Because it's worth it

Also available in the *Contemporary research issues* series

Making the most of it: Economic evaluation in the social welfare field
Tom Sefton, Sarah Byford, David McDaid, John Hills and Martin Knapp

Because it's worth it

A practical guide to conducting economic evaluations in the social welfare field

Sarah Byford, David McDaid and Tom Sefton

JR
JOSEPH
ROWNTREE
FOUNDATION

The **Joseph Rowntree Foundation** has supported this project as part of its programme of research and innovative development projects, which it hopes will be of value to policy makers, practitioners and service users. The facts presented and views expressed in this report are, however, those of the authors and not necessarily those of the Foundation.

This project was carried out jointly by the London School of Economics (LSE) and the Institute of Psychiatry (IOP). Sarah Byford is a Senior Lecturer in the Centre for the Economics of Mental Health at the IOP, David McDaid is a Research Officer in LSE Health and Social Care, and Tom Sefton is a Research Officer in the ESRC Research Centre for Analysis of Social Exclusion (CASE) at the LSE.

Joseph Rowntree Foundation
The Homestead
40 Water End
York YO30 6WP
Website: www.jrf.org.uk

First published 2003 by the Joseph Rowntree Foundation

ISBN 1 85935 111 5 (paperback)
ISBN 1 85935 112 3 (pdf: available at www.jrf.org.uk)

A CIP catalogue record for this report is available from the British Library.

Cover design by Adkins Design

Prepared and printed by:
York Publishing Services Ltd
64 Hallfield Road
Layerthorpe
York YO31 7ZQ
Tel: 01904 430033; Fax: 01904 430868; Website: www.yps-publishing.co.uk

Further copies of this report, or any other JRF publication, can be obtained either from the JRF website (www.jrf.org.uk/bookshop/) or from our distributor, York Publishing Services Ltd, at the above address.

Contents

Acknowledgements

This project was carried out jointly by the ESRC Research Centre for Analysis of Social Exclusion (CASE) at the London School of Economics (LSE), LSE Health and Social Care, and the Centre for the Economics of Mental Health at the Institute of Psychiatry. It was initiated and funded by the Joseph Rowntree Foundation. The research team is grateful for this support and for the help and advice of Janet Lewis, who also chaired the Advisory Group Meetings, and Anne Harrop.

The authors would like to thank the rest of the research team for their significant contributions to this report – Martin Knapp, John Hills, Jennifer Beecham, Howard Glennerster, Ann Netten and Franco Sassi – and the members of the external Advisory Group – Martin Buxton, James Foreman-Peck, Sally Holtermann, Robert Maxwell, Tim Newburn, David Pearce, Becky Sandhu and Elliot Stern.

1 Introduction

Purpose of the report

This report is a step-by-step guide to the practical application of economic evaluation within the social welfare field. It is the second volume of a report resulting from a two-year project to promote better understanding and use of economic evaluation in the social welfare field. Volume I (Sefton *et al*., 2002) focuses on the application of economic approaches to a social welfare context, while this report provides more detailed technical advice on how some of these approaches might be applied in practice. In addition, it discusses some of the issues to be considered and choices to be made when undertaking an economic evaluation in practice. Inevitably, there is some overlap between the two volumes to enable each to stand alone and to keep cross-referencing to a minimum. Readers are referred to Volume I, however, where more detailed information that is not central to the understanding of this report is available.

This report is aimed primarily at non-economists involved in the evaluation of social welfare programmes, but also at economists who may be unfamiliar with the social welfare field. It is intended to provide a greater understanding of the methods of economic evaluation and the practical processes involved. It should not be seen as an attempt to provide either an exhaustive or definitive guide to economic evaluation, since this is a broad and complex area. Instead, the report aims to provide the reader with an understanding of the basic building blocks involved and enough knowledge to put into place the beginnings of a successful economic evaluation. Throughout the report, references to further literature are provided for those readers who wish to explore any area of economic evaluation in more detail and readers are advised to consult with economists before attempting a full economic evaluation alone.

