questions such as:

- Did the service produce or contribute to the intended outcomes?
- For whom, in what ways and in what circumstances?

Some of these questions will necessitate quantitative data (often describing 'how many' 'how often') whereas others will need qualitative data (often describing 'why' 'how' 'in what way').

Examples of good practice

- Before choosing indicators, consider spending some time learning about what data is already being collected within your hospital or what existing tools (e.g. validated questionnaires published in academic journals) might be a good match to your service's outcomes. VIP sites searched for existing data and measures within: Information collected by QI teams, PALS questionnaires, Friends and Family Questionnaires, or hospital data regularly collected, such as number of falls in a given ward.
- Choose a small number of short, medium and long term outcomes that are distinct for patients, staff, volunteers and the system, as well as relevant to the service and the questions your evaluation is trying to answer.
- Be realistic about what is achievable. If there is relevant routinely collected data then try to use it (considering how long data collection can take).
- Map out your data timepoints against your project timeline.
- If collecting data, plan how this will happen. Will equipment be needed e.g. paper surveys or a voice recorder? Will surveys be self-report or be administered? Have a plan on where to securely store your data.
- Don't over-reach: only collect the data that you need to answer your research questions.
- Use a mix of qualitative and quantitative data.

Practice that could be improved

- Outcomes that are relevant, but use a bespoke untested tool (e.g. reduce patient's anxiety, relevant measure but no tool described on how to do this).

Questions to ask yourself



- Could the indicators plausibly reflect the expected outcomes: For patients? For staff? For volunteers? For the system?
- Are the anticipated long-term impact and outcomes realistic?
- Are they reliable indicators? i.e. will observations by different observers find the same thing?

Top tips



- Where timelines are short, it can be helpful to select shorter-term outcomes that if achieved one could reasonably believe that longer-term outcomes could also be achieved.
- If collecting new data that isn't routinely collected from patients, volunteers or staff, be realistic about what is achievable - consider the impact on their time and design the process to be as simple as possible.
- When choosing indicators: (a) ask stakeholders to describe what pattern of effects would be typical for a service – and then search for appropriate indicators, or (b) look to comparative case studies.

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
STEP 3	Context	24
	Ethics & Governance	25
STEP 4	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
STEP 3	Ethics & Governance	25
	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

Have you accounted for ‘context’? (1/1)

It is important to monitor the progress of both:

- the service...to see whether it is being implemented as you have described, and
- the evaluation...to see whether you are likely to achieve the success you outlined.

It is equally important to also outline any possible ‘events’ or ‘roles of other people outside of the project’ that might positively or negatively influence your service and evaluation – these are called enablers and barriers respectively. An enabler could be commitment of funding from the Trust to carry out the evaluation, whereas a barrier could be ward staff not being receptive to the evaluation and ‘feeling watched’. Monitoring these potential influences can be done through risk registers or noting possible enablers or barriers in your narrative form of your theory of change.

Examples of good practice

- Developing a list of enablers (e.g. commitment of Trust funding) and barriers with your working group.
- Documenting the list, and writing down how you will manage any potential change.
- Failing to put appropriate plans in place to monitor the progress of the service or the evaluation.
- Working alone to develop a list of enablers and barriers, instead of the working group or a wider group of stakeholders.
- Acting rapidly (and in consultation with the working group) to change evaluation plans when they are not working.

Practice that could be improved

- Verbally monitoring the enablers and barriers or agreeing any changes, rather than documenting them.
- Failing to put appropriate plans in place to monitor the progress of the service or the evaluation.
- Working alone to develop a list of enablers and barriers, instead of the working group or a wider group of stakeholders.

Questions to ask yourself



- Have internal / external enablers and barriers been identified?
- Are there plausible plans to monitor these in any practicable way?

Top tips



- Things do not always go to plan, so keep in mind that it is okay to change the data you want to collect and /or your approach to collecting it if you experience challenges.
- However, also keep in mind that this might have implications for your analysis.
- It is not always possible to overcome barriers (e.g. COVID-19), but instead it can be useful to just monitor the impact barriers are having to recognise these as potential limitations when reporting your evaluation findings.

