

Mental Health Improvement: Evidence and Practice

Guide 3: Getting results
evaluation guides

Mental Health Improvement: Evidence and Practice Evaluation Guides

Guide 3: Getting results

March 2005

Acknowledgements

The Mental Health Improvement: Evidence and Practice Evaluation Guides series was commissioned by Health Scotland, with funding from the National Programme for Improving Mental Health and Well-being. This work was developed by Lynne Friedli, Allyson McCollam, Margaret Maxwell and Amy Woodhouse of the Scottish Development Centre for Mental Health.

Scottish Development Centre for Mental Health
17a Graham Street, Edinburgh EH6 5QN
T 0131 555 5959
E sdcmh@sdcmh.org.uk
www.sdcmh.org.uk

Health Scotland would like to thank all those who participated in the consultation process during the development of the Guides:

- Aberdeen Healthy Living Network
- Argyll and Bute LHCC
- CHEX
- Choose Life National Implementation Support Team
- Clydebank LHCC
- Communities Scotland
- Dundee Healthy Living Initiative
- Glasgow Caledonian University
- Grampian Employment Trust
- Health Promotion Fife
- Health Promotion Greater Glasgow
- Health Promotion Tayside
- Inverclyde Choose Life
- Mental Health Foundation
- Moving into Health, West Lothian
- National Resource Centre for Ethnic Minority Health
- NIMHE
- NIMHE West Midlands
- NHS Health Scotland
- Research Unit in Health, Behaviour and Change, University of Edinburgh
- Scottish Executive Effective Interventions Unit
- Scottish Executive Mental Health Research Team
- Throughcare Support Team, North Lanarkshire
- University of Aberdeen

Contents

	Page no
1. Introduction	1
2. Planning an evaluation	5
3. Designing and conducting an evaluation	9
4. Implementing an evaluation	16
5. Building capacity and sustainability	21
6. References	22
Appendices	
Appendix A Glossary	24
Appendix B Evaluation websites	27
Appendix C A selection of evaluation tools used in current mental health improvement in Scotland	30
Appendix D Other useful resources	31

1.1 What is the purpose of this guide?

This is the third in a series of **Evaluation Guides**, which aim to encourage, support and improve standards in the evaluation of mental health improvement initiatives.

The guides are intended to help colleagues design evaluations that build on what is known about what works to improve mental health and that take account of the challenges of assessing the effectiveness of mental health improvement interventions.

The first four guides in the series are:

- *Guide 1: Evidence-based practice.* How can we use what we currently know to inform the design and delivery of interventions. This guide explores current debates about evidence of effectiveness and why they matter for mental health improvement. It also considers how the evidence base on mental health improvement can be used to inform the design of interventions and their evaluation.
- *Guide 2: Measuring success.* How can we develop indicators to gauge progress and assess the effectiveness of mental health improvement interventions. This guide covers the use of consultation to develop robust, valid and reliable indicators, examines the difference between mental illness indicators and mental health indicators and provides a useful source of indicators.
- *Guide 3: Getting results.* How can we plan and implement an evaluation. This guide gives an overview of the stages involved in planning and implementing an evaluation, and outlines the key issues for consideration. It also indicates sources of further, more detailed information on evaluation.
- *Guide 4: Making an impact.* How do we analyse and interpret the results from an evaluation and communicate the findings to key audiences. This guide discusses how to use the data gathered. It explores how evaluation can be used to inform practice and how the publication of results can add to the evidence base for mental health improvement.

Each guide contains a glossary.¹

¹ Terms in bold also appear in the glossary

The guides have been compiled as part of NHS Health Scotland's work to support evidence and practice in mental health improvement (<http://www.hebs.com/researchcentre/specialist/mhevidprog.cfm>) on behalf of the National Programme for Improving Mental Health and Well-being (www.wellontheweb.net) and complement other resources commissioned by Health Scotland and the Scottish Executive:

- *Mental health, mental well-being and mental health improvement: what do they mean? A practical guide to terms and definitions* (Scottish Executive, 2004, www.wellontheweb.net)
- *Mental Health Improvement: Evidence and Practice case studies* (NHS Health Scotland, 2004). A selection of case studies of current mental health improvement practice in Scotland. This resource provides 22 case examples from a range of sectors and settings of work that is evidence-based, follows good practice guidelines and gives indications of effectiveness. The evaluation guides cross-refer to these case study examples, where appropriate, for illustrative purposes.
<http://www.hebs.com/researchcentre/pdf/FinalReport200304RE041.pdf>
- *Mental Health and Well-being Indicators Project* (<http://www.phis.org.uk/info/mental.asp?p=bg>). In support of the National Programme for Improving Mental Health and Well-being, NHS Health Scotland is currently developing a set of public mental health indicators for Scotland. The indicators will provide a way of monitoring the state of mental health and well-being in Scotland at a national level. Although the indicators will be designed for use at a national level, some of them may be collected and applicable at a local level and will be helpful for those working locally in mental health improvement.

The guides are designed to strengthen evidence-based practice in mental health improvement and to support evidence from practice.

Strengthening evidence-based practice involves:

- increasing knowledge and awareness of the existing evidence base among practitioners and managers, i.e. what we know about what works in mental health improvement. A summary of some of the literature on evidence of effectiveness is available in *Mental Health Improvement: What Works?* (Scottish Executive, 2003, www.hebs.com/topics/mentalhealth)
- involving practitioners in producing guidance on evidence of effectiveness in the context of local needs and priorities, to ensure local relevance
- disseminating guidance on evidence in ways that are accessible and relevant to practitioners and that acknowledge barriers to implementing evidence-based practice
- building capacity, confidence, knowledge and expertise in working with the evidence base, to ensure that the planning and delivery of interventions are informed by an understanding of what works.



Supporting evidence from practice involves:

- enabling practitioners to evaluate interventions in order to inform their own practice and to guide local service development
- supporting the publication of local evaluations in peer-reviewed journals to add to our collective understanding of effective mental health improvement interventions and strengthen the evidence base
- finding ways to bring together practitioner know-how and expertise drawn from their experience of ‘what works’ with findings from the research literature.

