

Service Evaluation

An introductory guide



Contents

Key Points.....	2
Introduction.....	3
Types of Evaluation.....	4
Designing an Evaluation – Seven Questions to Consider.....	5
Glossary.....	10
Further Information.....	12

Key Points

- All services that are delivered should be subject to an appropriate and proportionate evaluation
- Evaluations help identify what works and why, provide learning to improve effectiveness (of services), highlight good practice and unintended consequences and develop and test new ideas
- Evaluations help demonstrate potential cost savings, cost effectiveness and value for money of services being delivered
- The risk of not evaluating, or of poor evaluation, is that it will not be possible to know where services being delivered are ineffective, or worse, result in worsening the circumstances of older people
- Good evaluation evidence is dependent on the design of the evaluation and the design and implementation of the service being delivered
- Seven helpful questions to consider when planning an evaluation are:
 - What are the aims of the service to be delivered?
 - What are the questions to be answered by the evaluation?
 - How will attribution be measured?
 - What data will be collected and how will it be collected?
 - What resources will be required to carry out the evaluation?
 - How will the evaluation be quality assured?
 - How will evaluation findings be disseminated?

Introduction

An essential principle for all to adhere to is that of attempting to achieve the greatest impact at the lowest cost, as this ensures that a pool of funding can help the maximum number of older people.

This requires that services delivered are based on credible and reliable evidence, and high quality evaluation is vital to achieving this.

The risk of not evaluating, or of poor evaluation, is that we would not be aware of situations where services we deliver are ineffective, or worse, result in worsening the circumstances of older people. In addition without such evidence we would not be able to confidently claim that the funding is effectively spent, even where in reality, the services we are delivering are highly successful. Nor would we be able to identify why a service is or is not successful, and use the learning to improve the effectiveness of the service being delivered.

All services that we deliver should therefore be subject to an appropriate and proportionate evaluation.

This guidance document is not a textbook on evaluation; rather it provides the key questions that should be considered when developing an evaluation of a service. For further information, and more detailed texts, please see the references provided at the end of this document.

Types of Evaluation

An evaluation is an impartial process that attempts to answer two broad questions:

- How was the service delivered? **Process evaluations** assess whether a service is being implemented as intended and what, in practice, is felt to be working more or less well, and why.
- What were the impacts of the service? **Impact evaluations** attempt to provide an objective test of what changes have occurred, and the extent to which these can be attributed to the service.

Process evaluations include the collection of qualitative and quantitative data, covering both subjective issues and objective aspects of the implementation and delivery of policies or services. Impact evaluations tend to focus on the collection of quantitative data and the use of statistical analysis to identify the impact that can be attributed to the service. Qualitative research methods can also help assess the impact and in many cases triangulating between different sources of evidence is recommended.

Often understanding the cost-saving, cost-effectiveness or value for money of a service delivered is of interest, and involves combining information from both types of evaluation approaches. Both types of evaluation approaches should therefore be designed and planned at the same time to ensure all relevant information is captured.

Designing an Evaluation – Seven Questions to Consider

There are a number of stages in designing an appropriate and proportionate evaluation, which will provide credible and reliable evidence. Good evaluation evidence is not simply dependent on the design of the evaluation, but also on the design and implementation of the service being delivered.

The design and implementation of a service affects the quality and type of information that can be collected and therefore the types of questions that can be answered from the evaluation.

In practice it is therefore vital that the evaluation is planned at the same time as the design of the service. These two are complementary to each other and working on both at the same time will improve the design and implementation aspects of the service, and the evaluation.

This section presents seven key questions to consider when planning an evaluation. For more detailed guidance on which approaches and methods are most appropriate in which circumstances, see the list of documents and texts referenced at the end of this guidance document.

The seven key questions to consider are:

Question 1: What are the aims of the service to be delivered?

It is important to understand the assumptions and evidence that underpin the service that is being designed to achieve a series of aims and objectives. A *Theory of Change* or *Logic Model* which clearly sets out the links between the inputs, the activities that will then follow, the outputs that then result and the intended outcomes and impacts is beneficial for several reasons: it can help highlight gaps in the evidence, guide the design of data collection and monitoring processes and inform the objectives of the evaluation and development of research questions.

Question 2: What are the questions to be answered by the evaluation?

It is important to understand the gaps in the evidence and therefore the type of information that will need to be sought, along with an understanding of how the results are proposed to be used and for what audiences.

For example, if we already know that a service achieves the desired outcomes, is it necessary to carry out another evaluation? The answer to this question is dependent on a number of factors, but crucially on whether the service is being delivered under the same or similar circumstances as the intervention was tested in the evaluation and whether the service has been substantially changed or adapted.

Assessing these gaps will help provide focus on the types of questions that need to be answered (including assumptions underpinning the service model that need to be tested), and the types of information that need to be gathered. In particular it is important to determine what will be added to the existing body of knowledge by answering (through the evaluation) the questions identified.

Question 3: How will attribution be measured?

One of the most challenging aspects of an evaluation is attempting to obtain credible and reliable evidence of what would have happened in the absence of the service delivered (i.e. the counterfactual). This is important because in many cases a number of factors, other than the service, drive changes in outputs, outcomes and impacts. Therefore it is important to establish a *comparison* or *control group*, otherwise it is more difficult to claim with confidence the extent to which changes can be attributed to (i.e. are the result of) the service delivered. There are several approaches to creating a *comparison* or *control group* and critical to this is how the recipients are chosen – an explanation for why good evaluation evidence is dependent on good policy and service delivery design¹.

Question 4: What data will be collected, and when & how will it be collected?