Economic evaluation

Economic evaluation, summarised in Box 1, is defined here as the systematic attempt to identify, measure and compare the costs and outcomes of alternative interventions. The main purpose is to assess the value for money of alternative services, in order to support decisions of which services should be funded and which should not. The overarching aim is to maximise the benefits to service users by funding those services that generate the greatest outcomes for the money available.

Economic evaluation is built on the notion of scarcity. Resources are limited and often insufficient to meet total needs, thus decisions on who should receive services have to be made. This is true in all sectors in society, including social welfare. By choosing to devote resources to one area of social welfare, society is forgoing the benefits that would have arisen had those resources been used to fund an alternative social welfare service or indeed a service outside the social welfare system. Thus, priorities for resource allocation must be set. Economics is concerned with making these priority-setting decisions in an explicit and rational manner.

To set priorities in any area where resources are limited, two variables need to be considered – costs and benefits. Economics is the study of welfare or well-being and the viewpoint taken is societal. It is concerned with the effect of an action on the well-being of the whole of society,

not just on the individuals directly involved (Arrow, 1963). The provision of residential services for people who misuse drugs, for example, will not only impact on service users and providers, but will also affect families and friends, health and social services, the criminal justice sector and the general public. These effects may be negative (costs) or positive (benefits) and the aim is to ensure that total benefits to society outweigh total costs, resulting in an overall improvement in society's total well-being. Although, in practice, more limited perspectives are acceptable, the improvement of society's total well-being is a central component of welfare economics – the theoretical framework on which economic evaluation is based – and broader perspectives are encouraged.

Using information on costs and benefits, the economist's criterion for deciding on a desirable allocation of resources is efficiency. In simplistic terms, an efficient allocation of resources occurs when the benefits of a service are maximised for a given cost. To ensure efficiency in the allocation of resources, it is necessary for social welfare interventions to be evaluated to establish their cost-effectiveness, not just their effectiveness.

Directing resources towards all services shown to be cost-effective (or efficient), however, is limited in two important ways. First, funds may be depleted before all cost-effective services are provided, if these services are large in number, particularly expensive or needed by a large proportion of the population. Second, society may prefer to fund a less efficient service if it believes that service will produce a more equitable distribution of resources. The efficiency criterion ensures that societal well-being is increased, but makes no judgement about which members of society benefit from this increase. Thus, when undertaking an economic evaluation, which is concerned primarily with efficiency, researchers should also explore the equity implications of their research.

The main purpose of economic evaluations is therefore to determine the best use of society's scarce resources – to get the most from your money. Economic evaluations can, however, be useful in many other ways (Knapp, 1997). They can provide broad information on the service

Box 1 Economic evaluation

- Economic evaluation is defined as the systematic attempt to identify, measure and compare the costs and outcomes of alternative interventions.
- Resources are limited and insufficient to meet total needs, thus decisions on who should receive services and which services they should receive have to be made.
- To set priorities in any area where resources are limited, two variables need to be considered – costs and benefits.
- The aim is to direct resources towards those services that produce the greatest benefit for the money available in order to maximise overall societal well-being.

and cost implications of social welfare problems, from the perspective of the service provider, the funding bodies, other providing agencies and service users and their families. They can help to assess the relationship between costs and effectiveness of services – in other words, they can help to determine the appropriateness of expenditure by examining whether the money spent on services is having a positive impact on outcomes for users. They can be used to explore which personal, family or situational circumstances increase or reduce the need for services, thereby providing information to aid budget planning. And, given the increasing cost-consciousness of many funding bodies, they can help researchers to gain funding for evaluations and agencies to gain funding for their services.

Economic evaluation thus has a number of uses and should be carefully considered for inclusion in future evaluations of social welfare services. For a useful introduction to the application of economic evaluation to social welfare programmes, see Holtermann (1998). More detailed discussions of the methods of economic evaluation can be found in Knapp (1984), Drummond *et al*. (1997) or HM Treasury (1997). Brief definitions and descriptions of relevant economic terminology can be found in Kielhorn and von der Schulenburg (2000).