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
Ethics & Governance 25		
STEP 3	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

Have you considered ethics and governance approvals? (1/1)

Having your project reviewed by an ethics committees will demonstrate to your potential participants that your research is worthwhile and will have beneficial effects that outweigh the risks.

Early in the design process, engage with your local Research and Development (R&D) team about what approvals might be needed. They can advise whether approvals can be given in-house or whether you will need to speak to the Health Research Authority and/or Research Ethics Committee.

You will also want to speak with your hospital’s team responsible for developing information governance (including GDPR), as they will be able to provide advice on how you collect and store data during and after the evaluation. Gaining approvals can be a lengthy process, as committees often meet monthly and need to review your initial proposal and any changes over time.

Examples of good practice

- Contacting your R&D team for advice as soon as possible
- Including a R&D team representative in your working group.
- Informing the R&D team of changes to the evaluation plan and the implications this will have on your potential and existing participants and their data in a timely manner

Practice that could be improved

- Waiting too long: Seeking out approvals too late, delaying the evaluation timeframe, not keeping up to date with annual reports requested by committees.
- Lacking detail: Not including the contact address for the study sponsor on information sheets, in case breaches of ethical conduct need to be reported by participants.

Questions to ask yourself



- Is research ethics approval required for the planned evaluation?

Top tips



- Determining whether you need ethical approval is challenging – especially where you have little guidance to help you choose.
- Most local evaluations are classed as service evaluations because their results are not meant to be generalisable.
- Your R&D team are best placed to advise on what approvals you will need.
- The Health Research Authority (HRA) can also tell you whether you need ethical approval – and if yes, whether you need their approval and/or that of the Research Ethics Committee. The HRA also provides templates for project documents such as consent forms and information sheets.

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
	Ethics & Governance	25
STEP 3	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

Step 3: Collect and analyse data

After this step, you should be able to answer:

- What baseline data will you use?
- Who will form your comparison group?
- How consistent will data collection be? How frequent?
- Will the sample be big enough for quantitative analysis – or would qualitative approaches be better?



STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
	Ethics & Governance	25
STEP 3	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

Do you have appropriate baseline data and comparison groups? (1/2)

Baseline data

Ideally all projects should have some understanding of the conditions at the start of the service or before the service was put in place (i.e. baseline data). This could include:

- patient data before they were exposed to the service (e.g. mobility levels).
- staff data (e.g. data on stress or wellbeing).
- system measures (e.g. monthly rates for length of stay greater than 21 days or number of referrals to physiotherapists).

Baseline data acts as a fixed reference point or benchmark to compare your service against, from which change and progress can be measured. Some baseline data may be readily available (e.g. if routinely collected); however, some baseline data might need to be collected as part of the evaluation (e.g. surveys or at patient appointments). Without baseline data, it can be challenging to monitor and evaluate a service.

Comparison groups

Having a comparison group provides an understanding of what would have happened to the patients, staff, volunteers and system if the service was not in place (often called the ‘counterfactual’). There are a few straightforward ways you can calculate a counterfactual with the help of your working group:

- Logically constructing a counterfactual using the baseline as an estimate of the counterfactual (and then using process tracking at each step of the theory of change).
- Comparing outcomes with a comparison group. Some VIP sites compared historic patient data. For example, one site rolled out its service in a staggered way across wards and compared each new ward against the data from the earlier wards. Another group of VIP Trusts compared their results with trends in a group of similar patients on another local ward who had not been exposed to the service. Others compared their findings with national data from a set of similar patients.
- Comparing the before-and-after difference for the group receiving the service (and then comparing with a before-after difference with a group who did not receive the service).
- In addition, it can be helpful to ask people who know the service well to predict what would have happened in its absence.

Baseline data and comparison groups (2/2)

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
	Ethics & Governance	25
STEP 3	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

Examples of good practice

Baseline data

- Where baseline surveys were carried out before the service, good practice requires having access to the survey tool, a clear sense of how sampling was carried out, access to all raw data.
- Where baseline data is in the form of local or national statistics – the level of data would ideally be at the ward / unit level (rather than hospital or local area) and available for a number of months / years before your service began.