The evaluation questions that need to be answered (along with the information required for ongoing monitoring of implementation and delivery) will determine the types of data that need to be collected, and when those data need to be collected. In most cases the requirements will involve the collection of both qualitative and quantitative data, and will need the use of surveys and interviews. Data collection will often need to commence before the service is implemented in order to collect *baseline data* from which changes in outputs, outcomes and impacts can be observed.

¹ [Quality in policy impact evaluation: understanding the effects of policy from other influences](#) (HM Treasury); [Test, Learn, Adapt: Developing Public Policy with randomised control trials](#) (Behavioural Insight Unit, Cabinet Office) – these two texts are recommended starting points for detailed guidance on approaches for creating a counterfactual (including their strengths and weakness)

Question 5: What resources will be required?

Carrying out an evaluation is usually straightforward – the challenge is to do so in a way that leads to credible and reliable evidence being obtained. It is therefore advisable to start early and involve from the beginning people that have an understanding and experience of developing and carrying out high quality evaluations. This may be in the form of in-house expertise or external expertise. In addition to evaluation expertise, resources will be required for the project management of the evaluation strand of the service being delivered: from the planning of the evaluation, to commissioning (if appropriate), to the day-to-day management, to arranging for quality assurance, to the dissemination of the findings. The resources available will influence the scale and form that the evaluation can take.

Question 6: How will the evaluation be quality assured?

Quality control and quality assurance are crucial aspects of evaluation; without these it is difficult to have confidence in the credibility and reliability of the findings. Therefore it is important to consider the arrangements required to ensure the evaluation fulfils the principles of *independence, inclusivity, transparency* and *robustness*. This can be achieved through having processes in place and people with the appropriate knowledge and skills to ensure evaluations are designed, planned and delivered to professional standards, are informed by a governance community (such as a steering group that includes recipients, delivery bodies and stakeholders) and that involve consistency in data collection, methodology, interpretation of findings and reporting.

Question 7: How will evaluation findings be disseminated?

The dissemination of the evaluation findings, including how and to whom they will be presented, plus how they will feed back into the policy process, will influence the type and format of evidence that is needed. A range of activities should be considered to disseminate the findings (in addition to the publication of a final report), and it is important that the key conclusions and messages are conveyed with brevity and clarity.

All stakeholders should be involved in seeing the evaluation findings, and to help with this process, where possible, findings should be shared at the earliest opportunity. The evaluation findings should also inform *forward planning*. They should be seen as providing vital intelligence for service development and not simply an historic record of what has happen.

The evaluation findings should also be made publically available. This will help improve the credibility of the findings by opening them up for wider peer review and helping share the learning to as wide an audience as possible.

Glossary

Additionality – an impact arising from an intervention is additional if it would not have occurred in the absence of the intervention

Appraisal – the process of defining objectives, examining options and weighing up the benefits, costs, risks and uncertainties of those options before a decision is made

Counterfactual – the circumstances that would have arisen (i.e. what would have happen) in the absence of the intervention

Cost Benefit Analysis – analysis which quantifies in monetary terms as many of the costs and benefits of an intervention as feasible, including items for which the market does not provide a satisfactory measure of economic value

Control Group – the group of people that does not participate in the intervention

Cost Effectiveness Analysis – analysis that compares the costs of alternative ways of producing the same or similar outputs or outcomes

Deadweight – expenditure to promote a desired activity or result that would in fact have occurred without the expenditure

Evaluation – an impartial process to assess how an intervention has been implemented and delivered, and the impact that it has had

Inputs – resources required to deliver the intervention

Impact evaluation – an objective assessment of what changes have occurred, and the extent to which these can be attributed to the intervention

Impact – the wider and longer term economic and social changes that result from outcomes

Monitoring – systematic collection and analysis of information as an intervention progresses, to help assess delivery against plans and milestones along with reviewing key performance indicators

Opportunity Cost – the value of the best alternative forgone

Outcomes – the changes that are brought about by the intervention

Output – the products, goods or services and other immediate results that the service has delivered, which are expected to lead to the achievement of outcomes

Process Evaluation – assessment of whether a service is being implemented as intended and what is working more or less well and why

Social Benefit – the total increase in the welfare of society from an action; the sum of the benefit to the persons carrying out the activity, plus the benefit accruing to society as a result of the action

Social Cost – the total cost to society from an action; the sum of the opportunity cost to the persons carrying out the activity, plus any additional costs imposed on society from the action

Further Information

Age UK (2017) [Decision Tree – What Research Methods to Use](#), London, Age UK

Age UK (2017) [Top Tips – Designing & Developing Surveys](#), London, Age UK

Age UK (2017) [Guidance – Sample Size Estimation for Qualitative Methods](#), Age UK

Haynes, L., Service, O., Goldacre, B. and Torgerson, D. (2012) [Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials](#), London, Cabinet Office

HM Treasury (2011) [The Magenta Book - Guidance for evaluation](#), London, HM Treasury

HM Treasury (2012) [Quality in policy impact evaluation: understanding the effects of policy from other influences](#) (supplementary Magenta Book guidance), London, HM Treasury

HM Treasury (2012) [Quality in qualitative evaluation: a framework for assessing research evidence](#) (supplementary Magenta Book guidance), London, HM Treasury

Kail, A. and Lumley, T. (2012) [Theory of Change The beginning of making a difference](#), London, New Philanthropy Capital

NPC & CLiNKS (unknown) [Using comparison group approaches to understand impact](#), London, CLiNKS

Puttick, R. and Ludlow, J. (2012) [Standards of Evidence for Impact Investing](#), London, Nesta

New Philanthropy Capital - <http://www.thinknpc.org/>