Scope of the report

The focus of this report and the accompanying volume (Sefton *et al*., 2002) is on economic evaluation at the practice and individual level, as opposed to broader or more strategic policy questions. The report is thus primarily concerned with the economic evaluation of individual services and the impact of these services on user outcomes, costs and cost-effectiveness. Although economic jargon has been kept to a minimum, it is important for the reader to become familiar with some of the economic terminology in common usage and, where its use was unavoidable, such terminology has been clearly defined.

Our definition of social welfare encompasses social care, early intervention schemes, housing, urban regeneration, community development, work with families and welfare to work. Although the scope of this initiative excludes health care, this report draws heavily on the relatively well developed field of health economics. Economic evaluations in health care are common and there is much theoretical, methodological and empirical literature from which lessons can be learnt, as well as much overlap between the fields of health and social welfare. The primary purpose of any specific focus on the evaluation of health-care interventions within this report, however, is to provide examples of how economic evaluation may be applied in the area of social welfare. Thus, there is a stronger focus on the evaluation of complex health-care interventions, such as those in the mental health field, which arguably have much in common with many social welfare interventions (Byford and Sefton, 2002). Where available, economic evaluations carried out in the social welfare field have been used as examples to illustrate the processes of economic evaluation described.

Structure of the report

Chapter 2 lists the basic ingredients for a successful economic evaluation, providing an

overview of some of the issues that must be considered and the decisions that need to be made before an economic evaluation can take place. Chapter 3 outlines the main methods of economic evaluation and discusses some of the advantages and disadvantages of each and their applicability to social welfare evaluation. Chapters 4 and 5 discuss in more detail methods for the identification, measurement, valuation and analysis of costs and outcomes, respectively. Chapter 6 attempts to show how best to combine data on costs and effects in order to assess relative cost-effectiveness and provides advice on the interpretation, presentation and implementation of the results.

Different readers will require different levels of detail, ranging from a fairly basic understanding of the language and concepts involved, perhaps to aid their understanding of the impact of an economic evaluation on a proposal they are developing with the support of an economist, to a much deeper level of understanding of the methodology, in order to undertake some of the work themselves. For this reason, some of the more technical aspects of this report, which some readers may prefer to skip and which are not essential to a basic understanding, have been shaded.

2 Basic ingredients of an economic evaluation

The aim of an economic evaluation is to compare the costs of different interventions with the consequences, in order to determine whether a particular intervention is worth doing (in terms of the outcomes it generates) in comparison to other things that could be undertaken with the scarce resources available. In order to achieve this aim successfully, there are a number of requirements that should be explicitly considered before the onset of a study. The main ingredients are discussed below and summarised in Box 5 at the end of this chapter. A more detailed checklist of the criteria for assessing economic evaluations can be found in Box 27 in Chapter 7 of this report and a useful guide is provided by Drummond and Jefferson (1996).

Evaluation question

Economic analysis can be used to answer many different questions (see Box 2) and the design and data requirements of a study will depend on the purpose of the evaluation. Some of the questions listed in Box 2 would involve full economic evaluations (e.g. which services or combinations of services are the most cost-effective?). Others can be classed as 'building blocks' towards a full economic evaluation (e.g. what are the service implications and costs

Box 2 Typical economic questions

- What are the service implications and costs associated with the social welfare problem of concern, both now and in the future?
- What services are available and what are their costs?
- Are service costs related to service effectiveness?
- Which services or combinations of services are the most cost-effective?
- Are there services or combinations of services that reduce indirect service use (use of services elsewhere in the social welfare system)?
- Will higher investment in social welfare services today reap longer-term benefits in terms of cost savings and improved outcomes?
- Are there personal, family or situational needs or circumstances that increase or reduce the likelihood of greater service utilisation?
- What are the long-term effects of the social welfare problem of concern on the health, social care and education needs and service use patterns of the children of current service users?