Comparison group

- Clearly describing how the comparison group compares to the group of people receiving your service in terms of their age, gender, or illness (e.g. on wards with a similar make-up of patients in a sister hospital with an older persons assessment unit) – although understandably this is not always possible due to costs or no local comparable ward.
- Using historical controls where the baseline data is available (e.g. comparing ‘old’ data from patients with the same condition as the patients receiving your service – or using the patients’ own medical history).
- Having raw data available on the comparison group (rather than just summary statistics).

Practice that could be improved

Baseline data

- Outcomes that are relevant, but use a bespoke untested tool (e.g. reduce patient’s anxiety, relevant measure but no tool described on how to do this).

Comparison group

- A comparison group that is very different from the intervention group (i.e. not similar from a demographics perspective), or very small sample size or lack of transparency about the comparison group’s characteristics.

Questions to ask yourself



- ❑ Are there specific plans to obtain baseline data (i.e. from surveys, patient records or local statistics)?
- ❑ How does the comparison group compare to your sample of patients / staff in terms of health condition and demographics?

Top tips



- Try to include a statistician in your team to widen your opportunities to explore different analysis approaches
- **Baseline data:** Consider the time schedule for the evaluation - what baseline data is feasible and what resources are required to collect this data (e.g. what routine data is available) – only measure what is necessary and sufficient. Consider the sampling requirements.
- **Comparison groups:** It is good practice to have a comparator group for patient outcomes (e.g. a sister ward or historical patient data), but if you don’t– it isn’t the end of the world, the evaluation is still worthwhile. Don’t try to use a comparator that isn’t appropriate. Be realistic about what might be possible given the resources, time and already available data.

What data do you plan to collect? How will you analyse it? (1/5)

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
STEP 3	Ethics & Governance	25
	Baselines & Comparisons	27
STEP 4	Data collection & Analysis	29
	Reporting	35

Collecting your own data

Draw on the resources of in-house analysts and other staff/ students/ volunteers who are well-positioned to collect data where possible. The Better Evaluation website provides [data collection toolkits](#) on some of the most common forms of data collection, including observation, surveys, focus groups. As the team collects data, save them to a secure database every day – and back this up. Keep GDPR guidance in mind when collecting and storing data and consult the R&D team with any queries. Some of the most common sources of data collected during the Helpforce VIP evaluations included:

- Adapted versions of surveys created by Helpforce to measure patient and staff experience. If you want to create your own surveys, you can borrow questions from large scale surveys such as the NHS Staff Survey (you will need to acknowledge the survey you have borrowed questions from in outputs and may need permissions depending on the survey)
- ‘Bespoke’ metrics that were developed by Trust data teams to measure the service’s impact on the hospital.

A common evaluation challenge is knowing when you have collected enough data. Using a ‘study design’ table (see an example on [slide 32](#)) can help you map your evaluation questions, approaches to collecting and examining data on your service, information you think you will obtain based on your planned approaches, and the stage at which you think you will have ‘good enough’ data.

Keep your theory of change in mind

- A significant part of the analysis process will require the team trying to understand the causes of outcomes by critically assessing the data collected against the ‘theory of change’.
- It can be helpful to undertake some early analysis. This will allow your emerging results and your interpretations of their causes to inform how the service (or evaluation approaches) can be improved. Be careful about publishing early results, as it is possible the data trends could change directions.

Sample size

- An important consideration for the evaluation is the possible sample size. For example, it may be that you have 200 volunteers, which would allow you to look at a very common outcome, but not a really rare outcome.
- Some VIP Trusts chose to undertake very specific and targeted interventions and as a result had very small numbers of volunteers and patients involved – for example in end of life care or restraint settings in particular wards. These small numbers made it difficult to determine any quantitative metrics that would demonstrate impact.

What data do you plan to collect? How will you analyse it? (2/5)

Quantitative data (i.e. numbers)

Quantitative data are any data that can be expressed as numbers. Analysis can be very simple or very complicated, depending on what data you have collected and what you are trying to find out. Make sure you know exactly what you want to measure, and how you are going to do that.

Collect baseline data – you cannot demonstrate any change in the outcome of a service if you don't know what was happening before the start of the intervention/change being evaluated. Note: data can be expressed as numbers or categories (e.g. male, female etc.).