1.2 Who are the guides for?

The guides are intended as a resource for colleagues across all sectors and settings. It is anticipated that they will be relevant to those working in a wide range of disciplines and services, both those with an explicit remit for mental health improvement and those for whom mental health improvement is an integral but implicit aspect of their work. The guides relate to areas of activity that are central to the responsibilities and interests of Community Planning Partnerships, Community Health Partnerships and multi-agency service planning groups for children and young people and for adults of all ages.

They have been developed in response to a clearly identified need among practitioners and service managers and programme managers for information and guidance on the evaluation of mental health improvement interventions. The guides therefore bring together information on evaluation theory and practice and a discussion of current debates and challenges in the field of mental health improvement, as well as pointers for practical application in designing and evaluating interventions. The series is not intended to be an evaluation manual – more detailed advice on evaluation for those who require it can be obtained from the resources listed in this guide and in Appendix B of *Guide 1: Evidence-based practice*.

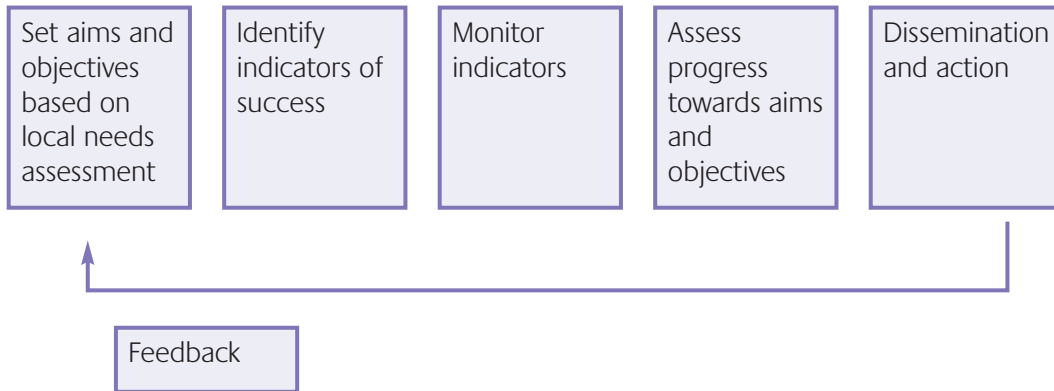
The Evaluation Process



Planning an evaluation

2.1 Planning for action

To be effective, evaluation forms part of a cycle of activity as summarised below:



Evaluation process in *Making it Happen – A guide to delivering mental health promotion*, Dept of Health, 2001.

Working out how you might evaluate an initiative or intervention when you are at the planning stage of setting up the initiative can be very helpful. It encourages you to address critical questions that are important for the success of any intervention, as well-being helpful in preparing for evaluation.

- *Defining the problem or issue*
What are the problems or issues that your programme is attempting to address or solve? What does previous research suggest are likely to be the outcomes of the interventions?
- *Needs assessment*
What are the needs of the community or target group your programme is intended to serve? Do the identified needs provide sufficient rationale to justify implementing the programme in the local context? (see also *Guide 1: Evidence-based practice*)
- *Expected processes, outputs, outcomes and impacts*
How is the intervention intended to work? What do you expect the programme to achieve? What would indicate that the desired outcomes had been achieved? Will the intended programme meet the needs of the community?
- *Assumptions*
How and why do you think the intended programme will work in this community or with this target group? How does this relate to the risk and protective factors for mental health, as explored in *Guide 1: Evidence-based practice*?

EXAMPLE Planning for action

Identified problem

Social isolation and lack of social contact are contributory factors for psychological distress in older people. The planned initiative involves strategies to include older people in the community more, and to facilitate social contact.

Needs in context

However, you may need to assess whether this applies to any or all sets of older people living in your community; older people living in urban and rural communities may have different perceptions and experiences, as may older people from minority ethnic communities. It may also be the case that there are specific issues within a community (such as fear of violence) that inhibit opportunities for social contact.

Implications for intervention

Setting up a social club, for example, without additional strategies to address these concerns (e.g. offering a pick-up service) may fail to meet the needs of local older people.

2.2 Involving key stakeholders

Having established the rationale or logic surrounding, and the need for, the programme and what it is hoped it will achieve, the next stage is to identify and involve the key **stakeholders**. These can include: those who use or might use the project; volunteers; project staff; those who manage and plan related local services; funders; and members of the local community itself.

The evaluation should consider the needs and interests of all key stakeholders (perhaps through representatives of stakeholder groups) and work with them to agree what the evaluation will seek to achieve and what outputs, outcomes and impacts the evaluation should assess. The early involvement of stakeholders encourages opportunity for their views and perspectives to be taken on board.

Stakeholders will have different interests and views and also hold different levels of power and influence over what the programme is to achieve and how. It may not be possible to meet everyone's criteria, but if everyone is aware at the outset of the realistic limits of what can be achieved within the timescale and resources available for the evaluation, there is less likely to be disappointment and disagreement when the evaluation results are reported.

Involving representatives of each stakeholder group will facilitate the evaluation process, by securing the co-operation of key groups in contributing to data collection processes. It can also help reduce anxiety among staff and service users/community members about the purpose of the evaluation. Having the involvement of stakeholders built in from the early stages is also valuable in ensuring that emerging issues and findings can be fed back to different groups. This is discussed further in *Guide 4: Making an impact*, in relation to reporting results and making use of findings from an evaluation.

Guide 2: Measuring success flags up the importance of involving stakeholders in identifying the nature of the problems, deciding on the intervention and selecting indicators to gauge success as a critical feature of mental health improvement evaluation.

2.3 Agreeing the goals, aims and objectives of the evaluation

This involves determining what questions the evaluation is seeking to answer. For example:

- what are the overall aims of the project?
- how will you know if they have been achieved?
- what activities will be undertaken to implement aims/achieve goals?
- what factors may help/hinder the achievement of aims/goals?