Adapted from Knapp (1997).

associated with the social welfare problem of concern?). For the latter, the data requirements are likely to be more limited than for a full economic evaluation. In order to ensure that the evaluation is adequate to answer the questions posed, the purpose of a study must be clarified in advance. For a more detailed discussion of the clarification of objectives, see Chapter 6 of Volume I (Sefton *et al*., 2002).

Comparison group

Since economics is concerned with the use of scarce resources, implying that the provision of one service must be at the expense of others, economic analysis by definition should involve comparisons between alternative courses of action. The choice of comparison (or control) group can have significant implications for the design of a study and, ultimately, on the cost-effectiveness of the interventions under investigation. The less effective the control intervention, the more effective the alternative will appear in comparison, and the greater the danger of giving an overly optimistic impression of an intervention's effectiveness.

The most appropriate comparison is the most cost-effective alternative intervention currently available (the 'next best' alternative). Where the next best alternative is not clear-cut, a number of alternatives may be considered, including the most widely practised alternative or current local practice (Byford and Palmer, 1998). Where there is little evidence of effectiveness or cost-effectiveness, a 'do nothing' alternative may be required, as it cannot be assumed that existing services are better than doing nothing.

For example, a new adolescent support service for young people experiencing difficulties that is provided by a voluntary agency could be compared to alternative support services, such as a service provided by a local authority, where one exists. A 'do nothing' option may be more appropriate in an area where no alternative support service specifically for adolescents exists. In this case, a study may attempt to follow up young people who are having difficulties at school or who present to social services with difficulties at home but who are not considered a high enough risk for care or accommodation and compare those who receive the new service with those who do not.

Once selected, the alternatives being compared should be described in sufficient detail to enable readers to fully understand what is being evaluated. This allows the applicability of the results of a study to similar services elsewhere to be assessed and enables the services under evaluation to be replicated by other service providers, should they prove to be a cost-effective use of resources.

Perspective

The identification of costs and outcomes relevant to an individual study will depend, to a large extent, on the perspective of the study. From a cost point of view, economic evaluations in health care often take the perspective of the health service alone. Economics, however, is concerned with the impact of an action on the well-being of the whole of society, not just on the individuals or organisations directly involved, and the exclusion of certain sectors may alter the conclusions of a study (Johannesson, 1995).

The identification of all costs relevant to a particular study is not always straightforward, particularly in areas where many agencies are involved in providing services to one individual, such as social care and mental health. Box 3 lists the broad range of costs that may be relevant to users of social welfare services.

To certain interest groups, an evaluation carried out from a societal perspective may seem unnecessary. The inclusion of all relevant costs, however, will enable such groups to isolate the information relevant to their own perspective and will also allow the effect of their actions on other sectors to be determined. Narrow perspectives can limit the usefulness of an evaluation and may in fact be detrimental to the purpose, since they may result in a sub-optimal allocation of resources and a corresponding loss in societal well-being. Examples of how the perspective chosen can alter the results of the study are given in Box 4.

Although the societal approach is, philosophically speaking, the ideal, this will not always be possible in practice and indeed it is rarely adopted in health-care evaluations (Sassi *et al*., 2002). For example, many health-care evaluations do not include patient costs or the cost of informal care (care provided by friends or family), which may make community-based interventions appear unduly favourable compared to residential care. A recent review of the economic literature in the social welfare field suggests that, similarly, most evaluations tend to adopt the perspective of the programme funder or provider, rather than a broad societal perspective (McDaid *et al*., 2003). A narrow perspective will often be selected because those who commission economic evaluations are interested in a particular focus or because the societal perspective makes for a more expensive and time-consuming evaluation.