- Survey – if you are using a survey, it is really important that it is well designed, or you won't be able to answer the questions you want to ask. See the CLAHRC West website for details of their courses <https://clahrc-west.nihr.ac.uk/training-and-capacity-building/>.
- Descriptive statistics – these are used when you want to show what is happening at one moment or over time, e.g. the number of referrals made to a new service. They can be presented in various ways, e.g. tables, graphs, bar charts, pie charts, run charts, etc. See <https://baselinesupport.campuslabs.com/hc/en-us/articles/204305665-Types-of-Descriptive-Statistics> for more general information and <https://qi.elft.nhs.uk/resource/run-charts/> for information about run charts. Be careful not to assume that any noted change has occurred because of the intervention, there may be other factors at play.
- Inferential statistics – these are used when you want to decide whether the intervention/change has led directly to any change in an outcome. There are many statistical tests available to use – choosing the correct one is very important, and depends on the test conditions and the level of data. See <https://baselinesupport.campuslabs.com/hc/en-us/articles/204305685-Inferential-Statistics> for more information. **If you are not sure which test to use, ask a statistician.**
- Analysis for economic evaluation – conducting a good economic evaluation is always a complex undertaking. **If your team does not involve someone with relevant expertise, talk to a health economist.**

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
STEP 3	Ethics & Governance	25
	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

What data do you plan to collect? How will you analyse it? (3/5)

Qualitative data (i.e. stories)

- If you are able to, record interviews and focus groups and transcribe them verbatim. If this isn't possible, take comprehensive notes either during data collection or when listening to a recording afterwards.
- **Thematic analysis** is the most common method used for analysis. This involves reading and re-reading the data transcript until you are very familiar with it, highlighting issues you think are important, and grouping them in a way that makes sense to you. You then draw your own conclusions about the key messages emerging from the data. See <https://sites.google.com/site/howtousethematicanalysis/home/what-is-thematic-analysis> for a step-by-step guide to thematic analysis.
- There are many other methods for conducting qualitative analysis. See <https://research-methodology.net/research-methods/data-analysis/qualitative-data-analysis/> for more information.
- By its very nature, qualitative analysis is a subjective exercise. This should not be thought of as an inherent weakness of the approach, it is simply one of its features; but it is important for researchers to be aware of how they may be influencing analysis. So **researcher reflexivity** is a key issue in qualitative research – you can find a good discussion about it at <https://researchdesignreview.com/2012/11/14/interviewer-bias-reflexivity-in-qualitative-research/>

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
STEP 3	Ethics & Governance	25
	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

Data collection and analysis (4/5)

Completing a study design table will make you think through all of the stages of the evaluation and make sure your approach matches your evaluation question. The table can also help you decide when you have all of the data you need to answer your evaluation questions. You can make the table as simple or complex as is helpful to you. See the example below.

Evaluation question	Data collection methods	Analysis	Type of data you think you will obtain	When will you have 'good enough' data?
Example service: Debrief Volunteers (who are mental health patients themselves) to support other patients who had been restrained to reduce the number of times patients will be restrained in the future.				
What is the impact of peer debriefing on patients?	<ul style="list-style-type: none"> • Patient experience survey of the entire hospital service (among those who had been restrained and chose to either reject or accept peer debriefing) • Interviews with patients experienced the debriefing • Observations of debriefings • Reviews of relevant ward documents and patient files 	Thematic analysis	<ul style="list-style-type: none"> • Reasons patients chose to accept or reject the peer debriefing • Patient experience of the debriefing (i.e. what went well and what could have been done differently) • Patient satisfaction with the debriefing process (including how it was introduced, carried out and then reviewed with staff and volunteers) 	<ul style="list-style-type: none"> • Captured the perceptions from a wide range of patients across at least half of the 10 wards (choosing 2-3 wards where restraint use or violence is most common) • Captured perceptions of some patients who rejected debriefing • Reviewed protocols for debriefing and was able to observe a few instances when these were put into practice (tracked if these were followed as well as deviations)
How is the debriefing exercise affecting staff?	<ul style="list-style-type: none"> • Interviews with staff • Staff experience survey • Staff sickness rate 	Thematic analysis	<ul style="list-style-type: none"> • Staff views on how and why the debriefing service is or is not working for them and for patients • Staff views on whether violence towards them has decreased since debriefing began 	<ul style="list-style-type: none"> • Captured the perceptions of staff from varied levels of seniority • Surveys received by at least half of the staff

Table source: Vindrola-Padros, C. (2020). Rapid ethnographies: A practical guide. Cambridge University Press.