The evaluation questions (which must relate to the purpose and goals of the programme) are critical in determining what type of knowledge or information the evaluation should collect. However, different stakeholders will have different questions and it is important to explore these different perspectives to be able to build consensus on the focus for the evaluation and settle whether the evaluation is to be:

- concerned with the **process** of the programme – how it operates, what staff and service users think about the service, what their experiences have been and what could be improved
- **outcome** focused and therefore concentrate on collecting measures that can demonstrate change and measure **effectiveness**
- interested in issues of cost and **efficiency**
- comparative, looking at different groups or different interventions
- a snapshot in time or looking at trends over time
- **prospective** or **retrospective** – looking back or forward
- short term or long term.

Service funders may be more concerned with quantifying the **impact** of the programme in order to justify expenditure. Planners may want to know about the acceptability and effectiveness of a programme. Those involved directly as staff or as recipients of the intervention may be more interested in process and quality, to document what it is about the programme that people value. Outcome evaluations need to be able to provide some account of ‘what went on’ in the project, if the work of the project is to be replicated in other places. It may be that all or a combination of two or more of these aims/objectives of the evaluation are to be pursued.



Case study 22 (see *Mental Health Improvement: Evidence and Practice case studies* NHS Health Scotland, 2004) provides an example of an innovative project targeted at women with common mental health problems attending primary care. The evaluation of this demonstration project is focusing on both process and outcome to gauge its impact and to identify the salient features of how it works.

The focus, aims and objectives of the evaluation may be partly influenced by the funding for the evaluation. Evaluation funders may want to prioritise a particular aspect of the programme for evaluation.

Points for reflection

- Which stakeholders are not being heard in consultations on the evaluation? Why not?
- How can the evaluation best reflect different perspectives?
- What are the key questions that the evaluation should address?
- What are the main drivers behind the evaluation? To pilot an innovative service? To inform practice and improve service quality? To secure funding?

Designing and conducting an evaluation

3.1 Study design

Once the focus of the evaluation has been agreed, you need to consider what types of information, knowledge or data is required, from whom and how the information can best be obtained to address the key questions posed. This will shape the design of the evaluation. The design may also be constrained by what is feasible within the resources and timescale available.

The following information outlines some potential methods for evaluation.

Focus of evaluation	Questions of interest	Study design	Key considerations
Process	How people get access to the project. Who attends. Experiences of using the service.	Descriptive/observational. Snapshot in time.	Useful in piloting an initiative to assess its feasibility and acceptability.
Outcomes	Effectiveness What changes and impacts are achieved? For whom?	Comparative Comparing different groups at one point in time. Comparing different interventions. Measuring the same group at more than one point in time, e.g. before/during/after.	Requires steps to control for: <ul style="list-style-type: none"> • differences between groups • differences in the interventions received

If the focus is on the process of the intervention, then a more **descriptive/observational** study may be indicated to find out who uses the service, how often, and their experiences of the service. This design is most appropriate when piloting an initiative to assess its feasibility and acceptability. Once this has been assessed, it may be appropriate to conduct an evaluation of its **effectiveness** to examine **outcomes**.

To focus on outcomes, the evaluation design can assess the target group before and after the intervention or can compare two interventions (normally the new initiative with the status quo). A comparative design usually relies on controlling both the interventions so that they remain consistent throughout the project, and also controlling for other potential influences relating to the client group (e.g. age, gender, severity of problem).

If the client group (with the same characteristics) is randomly allocated to either intervention, then this constitutes an **experimental** or **randomised controlled trial** (RCT). This latter design allows you to be more certain that the intervention itself caused the results observed in the intervention group, as the status quo group acts as the 'control'.

Although RCTs are often seen as the ‘gold standard’ and have the greatest potential for ruling out bias in the study design, they are also complex, require large sample sizes, specialist research skills, and have additional ethical and cost implications. In addition many mental health improvement interventions do not readily lend themselves to evaluation using an experimental method of this type.

Guide 1: Evidence-based practice provides a critique of RCTs and considers the limitations of this method in evaluating mental health improvement interventions, primarily because of the need to take account of context and of the subjective experiences of those *involved in mental health improvement programmes*.

A programme or initiative may involve multiple sites, implementing interventions which are similar but not necessarily the same or consistent throughout the life of the project, in different local contexts. This provides an opportunity to examine the collective outcomes across these sites or to consider the specific contexts in which the intervention works well or less well. A comparative or RCT design may not be feasible or appropriate in these circumstances.

As discussed in *Guide 1: Evidence-based practice*, realistic evaluation provides a framework to investigate the context in which an intervention takes place and the mechanisms by which the intervention works, which in turn give rise to the observed outcomes. This makes it possible to address three key questions: what works, for whom, in what circumstances? (Pawson and Tilley, 1997).

It may be necessary to focus in some depth on a small number of intervention sites depending on the level of detail required, as this is often a resource-intensive exercise. A greater understanding of the programme (and what works in what circumstances) can be achieved using a smaller number of **case studies**. However, the selection of case study sites requires careful consideration (Yin, 2003).

Points for reflection:

- Create a flexible and responsive design, collect and analyse information from multiple perspectives.
- Always return to your evaluation questions.
- Be sure that the design you have chosen is the most appropriate for the intervention to be evaluated and for the questions that you want to consider.

3.2 Types of data and data collection

All evaluations, regardless of design, can utilise both **quantitative** and **qualitative** methods for collecting data.

Quantitative methods generally involve collecting and collating numerical data which can include:

- basic demographic data to give a profile of those who use a service (e.g. by age, gender, ethnicity, community they are from)
- activity data to indicate the number of clients using the service, number of visits or sessions per client
- rating or assessment of problems/needs, obtained by using instruments such as questionnaires or standard assessment tools.

Guide 2: Measuring success provides some examples and references for tools that can be used to rate mental health as well as those commonly used to assess levels of mental ill health.

It is worth noting that many standard tools are subject to copyright and there may be fees associated with their use, and with purchasing copies of instruments, manuals or handbooks.

If your evaluation is likely to require the collection of a considerable volume of quantitative data over and above routine statistics, it is worth seeking the advice of a statistician at an early stage. This can help ensure you have made adequate plans to recruit a large enough sample to allow meaningful analysis. It is also important to plan in advance how you will approach data analysis as this can influence how you go about collecting and recording data.

Qualitative methods collect data that reflect people's perceptions, experiences or attitudes. This generally entails using methods such as interviews, focus groups or observation. It can also include written responses to open-ended questions in questionnaires, for example: How did you find out about this service? What for you have been the main benefits of attending this parenting group?

