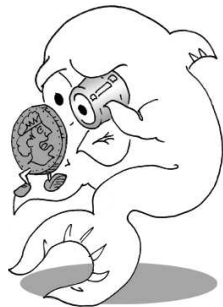


1 What is an Evaluation?



What is an evaluation?

What is an evaluation?
 What do we mean by small-scale evaluation?
 Why evaluate?
 What shapes an evaluation?
 What evaluation can and cannot do?
 Who should do the evaluation?
 How do we do an evaluation?

OVERVIEW

What is an evaluation

Evaluation means to assess the value of something. This raises a very important question for any evaluation – valuable in terms of what? Why an evaluation is being done, and who it is being done for, will determine how it is carried out and what it aims to assess.

Results of evaluations are often used to make decisions about competing health care options. When we do an evaluation we are normally asking a question like; ‘does it work?’ and ‘is it worth it?’

Importantly, evaluation should be systematic in terms of assessment and reporting. It therefore needs to be carried out in a way that allows other people to follow the aims and methods and perhaps to repeat the evaluation themselves elsewhere. It should always be easy to see from the write-up of an evaluation how it was done, what information was collected and how it was analysed. Results should be provided as well as conclusions so that their validity (how believable they are) can be assessed.

What do we mean by small-scale evaluation?

Carrying out research can be time consuming and expensive. Evidence that is published about the clinical effectiveness of treatments or services is often produced through large-scale – usually multi-centre – studies that can cost hundreds of thousands of pounds to complete. In this book we aim to provide guidance for those who wish to undertake local, small-scale

2 Small-Scale Evaluation in Health

studies that are run from within existing budgets, or with a modest amount set aside to carry out or commission an evaluation. It is important to recognise that, within these constraints, it is usually not possible to answer questions that require a complex study design, recruitment of hundreds or even thousands of patients, and sophisticated statistical or qualitative analysis of data. Questions concerning the clinical effectiveness of new treatments or services that include the assessment and comparison of health outcomes of patients (for example, measurement of symptoms; assessment of disability; assessment of quality of life) are unlikely to be answerable within small-scale, local, low budget evaluations. But there are important questions that service commissioners and planners do want to know about how new services are running, such as: are patients seen more quickly? Are staff happy with the new arrangements? Are sessions well attended? This book is designed to help practitioners and managers to plan and carry out evaluations of services that are relevant to their needs, and which will produce findings that are valid and can be relied upon for informing decision making.

In summary: Evaluation is a systematic assessment of the value of something. Understanding gained from an evaluation allows us to help plan a programme better, make improvements and inform decisions about whether a treatment or process should continue, and whether it should be more widely implemented.

Why evaluate?

Effects of treatments or services are not always predictable. Sometimes a service or treatment is introduced with the best possible intention but the consequences can actually result in people being worse off. For example, the introduction of fruit stalls in schools led to children eating less fruit (see Case Study 1.1). Evaluation can identify any negative impacts of services as well as the positive.

CASE STUDY 1.1

Evaluation of fruit stalls in four schools in South Wales

Background: In an initiative to encourage healthy eating, the Local Health Board provided the equipment to develop fruit stalls in participating local schools. These stalls were set up as non-profit-making shops offering low-cost fruit and vegetable to children along with healthy snacks.

Methods: Stalls were implemented in stages, with two schools introducing stalls early and two control schools introducing stalls at a later stage. Number of items of fruit eaten was compared by

observation of the behaviour of randomly selected children (n=40) in control and intervention schools.

Results: Before the intervention children in both the control and intervention school consumed on average 1.5 portions of fruit or vegetable per child during school time. After the intervention the control school children consumed an average of two portions and the intervention school children consumed an average of 0.5 portions per child (much of this was due to the intake of tomato sauce). [$p < 0.001$ Difference 1.5 (95 per cent CI: 1.01-1.99)].

Observations of lunch box contents showed that parents in control schools gave children fruit and juice. However, in the intervention schools, parents gave the children money to buy low-price fruit and no longer provided fruit and juice.

Conclusion: In the intervention schools, children choose to keep the money rather than buy fruit and are more likely to purchase non-health items such as crisps and chips and sweets. Therefore, fruit stalls in schools do not encourage fruit consumption, but schools that encourage parents to provide a health lunch and encourage a healthy lifestyle can influence total fruit and vegetable consumption during the school period. These findings are based on schools that already have a good record in healthy eating. Findings may differ in schools with a poor record in fruit consumption.