Data collection and analysis (5/5)

Good practice Practice that could be improved

- Training is important so that data is all collected in a similar way
 - Be clear about the quality of data, and when significant data is missing, describing this as a limitation.
 - Monitor data quality as you go along – for missing data or any issues – adapt your approach if needed and be clear about if and when data collection tools and processes changed. Likewise, undertake analysis early and throughout the evaluation to inform changes to the service and evaluation.
 - Ideally you would have data on the people you expect to benefit from your volunteering service, but sometimes data is only available for a wider population (e.g. within certain age bands or within a set of wards, or only at the trust level).
 - Always be clear about your denominators - that is, the pool from which you pulled your sample (e.g. patients who improved their HbAc1 levels as a sample of all patients who received the service).
 - Where appropriate and justifiable, bring together samples from various implementation phases or across sites.
- Using capacity of students or volunteers to administer data collection without sufficient support, training or recognition for their contributions.
 - Selecting data with a significant time lag that causes a delay to the evaluation or failing to clarify the data accessibility early on
 - Not differentiating between target and actual sample size in reporting
 - Bringing together samples from different interventions in an effort to increase sample
 - Not being realistic about what can be achieved – quantitative data often requires large sample sizes

Questions to ask yourself



- Are the proposed indicators of good quality?
- Will the data collected be consistent?
- Is this data going to be easily accessible?
- Is there significant missing data?
- Is data available with sufficient frequency?
- Is the sample size likely to be sufficient for meaningful analysis?

Top tips



- Data collection could take twice as long as you expect. If you have not clearly defined what data is needed and why – you may also end up ‘wasting time’ analysing unnecessary data.
- If the data you plan to collect only covers a small group of people, qualitative rather than quantitative might be better.
- If collecting qualitative data, consider testing your planned data collection approach on a small sample (e.g. trying out your interview guide on a ‘test participant’). This can improve the ordering and phrasing of questions and improve the quality of data.
- Have a statistician or someone with relevant expertise look over your initial analyses to check you have the data you need.
- If data is not routinely collected, consider the time and resources required and available for data collection. If the data is being accessed via a data analyst or ‘gatekeeper’ agree a time schedule for how regularly you would like to receive the data.

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
STEP 3	Ethics & Governance	25
	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
	Ethics & Governance	25
STEP 3	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35

Step 4: Reflect and report on your findings

After this step, you should be able to answer:

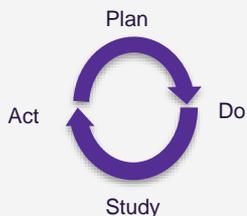
- Will you report findings throughout (for service change) or at the end?
- In what ways can you share your findings with varied audiences?



How will you reflect and report on the evaluation findings? (1/2)

Reporting findings throughout your evaluation using the PDSA cycle

- The Plan Do Study Act (PDSA) cycle is an approach to evaluation that is popular within the NHS and wider health services. It is an evaluation approach that is implemented alongside service change.
- This example of PDSA implementation, from the publication linked below, shows how the stages are implemented as part of service implementation:
 - **PLAN:** Focus group with black and minority ethnic women at risk of developing Type 2 gestational diabetes to develop a plan to support healthy lifestyle and weight loss.
 - **DO:** Implementation of the programme in a hospital setting.
 - **STUDY:** Evaluation data collected and analysed– food diaries, weight, blood glucose levels, interviews.
 - **ACT:** Programme content and delivery setting/ times changed to reflect evaluation learning. Programme re-implemented. See [here](#) and [here](#) for more details on PDSA cycles.