In addition other media such as minutes of meetings, work logs, diaries, pictures, photos and videos can be rich sources of data. It is important to ensure that the methods selected maximise the inclusion and participation of those who use the service and are culturally appropriate.

Case studies or personal narratives can, with the consent of the person or people involved, provide a considerable amount of in-depth information drawing together material collected through interviews, records, pictures, etc.



Case study 14 (see *Mental Health Improvement: Evidence and Practice case studies* NHS Health Scotland, 2004) illustrates how methods of evaluation can be adapted for work with primary school age children. In this project, which works with children who have experienced abuse, the project evaluation employed a combination of methods including: questionnaires designed for children to complete to rate their feelings and emotions using pictures; videos to gather feedback from children involved in the project; and formal feedback from professionals in other services.

Each project will have a certain amount of both qualitative and quantitative data that it collects as part its routine work or that was gathered to inform its planning and development. The evaluation can often build on this. For example, collecting client feedback can be incorporated into the routine work of a project, and systems can be established to gather this information together periodically.

What is essential is that the evaluation questions determine which type of data collection methods are most appropriate.

EXAMPLES of data collection methods

In a project aiming to reduce social isolation in older people, some of the questions for the evaluation might be:

1. What has been the uptake of the service by older people, and which type of older people are more likely to use the service?
2. What have been the experiences of older people using the social club and how do they feel it has impacted on their lives?
3. Has the provision of a community social club for older people reduced social isolation or impacted otherwise on mental health and well-being?

Each of these questions will require different types of data collection methods

Evaluation question	Method of data collection
What has been the uptake of the service by older people?	Routine data collected by the service on who uses it.
Which type of older people are more likely to use the service?	Census data on profile of local population of older people to identify groups who may not be attending.
What have been the experiences of older people using the social club and how do they feel it has impacted on their lives?	User feedback. This could be in the form of a structured or a semi-structured questionnaire , more in-depth qualitative interviews, focus groups or diaries kept over an agreed period to record activities and views.

Evaluation question	Method of data collection
Has the provision of a community social club for older people reduced social isolation or impacted in other ways on their mental health and well-being?	Before and after measurement of social isolation and/or other measure(s) of mental health and well-being of either the community in which the initiative is based or the individuals who utilise the service.

Boynton (2004) provides helpful information on selecting, designing and developing questionnaires. Boynton *et al* (2004) consider how to overcome problems with accessing disempowered and socially excluded groups, cross-cultural issues and participants whose physical or mental health may interfere with their ability to complete a questionnaire.

In the example above, questionnaire methods would assess the perceived benefits experienced by older people and levels of satisfaction with the service. Questionnaires can be completed in a variety of ways: by an interviewer who meets the person face to face; by self-completion on paper or by email; by a telephone interviewer.

Qualitative interviews and focus groups allow for more in-depth evaluation of a service and how it meets needs, what works and what could be improved, and factors that might hinder the potential of the service.

Other possible methods are the **nominal group technique** (useful if individuals can be brought together face-to-face) and the **Delphi technique** (normally using a series of mailed questionnaires). Both are helpful if you want to pose a specific question to a group of individuals to generate a list of responses which the group then ranks in order of importance.

3.3 Sampling

When identifying how to collect data, initial questions to address are:

- whether to include all participants/clients/stakeholder groups or a sample
- whether the evaluation should take account of the views of people who do not access the service/initiative; this might be important if you want to determine how the service might be more inclusive
- whether to follow up people who drop out of the intervention; this might be important to understand how the evaluation might need to be modified or adapted to meet the needs of different groups.

Sampling is the process of selecting units (e.g. individuals, families/households, organisations) from the population you are interested in, so that, by studying the sample, you can generalise the findings to the wider population from which they were chosen.

Sampling methods are classified as either **probability** or **non-probability**. In probability samples, each member of the population has a known probability of being selected. This method means that it is possible to indicate how results relate to the wider population. In non-probability sampling, members are selected from the population in some non-random manner and the degree to which the sample differs from the population remains unknown. It is therefore not possible to indicate how the findings can be applied more widely.

The population to which you have access in selecting a sample is normally called the sampling frame. Identifying a sample from the sampling frame is where **bias** can occur. You may want to consider the following questions:

- Is the sampling frame complete and accurate (e.g. is it up-to-date, are records missing)?
- Are some sets of people easier to access than others?
- Do some sets of people respond/participate more readily than others?
- Are some sets of people more likely to drop out than others?

Points for reflection

- Developing a sampling strategy that effectively takes these considerations on board will be more likely to produce 'generalisable' results.
- Be clear in reporting your results how you selected your sample and what the consequences of that may be for the interpretation of your findings.

3.4 Timing of data collection

You will also want to consider when and how often data collection should be undertaken: whether this is before and after the intervention; on-going throughout the project; or whether this will be conducted as a follow-up exercise.

3.5 Levels of data

There are different levels of data that can be accumulated within an evaluation:

Inputs: detailing the inputs or resources required in order to set up and deliver an intervention is important for those who want to replicate the intervention elsewhere. This can include staff numbers and skills, premises, equipment, opening times or accessibility of resource. It also includes economic inputs, e.g. set-up costs and running costs.

Processes: how the intervention is delivered, its quality and reach and subjective experiences of those who deliver and receive it and/or the wider community and other services and resources, relationships between those involved in the service, what arrangements are made to ‘recruit’ people to the project and to prepare and support them when they move on.

Outputs: the ‘activities’ undertaken or the service delivered. In relation to the social club for older people, this would include: the number of older people who attended; the number of times each has attended; the types of people the service is reaching (men and women or mostly women?); the number of hobby groups established (e.g. bingo, lunch club).

Impacts: consequences of the intervention for staff or service users or the wider community. Impacts differ from ‘outcomes’ in providing a more qualitative assessment of the overall effects of the intervention and whether it has met its objectives, rather than only the direct immediate effect of the intervention, as assessed by outcomes. For example, the opportunity to participate in the social club may have had a wider impact on older people feeling more able to organise themselves to lobby the local council for more street lighting. This impact would not have been anticipated or measured within the data collected to assess the immediate and direct effect of the intervention on participants.