If a broad perspective is not possible, the chosen perspective should be explicit and the

Box 3 Costs that may be relevant to users of social welfare services

- Social-care services (e.g. social work, accommodation, day care).
- Primary health-care services (e.g. general practitioners, health visitors).
- Secondary health-care services (e.g. psychiatric services, clinical psychology).
- Education services (e.g. educational psychologists or education welfare officers).
- Education facilities (e.g. school for learning difficulties, pupil referral units).
- Voluntary-sector services (e.g. Childline, Barnardo's, Alcoholics Anonymous).
- Private-sector services (e.g. counselling, alternative therapies).
- Accommodation (e.g. sheltered living, staffed hostel accommodation).
- Criminal justice (e.g. family courts, youth offending team).
- User and family costs (e.g. travel to services and child care).
- Informal care (e.g. care provided by family or friends).
- Productivity costs (e.g. time off work or unemployment because of disability).

Box 4 Examples of the influence of perspective

The benefits and costs of JTPA title II-A programmes: key findings from the national job training partnership act study (Bloom *et al.*, 1997)

This study involved a randomised evaluation of three job training programmes as compared to the usual alternatives available for disadvantaged adults and out-of-school youths. Outcomes assessed included incremental earnings, educational attainment and receipt of social security benefits. The results were presented from two perspectives: the impact on overall earnings of the participants and the broader impact on employment levels. From the perspective of the adult participants, training programmes were beneficial; the additional earnings gained from employment as a result of the training, as well as the reduction in out-of-pocket expenses for other training courses, more than outweighed any loss of welfare benefits and additional taxes paid. However, the net benefits of the programme were much lower from a broader societal perspective because of the additional costs of running the new courses and the loss of revenue from other courses that participants would otherwise have attended. From the perspective of disadvantaged youths, training programmes did not produce any benefit; there was a small net loss from the perspective of the participants and an even greater loss from the societal perspective.

Multisystemic therapy treatment of substance abusing or dependent adolescent offenders: costs of reducing incarceration, in-patient and residential placement (Schoenwald *et al.*, 1996)

This study was an evaluation of multisystemic therapy (MST) in terms of the costs and reductions in days in prison, hospitalisation and residential treatment for substance-abusing or substance-dependent young offenders. Although MST itself was a cost-increasing intervention, taking a societal perspective, it was found to have a positive impact in terms of reducing the cost of subsequent days in hospital, residential care and prisons. These reductions in cost almost offset the additional costs of MST.

exclusion of any items should be explained and discussed in terms of their likely influence on the final results. A pragmatic approach, though not ideal, may be to adopt a public sector perspective, where investment of public resources in, say, social services is evaluated from the social services perspective, but also includes assessment of any impact on the resources of other public sectors, such as the health service or education department. Such a scenario is now more realistic given the greater degree of cross-departmental co-operation within the UK government (Edwards and Thalanany, 2001).

The perspective of an evaluation will also influence the evaluation question and the type of outcomes measured. At the individual (micro) level, interest will be focused on the

impact of an intervention on the service users. The evaluation question may be 'Is this new service a more cost-effective intervention for our clients than our current service?' and the most appropriate outcome measures to answer such individual-level questions will be measures that are specific to the particular needs of the client group. This can include condition-specific or service-specific measures that enable a particular condition to be monitored over time, such as activities of daily living, or more general measures of user satisfaction or quality of life.

At the level of the purchaser or provider of services, broader questions may be raised regarding the overall level of need in a particular area or issues of resource targeting requiring more generic measures of unmet need, satisfaction and quality of life that can be compared across different user groups. At the level of central government (the macro-level), even broader issues of resource allocation become significant, such as the relative importance of education services as compared to health or social services. These concerns are best evaluated using national indicators of such things as social exclusion, poverty, employment, mortality and morbidity, and outcome scales that are comparable across all sectors in society, such as quality of life. The focus of this report, however, is at the micro-level.

Costs

Calculation of relevant costs to be included in an evaluation involves three main steps: the identification of resources used (for example, contacts with social workers), the collection of data on the use of these resources and the calculation of appropriate unit costs for each resource element. Services to be included will be determined partly by the chosen perspective of the analysis, discussed in the previous section on 'Perspective', and should be built up from an understanding of the client group involved, through consultation with users and professionals, or within pilot or feasibility studies. Data on the use of these services can then be collected in a number of different ways, including interviews with study participants or searches of case notes. Unit costs must then be attached to every service used in order to calculate a total cost per study participant. Each of these steps in described in detail in Chapter 4.