Reporting findings at the end of your evaluation

- It is important to consolidate all of your learning into a report, summary or email/leaflet at the end of your evaluation. Remember that while the focus of the report should be the findings, it is important to describe your intervention clearly. It is also important that you describe your evaluation approach and state the limitations / challenges you faced in the evaluation.
- Consider the various audiences who will be interested in your findings and where possible include recommendations for these groups directly linked to your findings. These recommendations will often be linked to the purpose of your evaluation. Think about why varied stakeholders will be interested and what they might want to do with information and design your output accordingly. For example, if you want someone in a senior position to make a specific decision or change based on your document, ensure that the relevant information is clear and ‘centre-stage’ and not hidden in the detail of a report.
- Your evaluation publication can take many forms: summaries, memos, news communications, website communications, MS PowerPoint presentations, posters, feedback workshops, conferences. Different audiences will want different output styles, be sure to ask your stakeholders (including patients) for their preferences and input in the ‘writing up’ phase.

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
STEP 3	Context	24
	Ethics & Governance	25
	Baselines & Comparisons	27
STEP 4	Data collection & Analysis	29
	Reporting	35

How will you reflect and report on the evaluation findings? (2/2)

Output type	For whom?	Top tips	Examples
Recommendations	<p>Policymakers</p> <p>Hospital boards</p>	Set out clear recommendations and try to keep them succinct – one page should do!	Transforming Ageing report
Case studies	<p>Volunteers</p> <p>Patients</p>	Case studies can be a nice format to show other volunteers or patients what you are doing and your findings in a digestible way. Make use of different websites: Helpforce, the hospital's – or create your own 'stories' page or send regular newsletters. Be sure to include a clear description of your programme.	<p>Volunteering Matters</p> <p>Voluntary Impact</p> <p>NESTA guidance (see item 28 for a newsletter example)</p> <p>NCVO</p>
Evaluation reports	<p>Policymakers</p> <p>Commissioners</p> <p>Hospital boards</p> <p>Other volunteering services</p>	Evaluation reports are a comprehensive and robust way to demonstrate your findings and make the case that your service has impact. If you find writing a full report a bit daunting, consider pairing up with a research organisation and asking them for support in writing it. Try and keep it concise – and make use of diagrams to make the content more accessible.	<p>Helpforce VIP evaluation</p> <p>Nesta Helping in Hospitals evaluation</p> <p>Corporate Volunteer Programme Valencia</p> <p>King's Fund evaluation of King's College Hospital volunteering programme</p>
Academic publications	<p>Academics</p> <p>Policymakers</p>	Writing academic publications can help cement your findings into a wider evidence base. Support from a research organisation might be helpful in targeting the right journals.	Evaluation of a volunteer hospital programme in rural Australia

STEP 1	Relevance	13
	Clarity & Complexity	14
STEP 2	Working group agreement	20
	Plausibility & Sustainability	22
	Indicators & Outcomes	23
	Context	24
STEP 3	Ethics & Governance	25
	Baselines & Comparisons	27
	Data collection & Analysis	29
STEP 4	Reporting	35



www.nuffieldtrust.org.uk



Follow us on Twitter – twitter.com/NuffieldTrust



Sign up for our newsletter – www.nuffieldtrust.org.uk/newsletter-signup

Contributing authors

Stephanie Kumpunen, Camille Oung, Nadia Crellin, Nina Hemmings,
Theo Georghiou, Natasha Curry (Nuffield Trust)

Jullie Tran-Graham, Mark Brooker (Helpforce)

Thank you to Rahel Spath for her comments on drafts of this guide.

Cite this guide as:

Nuffield Trust, Helpforce (2021). Evaluating the impacts of your volunteering service: A guide for managers of hospital volunteering services. Research report. Available at: INSERT URL

Appendix A: Glossary of evaluation terminology

Aim: An aim is the overall, or wider objective of a project or action OR the anticipated outcome that is intended or that guides one's planned action. It is useful to break aims down into two different categories: overall aim and specific aims.

Attribution: “causal link between observed (or expected to be observed) changes and a specific intervention.” WHO (2013) Evaluation Practice Handbook apps.who.int/iris/bitstream/10665/96311/1/9789241548687_eng.pdf

Baseline: “A set of measurements before any intervention starts (after any initial ‘run-in’ period with no intervention), with which subsequent results are compared.” <https://www.nice.org.uk/article/pmg20/chapter/glossary#baseline>

Benchmark: “evaluate (something) by comparison with a standard: we are benchmarking our performance against external criteria”

<http://www.oxforddictionaries.com/definition/english/benchmark>

Beneficiaries: The individuals, groups, or organizations that benefit from an intervention, project, or program.