Outcomes: what is defined as an outcome will depend on the aims and objectives of the intervention. A distinction is often made between ‘hard’ (objective) outcomes and ‘soft’ (subjective) outcomes. However, as stated in *Guide 1: Evidence-based practice*, this distinction becomes less important in the light of the growing evidence that how people feel has a direct bearing on their objectively measured health status.

There are many standard tools that are available for use as ‘before and after’ measures of change in mental health (Bowling, 2001; 2004; Mauthner and Platt, 1998; Stewart-Brown, 2002). A more detailed discussion of indicators and measures of change relevant to mental health improvement is provided in *Guide 2: Measuring success*.

Points for reflection

- Will the data be able to answer the evaluation questions?
- Are there other (cheaper, simpler) ways of collecting the same information?
- How can the evaluation build on data which the project collects routinely?
- Are the data collection methods biased in any way, i.e. in relation to the sample or populations that they draw on, the questions that they ask, or who asks the questions?

4

Implementing an evaluation

Guide 1: Evidence-based practice flagged up some of the general issues associated with evaluation of mental health improvement interventions. Below are some further practical considerations that require attention when you reach the stage of implementing your evaluation.

4.1 Piloting

It is worth making provision to pilot the data collection methods you propose to use, to test their feasibility. Many studies have experienced delays if they cannot recruit subjects willing to take part in in-depth interviews, if they encounter problems with systematically collating the routine output data, or if it proves difficult to achieve the necessary sample sizes required to demonstrate change.

Such problems and delays can be costly and can have a detrimental effect on the ability of the evaluation to answer the questions it set out to address.

4.2 Timetable

It is important to create a realistic timetable for the intervention and the evaluation and to detail when data collection will take place through the course of the project. For example, is it important to collect baseline data prior to the intervention starting? The intervention may require a run-in period to become established before it is likely to see sufficient numbers of clients coming through its doors. It could be a waste of evaluation resources if a full-time evaluator were on board during this period.

4.3 Costs

Conducting an evaluation requires time and money. Even an internal evaluation will have resource implications, sometimes utilising highly skilled staff, and for this reason alone it may be advisable to hire an independent evaluator (see below for further discussion of internal and external evaluation).

Various sources recommend that programmes or projects allocate a proportion of their budget for evaluation of between 5 and 10 per cent (e.g. Kellogg Foundation Evaluation Handbook – see Appendix 2). The items to consider in costing an evaluation are: staff time and expertise to perform some evaluation tasks; external consultants; travel; stationery, postage, phone calls; printing and duplication of questionnaires and reports; meeting room hire for focus groups; costs associated with data collection instruments; supplies and equipment.

4.4 Ethical issues

Research being carried out using NHS patients and staff is subject to ethical approval. Activities considered to be ‘audit’ are not, and therefore internal auditing of client satisfaction with services and monitoring of client outcomes need not necessarily require permission if this is conducted internally. However, there may still be issues surrounding informed consent in how you later use the information that service users have provided. An external evaluation team conducting research involving NHS patients and staff is required to have ethical approval and the informed consent of patient and, in some areas, staff groups.

In evaluations involving the clients of other service providers, including local authorities and voluntary sector organisations, it is good practice to secure the approval of the organisation involved as well as ensuring that individuals are asked to give informed consent prior to participation in the study.

In addition to informed consent, the key ethical issues in conducting an evaluation relate to confidentiality, promoting respect, acknowledging the contribution made by study participants and their entitlement to receive feedback on the results of the evaluation and on any changes or developments that ensue.

For further information on ethical approval and access to other helpful sites and codes of conduct see <http://www.corec.org.uk/>. Additional resources are listed in Appendix D.

4.5 Incentives

An additional ethical issue to consider is whether to provide incentives to participants in the evaluation. This requires careful thought if the client group involved is young or vulnerable (such as homeless people). There is a distinction between an incentive and an ‘inducement’, which may have the effect of persuading people to take part or be construed as presenting an offer that they cannot afford to refuse.

Incentives can have a positive impact on participation rates among harder to engage populations (such as those in more deprived communities). One approach can be to reward people for their time, but not offer incentives in recruiting to a study – in which case you would not tell people in advance that they would receive some form of recompense. Non-monetary incentives to subjects can include: token gifts such as pens or key rings; participation in a raffle or lottery; or a donation to a charity. Monetary incentives can include a direct cash payment or a gift voucher. Where study participants are unwaged, it is important to remember that payment can have an impact on benefit entitlements and it is worth seeking advice on this.

It may be reasonable to expect to reimburse respondents for any travel expenses they may incur or for childcare or other care costs (particularly important if you want to access mothers/carers).

Generally (although not consistently), non-monetary incentives have been shown to increase responses and participation rates. However, both monetary and non-monetary incentives are effective only when enclosed with the survey. The promise of an incentive for a returned questionnaire has not been shown to be effective in increasing response.

4.6 Legal issues

Legal issues to consider relate to data protection and responsibility for negligent/non-negligent harm resulting from the intervention. This latter issue may require clarification through the identification of a research 'sponsor' who is willing to take on the legal responsibility and liability attached.

The Data Protection Act 1998 provides protection for personal data. This includes information which, by itself or in conjunction with other easily obtainable information, can identify a specific person. Data must be securely held, accurate and relevant. Organisations must notify the Data Protection Register in relation to data held on computer (see www.dpr.gov.uk and www.data.protection.gov.uk).

Further information on data protection is listed in Appendix D.

4.7 Risk minimisation strategies

Work involving young and/or vulnerable groups may require special considerations in order to minimise the risk to these research subjects. For example, in relation to children and young people, permission may have to be sought from parents or guardians to allow participation in research. It may also be advisable to have a responsible adult (other than the researcher) close by. The degree of maturity of a child must be taken into account particularly in relation to the topics being covered, as young people should not be worried, confused or misled. Interviewers should also be screened for their suitability in working with young people and adequately trained for research involving young people. It would also be advisable to have support or advice available for young people during or following their participation.