Source: early unpublished results.

Question: What could we have done first which could have helped predict that fruit stalls would not work?

In addition, money for health care is in short supply. So that scarce resources are spent in a way that achieves benefits across communities, understanding is needed about the effects of innovations in treatments and services. We need to collect and report information about effects as well as about how the treatment or service is delivered. For example, we may wish to compare two different models of care (such as different types of anaesthesia, Case Study 1.2) designed to deliver similar results in order to decide which is more acceptable to patients. This information can help to plan both local service delivery, to the benefit of users, as well as potentially the wider population to inform larger-scale research for service and policy development.



Patients' self-evaluation after 4–12 weeks following xenon or propofol anaesthesia: a comparison

Aims: The evaluation set out to assess patients' self-assessment of the two types of anaesthesia after having an operation. The two types of anaesthesia evaluated were propofol (an injectable general anaesthetic) and xenon (gas anaesthetic).

Methods: The evaluation looked at 160 patients aged between 18 years to 60 years who were undergoing elective surgery. Patients were randomly given either an injected anaesthetic (propofol) or a gas anaesthetic (xenon). The patients were telephoned to ask their assessment of the anaesthetic. The patients did not know which type of anaesthetic they received and the people conducting the interview did not know which type of anaesthetic the patient had received. Patients were asked about: their evaluation of the anaesthesia; choice of the same anaesthesia for the future; recall of uncomfortable feelings after the anaesthesia.

Findings: 116 took part in the telephone interview. The two study groups (that is, those taking injected, those taking gas anaesthesia) were comparable with respect to age, weight, height and gender. There was no difference between the groups in terms of self-evaluation of anaesthesia, taking the same anaesthesia again or recall of uncomfortable feelings. However, the post-operative pain and appetite/thirst were higher (worse) for those taking the gas (xenon) anaesthesia compared to those receiving the injected anaesthesia (propofol).

Source: Adapted from an abstract published by M. Coburn, O. Kunitz, J. Baumert, K. Hecker and R. Rossaint *European Journal of Anaesthesiology*, 2005, 22 (11): 870–74.

Question: What conclusions can we come to from the results of this evaluation?

If services are not evaluated in a rigorous and systematic way, then we cannot be sure that resources are used to best effect – some services may continue to be funded when they are ineffective, and others may not be implemented, or be discontinued when they could achieve more. Basing decisions about how to use scarce resources only on the views of those delivering services may mean that those who are able to shout loudest – or those who have the most influential contacts – may be the most successful in attracting funding for their services. This cannot be the most rational way of spending money for the benefit of the whole community, and the

decision-making processes are unlikely to be understandable or justifiable to people not directly involved. We can be much more confident about making the best possible decision if we know that our decision is based on evidence obtained from a rigorous and systematic evaluation.

Therefore, evaluations are carried out to make services more effective and justify decisions for funding in a transparent way.

What shapes an evaluation?

The aims of evaluation can vary depending on who it's being done for. Project funders often commission evaluation because they want to assess whether their money was used successfully. For example, funders might wish to evaluate an exercise programme in terms of effectiveness (see Case Study 1.3); they would therefore examine in people who did the exercise whether their fitness improved (muscle strength and blood pressure). Who the evaluation is for is also very important in how the evaluation is carried out and what it is supposed to do. An evaluation can only provide the evidence it is set up to provide, and this relies on appropriate identification of objectives and methods to match. Evaluations rarely find anything useful if they have not been planned and designed properly.

Evaluation of a community-based exercise programme for elderly Korean immigrants

CASE STUDY 1.3

Aims: The evaluation set out to examine the feasibility and effectiveness of a modified exercise programme for elderly Korean immigrants (EKIs).

Methods: We recruited elderly Korean immigrants through posted fliers and with help from a Korean social worker. Participants of the exercise programme (n=13) were elderly Korean people whose average age was 77 years, but age ranged between 67–86 years. A Korean-American instructor taught a modified version of an evidence-based exercise programme 3 times weekly, 50 minutes per session, for 12 weeks.

The exercise programme was evaluated by examining exercise adherence (i.e. did the participants do the exercises?), measures of health before and after the exercise programme and satisfaction with the exercise programme. Group discussion was used to evaluate satisfaction with the exercise programme.

Findings: Participants showed improved muscle strength, agility/balance and blood pressure after the exercise programme.