Counterfactual: A hypothetical statement of what would have happened (or not) had the program not been implemented.

Evaluation: A process that attempts to determine as systematically and objectively as possible the relevance, effectiveness, and impact of activities in light of their objectives.

Findings: Factual statements about a project or program which are based on empirical evidence. Findings include statements and visual representations of the data, but not interpretations, judgments or conclusions about what the findings mean or imply.

Impacts: The anticipated end results or long-term effects of a program: for example, changes in health status, such as reduced disease incidence or improved nutritional status.

Indicator: An indicator is a ‘unit of measure’ that identifies change (in quality or quantity) within a defined period of time. It allows to judge if an intervention was successful or not. Indicators can be divided into output indicators or outcome indicators. Output indicators are used to assess whether and to what extent outputs have been delivered. Outcome indicators are used to assess whether or the degree to which the expected outcomes have occurred.

Key informant: Person with background, knowledge, or special skills relevant to topics examined by the evaluation; sometimes an informal leader or spokesperson in the targeted population.

Monitoring: Monitoring is the routine process of data collection and measurement of progress toward program objectives. It involves tracking what is being done and routinely looking at the types and levels of resources used; the activities conducted; the products and services generated by these activities, including the quality of services; and the outcomes of these services and products

Monitoring and evaluation (M&E) plan: A comprehensive planning document for all monitoring and evaluation activities within a program. This plan documents the key M&E questions to be addressed: what indicators will be collected, how, how often, from where, and why; baseline values, targets, and assumptions; how data are going to be analysed/interpreted; and how/how often report will be developed and distributed.

Appendix A: Glossary of evaluation terminology

Outcomes: Outcomes are the changes, benefits, learning or other effects that happen as a result of services, such as changes in targeted attitudes, values, behaviours or conditions between baseline measurement and subsequent points of measurement. Changes can be immediate, intermediate or long-term; positive or negative; expected or unexpected.

Outcomes can be relevant for individuals, families, whole communities, organisations or wider issues such as policy or the environment. Outcomes relate to specific aims/purpose. Outcomes are all the changes that may actually occur when you carry out activities to achieve a specific aims. They may not always be the same as the outcomes you planned. Outcomes can be a direct and/or indirect result of outputs. Since there are often smaller changes that need to happen before the main desired outcome can be reached, intermediate steps have to be acknowledged along the way to the final outcome. Such steps are called intermediate or interim outcomes. Usually two different types of outcomes are distinguished:

Soft outcomes are typically defined as intangible, a matter of degree and more difficult to measure (e.g. changes in attitudes, self-perception or certain skills areas). These are often, but not always, intermediate outcomes.

Hard outcomes are defined as quantifiable and often more easily measured (e.g. organisations raise more money as a result of improved fundraising through training). Hard outcomes are not better than soft outcomes, simply different.

Plausible: An explanation or statement that is **plausible** seems likely to be true or valid.

Qualitative data: Non-numerical data rich in detail and description that are usually presented in a textual or narrative format, such as data from case studies, focus groups, interviews or document reviews. Used with open-ended questions, this data has the ability to illuminate evaluation findings derived from quantitative methods.

Quantitative data: Numeric information, focusing on things that can be counted, scored and categorized; used with close-ended questions, where participants have a limited set of possible answers to a question. Quantitative data analysis utilises statistical methods.

Questionnaire: Highly structured series of written questions that is administered to a large number of people; questions have a limited set of possible responses.

Reliability: The consistency of a measurement or measurement instrument over time. Consistent results over time with similar populations and under similar conditions confirm the reliability of the measure. (Can the test produce reliable results each time it is used and in different locations or with different populations?)

Appendix A: Glossary of evaluation terminology

Stakeholders: Stakeholders are individuals, groups or organisations who influence or who are directly or indirectly influenced/affected by the service. Stakeholders have a significant interest in the success or failure of the service. The involvement of the largest possible number of stakeholders into the management of the service and its evaluation (planning, implementation, evaluation, reporting) will promote understanding, enlarge ownership, and foster sustainability of the service.