Similar and other risk minimisation strategies may apply to other vulnerable groups. See <http://www.esomar.org/main.php?a=2&p=79> and also www.research.qut.edu.au/oresearch/policyandpro/ethics/downloads/16CHILDREN.rtf

4.8 Internal or external evaluation?

The critical skills of an effective evaluator include the ability to listen, negotiate, bring together multiple perspectives, analyze the specific situation, and assist in developing a design with the evaluation team that will lead to the most useful and important information and final products. (Kellogg Foundation Evaluation Handbook, see Appendix B)

A crucial decision will be whether to conduct the evaluation internally or to enlist the services of an external, independent evaluator (or team). A third option is to conduct the evaluation internally with an external consultant to advise and assist with the technical aspects of the evaluation. For some projects an external evaluation is a requirement of their funding.

The decision will depend on the capacity and research expertise available and required to design and implement the evaluation and to collect and analyse the type of data required, in an unbiased manner. It will also depend on the availability of additional funds to cover the costs of appointing external evaluators. It is important to remember that who asks the questions in an evaluation – someone closely associated with the project or intervention, or an independent neutral party – can influence the responses received. Care must be taken to avoid bias when the evaluation is conducted internally.

Universities and Colleges will be able to advise on evaluation, as will many private and not-for-profit organisations with experience in research or management consultancy.



Case study 6 (see *Mental Health Improvement: Evidence and Practice case studies* NHS Health Scotland, 2004) illustrates how internal and external independent evaluation can combine to provide different perspectives at different stages in the life cycle of an initiative.

For projects that opt for internal/self-evaluation, there are a number of tools and resources available to assist: see LEAP for Health (Appendix C); ASH Scotland (Appendix B); Effective Interventions Unit (Appendix B).

4.9 Management arrangements

Be clear about who has overall responsibility for the evaluation of the project or initiative. Who will ensure that the evaluator(s) are doing what they said they would and within an agreed timescale? Who will ensure that staff involved in the initiative are co-operating with the evaluator(s)? Are there clear lines of communication throughout the course of the evaluation? Are there key points at which you want to get feedback from the evaluation to inform the planning and development of the project? It may be helpful to set up an advisory group that will meet regularly with the evaluation team in order to ensure the smooth running of the project.

Points for reflection

- Consider the relationship between the evaluator and the primary stakeholders or internal evaluation team – should this be distant or interactive?
- Who will control and manage the evaluation process?
- What are the key ethical considerations in conducting the evaluation?
- How actively involved should key staff and stakeholders be in the evaluation process?
- What difference would it make to the evaluation if it were undertaken by an impartial external organisation – what would be gained, what might be lost?

Building capacity and sustainability

The evaluation process should not be a 'one off'. If data collection systems are put in place for the initial evaluation, then what would it cost to maintain these systems as part of routine practice? Changes to the initiative may have been made as a result of the initial evaluation and it is worth monitoring these. Key stakeholders or staff involved in the original phase of the initiative may leave and this may have an impact on the future direction or continuing success of an initiative. Building evaluation capacity and sustainability as part of the on-going initiative will allow for continuing monitoring of success or highlight problems or issues that require action. Ideas about making use of an evaluation are explored in *Guide 4: Making an impact*.

References

- Bowling A (2004) *Measuring Health: A review of quality of life measurement scales*. Open University Press, Milton Keynes.
- Bowling A (2001) *Measuring Disease: A review of disease specific quality of life measurement scales*. Open University Press, Buckingham.
- Boynton P (2004) Hands-on guide to questionnaire research: Administering, analysing, and reporting your questionnaire. *British Medical Journal*, 328:1372–5.
<http://bmj.com/cgi/content/full/328/7452/1372?etoc>
- Boynton P *et al* (2004) Reaching beyond the white middle classes. *British Medical Journal*, 328:1433–36. <http://bmj.com/cgi/content/full/328/7453/1433?etoc>
- Department of Health (2001) *Making it Happen: A guide to developing mental health promotion*. HMSO, London.
- Fetterman DM, Kaftarian SJ and Wandersman A (eds) (1996) *Empowerment Evaluation: Knowledge and tools for self-assessment and accountability*. Sage Publications, Thousand Oaks, CA.
- Fink A and Kosecoff J (1989) *How to Conduct Surveys: A step-by-step guide*. Sage Publications, Newbury Park, CA.
- Mauthner N and Platt S (1998) *Selective Literature Review of Measures of Mental Health and Emotional Well-being*. Health Development Agency, London.
- NHS Health Scotland (2004) *Mental Health Improvement: Evidence and Practice case studies*. NHS Health Scotland, Edinburgh.
- Pawson R and Tilley N (1997) *Realistic Evaluation*. Sage, London.
- Scottish Executive (2003) *Mental Health Improvement: What works?* A briefing for the Scottish Executive. Scottish Enterprise, Edinburgh.
<http://www.hebs.scot.nhs.uk/topics/topiccontents.cfm?TxtTCode=1395&TA=topicitles&newsnav=1&topic=mental>
- Scottish Executive (2004) *Mental Health, Mental Well-being and Mental Health Improvement: What do they mean? A practical guide to terms and definitions*. Scottish Executive, Edinburgh.
- Silverman D (1993) *Interpreting Qualitative Data*. Sage, London.

Stewart-Brown S (2002) Measuring the parts most measures do not reach: a necessity for evaluation in mental health promotion *Journal of Mental Health Promotion*, 1(2):4–9.

Yin RK (2003) *Case Study Research: Design and methods*. Sage, Thousand Oaks, CA.

Appendix A Glossary

Bias Any factor that distorts the nature of an event or observation. In research, a bias is any systematic factor other than the intervention that affects the difference in the outcomes of a treatment group and a control group. Randomisation is used to decrease this form of bias.

Bias also refers to a viewpoint that would affect someone's interpretation of a problem. Double blinding is a technique used to decrease this type of bias, e.g. where neither the patient nor the doctor know whether a pill is a placebo (related categories: **randomised controlled trial**). Bias can also occur in analysis of responses and in reporting of a study and it is therefore important in reporting a study to give full information about sample selection (including who declined to take part or dropped out of the study), data collection methods and the results that led to any conclusions drawn.