(Continued)

6 Small-Scale Evaluation in Health

(Continued)

All participants were satisfied with the exercise programme, and participation rates were good (nine participants attended ≥ 80 per cent of classes).

Conclusions: The exercise programme was feasible for this sample and should be evaluated in a larger population of elderly Korean immigrants and in populations of other ethnic minorities.

Source: Adapted from M.K. Sin, B. Belza, J. Logerfo and S. Cunningham (2005) 'Evaluation of a Community-Based Exercise Programme for Elderly Korean Immigrants'. *Public Health Nursing*, 22 (5): 407–13.

Question: Why is this an evaluation and not audit?

Appropriate evaluation can allow us explicitly to include the views of everyone involved in a treatment or programme, such as patients, service users, carers and health care workers as well as managers and policy makers. People from these different groups are likely to have different priorities and perspectives about what the most important effects of a service or treatment might be. This is especially relevant, for example, to the re-location and centralisation of services. Without a well planned and appropriately designed evaluation, it is unlikely that this range of views will be reported.

Therefore, the questions asked in an evaluation are shaped by who wants the evaluation done (who is funding it) and who is going to use the results and findings in the end.

What evaluation can and cannot do

Evaluation can give a clear answer to a specific question. However, evaluation cannot ask every possible question and provide every possible answer. The questions asked and the answers sought must be clearly planned and defined from the beginning. For example, unpredictable effects are unlikely to be identified unless methods are explicitly included to ensure they are picked up.

Evaluation can explore whether programme targets, or aims, are met and provides evidence to inform decision making. However, it cannot look at everything a programme might do. For example, we might wish to find out through evaluation whether nurse-led telephone helplines reduce call-out of emergency ambulances. We might find out that they do not, and this could be seen by some as failure of the service. However, if we evaluated use of the helpline in terms of users' self-confidence to deal with their condition themselves, we might find it is a great success in these terms.

Most evaluations ask questions like 'Does it work?' but not necessarily 'How or why does it work?' These questions can be specifically built into an

evaluation, but are not essential to all evaluations – if this understanding is required it needs to be identified at the outset so that appropriate objectives and methods can be defined, planned and incorporated into the evaluation. Questions like how and why something works which involve describing the ways in which services or treatments are delivered may require more complex approaches and methods. Sometimes evaluations are carried out to gather information about impact, with an assumption that understanding of how and why the effects have happened will be gained as an automatic outcome of the study. It is important to recognise that this understanding is unlikely to be gained unless methods designed to capture this information are included. An evaluation may focus on the weaknesses of a system in order to investigate how to improve these areas. For example, interviewing people who do not attend regular screening tests for cervical cancer could help to highlight how the system could be made more accessible.



Understanding of how

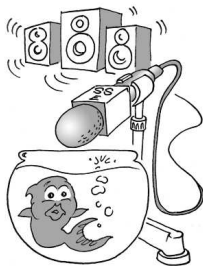
It is important to recognise the limitations of small-scale evaluation of an established service. For instance, in an evaluation of a diabetes educational programme it would be very difficult to show that people on the programme have fewer eye problems or kidney complications, and even harder to demonstrate that the programme has led to any improvements in health. So, alternative or 'proxy' measures could be selected, such as attendance rates or changes in people's health-related behaviour (exercise levels, diet changes), or people's views about the programme could be gathered through interviews or focus groups.

Who should do the evaluation?

Many experts recommend that evaluation of a programme or treatment should be carried out by a person who is independent, as this will make it more likely that they are not biased. If a person's job depends on a treatment working, then they may be more inclined to show beneficial effects of the treatment and their findings may not be accurate or unbiased. This approach requires the person doing the evaluation to be 'neutral'. However, in order to understand the effects of a treatment or programme, it may be important to involve the people who provide and use the service at the planning stages of the evaluation. They will be the ones who understand what aspects are important to include in the evaluation and the best way to incorporate the gathering of data alongside treatment delivery. Therefore, for successful evaluation, it is important to involve those giving and receiving the treatment. Indeed, in some forms of research, for example, action research, active participation of people involved in service delivery as well as service users is fundamental to the evaluation process.

It is important to make efforts to avoid any bias creeping in whilst the evaluation is underway, for example, in data collection or analysis. This can be

8 Small-Scale Evaluation in Health



done, for instance, by keeping records of those who are not invited to participate as well as patients recruited to the study, or by 'blinding' or keeping hidden information from the person carrying out the analysis which identifies patients and whether or not they received the new service. This will help to ensure that biases are minimised. This need not mean that stakeholders, such as service providers and users, are kept out of the process of interpreting results. It is normally encouraged as good practice.