Surveys: Uses structured questions from specially designed instruments to collect data about the feelings, attitudes and/or behaviours of individuals.

Sustainability: Sustainability describes the process of continued existence of benefits from an intervention after the concrete implementation has been completed. A service is sustainable if the changes purposely set in motion and supported (effects, processes, etc.) during the duration of the project/programme can be continuously developed and maintained over time.

Target groups: Target groups are those individuals or groups that a project or programme is targeting with its intervention. A target group consists of specific individuals, specific organisations, or specific institutions, etc. for whom project services are intended. Target groups can differ from beneficiaries of a project for whom the benefits of the intervention are intended. (E.g. an intervention might target parents through training in child care and the preparation of healthy food in order to eliminate obesity with the beneficiaries of the project, their children).

Theory of change: A set of assumptions about how and why desired change is most likely to occur as a result of your program, based on past research or existing theories of behaviour and development. Defines the evidenced-based strategies or approaches proven to address a particular problem. Forms the basis for logic model planning.

Validity: The extent to which a measure of a particular construct/concept actually measures what it purports to measure; how well a test actually measures what it is supposed to measure.

Bibliography for glossary:

<https://www.sportanddev.org/en/toolkit/monitoring-and-evaluation/glossary-key-terms-monitoring-and-evaluation>

<https://class.csuohio.edu/sites/csuohio.edu.class/files/EvaluationTerminology.pdf>

<https://www.measureevaluation.org/resources/publications/ms-07-20-en>

<http://www.nhsevaluationtoolkit.net/glossary/>

https://pdf.usaid.gov/pdf_docs/Pnado820.pdf

Appendix B: Questions to discuss with your working group

The service

Existing evidence: What is already known from the evidence and how to build on it (i.e. lessons learnt from previous internal/external efforts to solve the problem, academic literature)?

Improving the service's contents or its implementation: Are there ways in which the service could be improved (based on the evidence found) or potential risks that can be mitigated against (based on local knowledge or expectations)?

Sustainability: Are there plans for service sustainability in place beyond any initial trials/pilots?

Stakeholders

Key people to involve: Who else needs to be involved in the steering or working group, who could be helpful to engage for their opinions or access (e.g. patients for their views on the service, or Quality Improvement teams with access to data)? Is it possible to have participation from people at all levels within the organisation? Is it possible to agree up front in which areas and ways volunteers and staff may need to be involved, their responsibilities, and how their day-to-day work will change (if at all)?

Evaluation capacity: Who is available to help carry out or advise on the evaluation? Working with in-house analysts, QI teams, R&D teams, and information governance on site is recommended for gaining access to data and its interpretation. If wider support with evaluation is needed, consider connecting with other Helpforce sites to form a community of practice or learning circle, attending conferences or seek out coaching / peer coaching or mentoring.

Evaluation queries

Key questions: What are the key evaluation questions that should be prioritised? What questions cannot be asked / answered because of the resource implications or other challenges? Is data available and/or can data be collected to answer the key questions?

Unintended consequences: Are there any unintended consequences (positive or negative) that would be worth measuring in an evaluation?

Methods and data: Which approaches to data collection and analysis are most suitable?

Appendix B: Questions to discuss with your working group

Evaluation planning

Purpose: What is the purpose of the evaluation (e.g. will it be used to better understand if the volunteering service is achieving its aim in order to gain buy-in for it to be sustained/scaled to other hospitals at the Trust?)?

Resources: What resources (time, money, expertise) will be needed?

Governance: What ethics or governance approvals might be needed?

Defining success: What would a 'successful' evaluation look like? UK Evaluation Society Standards suggest that evaluations should be independent, transparent, ethical, impartial, of high quality (as defined by the Evaluation Office), timely and used to inform decision making.

Engagement and communications: How will you communicate with stakeholders throughout the evaluation? Talk through your evaluation plan with staff, commissioners and patients. Consider communicating regularly (via newsletters, emails, and at meetings) about the evaluation plans, progress and emerging findings as you go along (rather than just at the end) to engage audiences throughout the evaluation.