Case study In-depth analysis and systematic description of one or several groups of people in similar settings to promote a detailed understanding of their circumstances.

Delphi technique (see also **nominal group technique**) A method to ask a group of people a common question or set of questions and involve them in ranking the responses elicited. The Delphi technique does this through a written medium and by post.

Descriptive/observational study A study in which the investigators do not randomise participants to treatment and control groups, but only observe those who are (and sometimes those who are not) exposed to the intervention, and interpret the outcomes. Observational studies are often used to interpret research with pre-school children and other groups that are not able to take part in research methods that involve reading or talking.

Effectiveness The extent to which an intervention does people more good than harm. An effective treatment or intervention is effective in real life circumstances, not just an ideal situation. Effectiveness asks: are the intended impacts and outcomes being used to good effect?

Efficiency Whether resources being used to good effect.

Experimental See **randomised controlled trial**.

Impact The effect of the intervention and whether it has met its objectives.

Nominal group technique Similar to the **Delphi technique** (see above) but brings together to work face to face.

Non-probability sampling See **Probability sampling**.

Outcomes The effect or effects of an intervention, which might be measured in a variety of ways, for example health outcomes, economic outcomes.

Process. How the intervention is delivered, its quality and reach and the subjective experiences of those who deliver and receive it and/or the wider community.

Probability sampling This is a method to identify a sample for study that allows the sampling error to be calculated, i.e. the degree to which a sample might differ from the wider population from which it is drawn. When inferring to the population, results are reported, plus or minus the sampling error. Probability methods include random sampling, systematic sampling and stratified sampling. In **non-probability sampling**, a sample is selected from the population in some non-random manner, such as convenience/ purposive sampling, judgement sampling, quota sampling and snowball sampling. In this method of sampling, the degree to which the sample differs from the population remains unknown and this can affect how the results can be applied more widely.

Prospective study One in which the investigators design the study in advance. People are then recruited and studied according to the study's criteria. Some types of studies, such as **randomised controlled trials** are always prospective. In a prospective study, investigators do not know what the outcomes will be when they undertake the study (contrast with **retrospective study**).

Qualitative research A systematic, subjective approach used to describe life's experiences and give them meaning... conducted to describe and promote understanding of those experiences.

Quantitative research Systematic collection and analysis of numerical data to describe patterns and trends.

Randomised trial A trial of a treatment or intervention which does not include a control or comparison element.

Randomised controlled trial A trial of a treatment or intervention in which participants are randomly (i.e. by chance) assigned to two groups, one of which receives the intervention being tested (the experimental group) and one which receives no treatment, 'usual' or 'standard' treatment or a placebo (the control or comparison group). Both groups are followed up to see what effect, if any, the intervention has produced.

Quasi-experimental or **randomised trials** are when allocation is not by chance, e.g. using date of birth.

Retrospective study A study in which investigators select groups of patients that have already been treated and analyse data 'after the event'. They are considered less reliable than prospective studies (contrast with **prospective study**).

Semi-structured questionnaire Questionnaire design which includes open questions (What are your views about...?) as well as closed questions (Do you agree or disagree with the following statement?). Semi-structured questionnaires provide a menu of topics to be covered but tend to allow the interviewer discretion with regard to how to ask the questions and in what order.

Stakeholder Any person or group who has an interest in the project being evaluated or in the results of the evaluation, including funders, strategic planners of services, project staff, service users or clients of the service and other organisations/people for whom the programme is likely to have an impact. The evaluators are also stakeholders in the evaluation.

Structured questionnaire Questionnaire design which includes questions that have fixed response categories/closed questions (see **semi-structured questionnaire**). Structured questionnaires provide questions which must be posed as written and in a predetermined order.

User feedback Qualitative methodology which utilises participant experience of a service or intervention as a valid measure of impact.

Appendix B Evaluation websites

Evaluation UK

<http://www.evaluation.co.uk/>

A Guide to Programme Evaluation

<http://www.evaluation.co.uk/library/how/progs.htm>

A Guide to Research Evaluation

<http://www.evaluation.co.uk/library/how/researchprogs.htm>

So Did it Work?

<http://www.evaluation.co.uk/library/how/sodiditwork.pdf>

UK Evaluation Society

http://www.evaluation.org.uk/Pub_library/Good_Practice.htm

NHS Health Scotland

Research Centre R&E Toolbox

<http://www.hebs.scot.nhs.uk/research>

Scottish Community Development Centre

LEAP for Health

http://www.scdc.org.uk/leap_index.htm

Effective Interventions Unit Evaluation Guides for evaluating services for drug users

<http://www.drugmisuse.isdscotland.org/goodpractice/EIU>

See Guides 3 and 4 in particular

ASH Scotland

The Evaluation Journey: An Evaluation Resource Pack for Community Groups

<http://www.ashscotland.org.uk/inequalities/tobacco.html#eval>

Communities Scotland

SCARF jargon buster

http://www.communitiesscotland.gov.uk/Web/Site/Whatwedo/SCARF_Glossary.asp

Europa/European Commission

http://europa.eu.int/comm/budget/evaluation/keydocuments_en.htm

Ex Ante Evaluation: A Practical Guide for Preparing Proposals for Expenditure Programmes

http://europa.eu.int/comm/budget/evaluation/pdf/ex_ante_guide_en.pdf

Evaluating EU Activities: A Practical Guide for the Commission Services

http://europa.eu.int/comm/budget/evaluation/pdf/evaluation_guide_final.pdf

Evaluating EU Expenditure Programmes: A Guide: Ex post and intermediate evaluation