In summary: A good evaluation should be systematic, rigorous, and written up fully, so that others can see how the evaluation was carried out and conclusions drawn are based only on the results reported.

How do we do an evaluation?

The steps are:

- Develop a clear idea of what question we are asking.
- Examine what other people have done in this area.
- Form a plan of what we would like to do .
- Do a trial run of what we intend to do (a pilot to identify any problems).
- Do the collection of all the data.
- Analyse the data.
- Write a report that includes objectives, methods, results, interpretation and recommendations.

Summary

- Evaluation assesses the value of something.
- It can be used to: see what has been achieved; identify strengths and weakness of a system; compare the effects of a programme with other similar programmes; share experiences with others; assess whether the cost of something is worth the benefits it gives; make something more effective; help plan for the future.
- WHY you want to do an evaluation will determine HOW you do the evaluation.
- Good evaluation is systematic and rigorous, with methods and results described in full, and the conclusions drawn should be supported by evidence from the results of the evaluation.
- It is vital to recognise from the outset that any evaluation will only provide answers to questions which are clearly defined and addressed in the project. This means that unless evaluations have clearly identified aims and goals which are matched with appropriate methods, they are unlikely to succeed.

Frequently asked questions

What is the difference between clinical audit, research and evaluation?

Research in health care is enquiry designed to produce new knowledge, that can be applied more generally than in the setting in which it took place. The term can be used quite widely, and covers all types of approaches – including randomised controlled trials; exploratory enquiry; work to develop methods; and participative, action research.

Evaluation is a type of research, and examines the evidence to help decide what is best practice. Evaluation research always assesses the value of something, often in terms of costs and benefits.

Clinical audit examines whether treatments or services are being delivered to patients according to best practice. In audit we are comparing what we are doing against an agreed – and usually evidence-based – standard.

In audit, patients are not asked to have or do anything beyond the normal clinical care. However, in an evaluation (and research) we may ask patients to take or do things that are new and are not within normal clinical practice in order to examine the effect of these new treatments or procedures.

For more information see: www.geh-tr.wmichs.nhs.uk/services/orthopaedics/Audit%20and%20research/auditvsresearchleaflet.pdf
www.btuheks.nhs.uk/cg/difference-audit-research.pdf

Importantly, we normally need to gain ethical approval from a Research Ethics Committee before carrying out any research study in a health care setting. By contrast, clinical audit is seen as a fundamental part of the delivery of health care, for which ethical approval is not required. The carrying out of evaluation can require ethical approval, although it can be viewed as part of service delivery, depending on its aims and methods. More guidance about whether we need to apply for ethical approval for our proposed study can be found at www.nres.npsa.nhs.uk/docs/guidance/is_and_As_Differentiating_Research.pdf

Should we evaluate our own programme or commission others to do this?

We can choose to do either – or a bit of both. In designing the evaluation there should be input from the people involved in delivering the programme and involved in the programme. However, in conducting the evaluation (e.g. doing interviews, recording the results) the ‘researcher’ should be independent of the programme, unless an action research approach is explicitly taken. The interpretation should be based on the findings of the independent researcher put in context with discussions with the people involved in receiving and delivering the programme.

The important thing is to recognise the strengths and weaknesses of each approach, and decide what suits us best, considering the resources, skills and time that are available for our proposed evaluation.

We're not sure about how to evaluate our programme; isn't it best to collect and store a wide range of data and then work out our evaluation plan later?

No, if you don't know what you want to evaluate, how do you know what to collect? You might not collect something important, and will certainly collect huge reams of things you never needed. This is a great waste of resources and yours and participants' time. Conducting an evaluation like this will almost certainly lead to you not being able to answer your questions properly.

Further reading

Marie-Therese Feuerstein (1997) *Partners in Evaluation: Evaluating development and community programmes with participants*. London: Macmillan Education.

David Grembowski (2001) *The Practice of Health Program Evaluation*. London: Sage.

Colin Robson (2000) *Small Scale Evaluations: Principles and practice*. London: Sage.

Ray Pawson and Nick Tilley (1997) *Realistic Evaluation*. London: Sage.



Web pages

www.elwa.org.uk/elwaweb/elwa.aspx?pageid=2065
www.emro.who.int/rpc/pdf/healthresearchers_guide.pdf