http://europa.eu.int/comm/budget/evaluation/pdf/guide_en.pdf

Good Practice Guidelines for the Management of the Evaluation Function

http://europa.eu.int/comm/budget/evaluation/pdf/goodpractice_en.pdf

European Evaluation Society

www.europeanevaluation.org

WHO Evaluation workbooks

www.who.int/substance_abuse/pubs Psychoactive drugs.html

Learning Technology Dissemination Initiative

<http://www.icbl.hw.ac.uk/ltidi/index.html>

Evaluation Cookbook

<http://www.icbl.hw.ac.uk/ltidi/cookbook/contents.html>

Charities Evaluation Services

<http://www.ces-vol.org.uk/>

First Steps in Monitoring and Evaluation

http://www.ces-vol.org.uk/files/Monitoring_evaluation.pdf

MandE News

www.mande.co.uk

WK Kellogg Foundation

<http://www.wkkf.org>

WK Kellogg Foundation Evaluation Handbook

<http://www.wkkf.org/Pubs/Tools/Evaluation/Pub770.pdf>

American Evaluation Association

<http://www.eval.org>

UNICEF

<http://www.unicef.org/reseval/>

Policy Hub

<http://www.policyhub.gov.uk/publications/index.asp#eval>

Lancaster University – LTSM National Evaluation Project

<http://www.lancs.ac.uk/fss/projects/edres/ltsn-eval/>

Resources for Methods in Evaluation and Social Research

<http://gsociology.icaap.org/methods/>

Social Science Information Gateway

<http://www.sosig.ac.uk/>

Bill Trochim's Centre for Social Research Methods

<http://trochim.human.cornell.edu/>

BJA Evaluation Website

<http://www.bja.evaluationwebsite.org/html/roadmap/index.html>

Enterprise Development Impact Assessment Information Service

<http://www.enterprise-impact.org.uk/>

European Monitoring Centre for Drugs and Drugs Addiction

<http://www.emcdda.eu.int>

Evaluation Instruments Bank

http://eibdata.emcdda.eu.int/databases_eib.shtml

Bob Williams website

<http://users.actrix.co.nz/bobwill/>

A public service of StatPac Inc. Designing surveys and questionnaires

<http://www.statpac.com/surveys/index.htm#toc>

Appendix C A selection of evaluation tools used in current mental health improvement in Scotland

LEAP and LEAP for Health

www.scdc.org.uk/leap_index.htm

Rickter Scale

www.rickterscale.com/0101.htm

CORE (Clinical Outcomes Routine Evaluation)

<http://www.coreims.co.uk/>

How Good is Our School

www.hmie.gov.uk/documents/publication/cd/frames.htm

Step It Up

www.youthlink.co.uk/upload/stepitup-report.pdf

<http://stepitup.youngscot.org/>

Paths to Health

<http://www.pathsforall.org.uk/pth/index.shtml>

Active for Life (assessment tools)

<http://www.active-for-life.com/>

Eyberg Child Behavior Inventor

<http://www.parinc.com/product.cfm?ProductID=95>

Appendix D Other useful resources

EPI INFO (a public domain software package) <http://www.cdc.gov/epiinfo/about.htm>

The Cochrane Collaboration

<http://www.cochrane.org/cochrane/cngloss.htm>

Centre for Evidence-Based Medicine

www.cebm.net

Mental Health Promotion Toolkit – A practical resource for community initiatives

http://www.cmha.ca/mh_toolkit/intro/

NHS Centre for Reviews and Dissemination (2001) *Getting evidence into practice. Effective Health Care*, 5(1)

www.york.ac.uk/inst/crd/ehc51

NHS Health Scotland Guide to Planning Health Events

www.hebs.scot.nhs.uk/cleardirections/

www.who.dk/eprise/main/WHO/Progs/HEN/Home

access to sources of evidence including case studies, etc. for public health

Marks DF (2002) *Perspectives on Evidence-based Practice*. Health Development Agency, London.

www.hda.nhs.uk

Meyrick J and Sinkler P (1999) *An Evaluation Resource for Healthy Living Centres*. HEA.

MRC advice on developing and evaluating RCTs for complex interventions to improve health is at http://www.mrc.ac.uk/complex_packages.html

Scottish Executive (2004) *Mental Health, Mental Well-being and Mental Health Improvement: What do they mean? A practical guide to terms and definitions*.

www.wellontheweb.net

Resources related to ethics and codes of conduct for research

Economic and Social Research Council

<http://www.esrc.ac.uk/resfund.htm>

The ESRC's funding regulations including not only details required under ESRC funding but also useful guidance on research ethics and confidentiality.

Social Services Research Group

<http://www.ssrp.demon.co.uk>

The Social Services Research Group website, including Guidelines for Good Practice in Research and Collaborative Research.

Royal College of Paediatrics and Child Health Guidelines for the Ethical Conduct of

Medical Research Involving Children (prepared by the Ethics Advisory Committee August 1992) *Archives of Diseases in Childhood* 2000, 82:177–82

<http://www.rcpch.ac.uk>

Royal College of Nursing

<http://www.rcn.org.uk/library/library.htm>

Research ethics guidance for nurses involved in research or any investigative project involving human subjects.

The MRC has published guidance on the ethical conduct of research on the mentally incapacitated

http://www.mrc.ac.uk/ethics_c.html

HERO The Higher Educational and Research Opportunities site provides a valuable set of links targeted to new researchers.

http://www.hero.ac.uk/research/good_practice_for_new_researchers 230.cfm

Good Research Practice Guidelines (issued in 2000)

<http://www.mrc.ac.uk>

Institute for Healthcare Management

<http://www.ihm.org.uk/>

The largest UK professional body for managers working in healthcare and health services.

The Institute for Social Research

<http://www.soc.surrey.ac.uk/>

Includes consultancy services in the design and implementation of research, including evaluation.

The Help for Health Trust

<http://www.hfht.org/ConsumersinNHSResearch/index.htm>

Consumers in NHS Research aims to ensure that consumer involvement in R&D in the NHS improves the way that research is prioritised, commissioned, undertaken and disseminated.

The Social Research Association

<http://www.the-sra.org.uk/index2.htm>

Includes guidelines on ethics and safety of interviewers.

The Market Research Society's (MRS) Code of Conduct

<http://www.marketresearch.org.uk/>

Site also includes MRS's best practice guides including employee research, qualitative research, data collection and research with children and young people.

Data protection

Caldicott Report and Related Guidance (1997/98/99) Using confidential patient information in the modern NHS.

<http://www.dh.gov.uk>

Department of Health (1999) *For the Record: Managing records in NHS Trusts and Health Authorities*. Health Service Circular, March 1999/053.

Department of Health (1996) *The Protection and Use of Patient Information*.
Department of Health.