Outcomes

A number of different types of outcome can be measured and each serves a slightly different purpose. Process outcomes, such as proportion of clients maintaining contact with a service or days in local authority care, can be used to document the 'administrative' success of a programme in achieving what it set out to do. Intermediate outcome measures that focus more on the individual, such as truancy levels or disability free days, are useful short-term measures of change and can act as a motivational tool both for clients and project workers. 'Final' outcome measures that record changes in the well-being or quality of life of the individuals concerned are better long-term measures of success in achieving the overarching aim of most social welfare programmes, i.e. to improve quality or quantity of life.

Outcome measures chosen for an evaluation will thus depend on the objectives of the evaluation and the objectives of the

intervention, which can be clarified through discussions with key stakeholders, including policy makers, funders, service providers and users. Since economics is fundamentally concerned with the maximisation of well-being, final outcome measures are preferred. The measurement of final outcomes, however, can differ and will have an influence on the method of economic evaluation chosen, discussed further in Chapter 3. More detailed information on the measurement and valuation of outcomes is provided in Chapter 5.

Study design

Most economic evaluations are quantitative in nature and the design of evaluations capable of producing good quantitative evidence of both effectiveness and cost may differ from designs used more commonly in social welfare research, such as those that are primarily qualitative. Although many alternative evaluation designs exist (discussed in detail in Volume I [Sefton *et al*., 2002]), including observational data (e.g. case-series, before-and-after and case-control studies) and decision analysis (e.g. simulation models and decision trees), many economists favour the randomised controlled trial and other experimental approaches that are capable of producing unbiased estimates of costs and effects. Such study designs produce two groups similar in all respects apart from the intervention received, thus observed differences in outcome can be attributed to the interventions, rather than other confounding influences. Experimental designs also enable the application of statistical techniques capable of detecting these differences between groups with a specified level of confidence.

Experimental approaches are, however, a relatively new framework for many social welfare researchers and may conflict with some of the principles they hold. In addition, there are situations where approaches other than randomised controlled trials may be more appropriate, for instance on ethical grounds (Black, 1996) or to be able to better generalise to the real world. Recently, there has been a growing interest in combining quantitative data from traditional economic evaluations with more qualitatively focused contextual information in a more 'holistic' economic evaluation (Jan, 1998; Coast, 1999). This approach considers how the impact of such contextual information, for example the enthusiasm of service providers, can influence outcomes and therefore help to determine whether a successful intervention may be replicated elsewhere. While this report recognises the usefulness of this approach, it concentrates on quantitative economic evaluation techniques that may be less familiar to the social welfare researcher. A more detailed discussion of alternative evaluation frameworks is provided in Volume I of this report (Sefton *et al*., 2002) and will not be discussed further here.

The timescale of an evaluation is another important design issue. Health and social welfare problems vary enormously in terms of duration. Some relatively minor health conditions, such as a broken bone, are easily fixed. Similarly, some social welfare concerns can be relatively easily dealt with, such as overnight accommodation for a child whose single mother needs to spend a night in hospital. Such 'acute' problems are short-term in nature and both the costs and effects of interventions are felt soon afterwards.

Evaluation can thus be relatively short-term with little danger of missing important effects. More complex health and social welfare problems, however, can be chronic and enduring, thus costs and benefits may need to be recorded over a much longer period of time.

Involvement of economists and statisticians at an early stage in the design of an evaluation is fundamental to the success of an economic evaluation. A considerable amount of pre-planning is required to ensure that all of the ingredients listed in this chapter are adequately incorporated and a useful reference to the design and analysis of trial-based economic evaluations is provided by Glick *et al*. (2001).

Box 5 Summary of the basic ingredients of an economic evaluation

- Evaluation question: should be driven by the purpose of the evaluation and achievable within the resources available.
- Comparison group: requiring identification of the 'next best alternative', which may be a 'do nothing' alternative.
- Perspective of the evaluation: to help determine which costs and outcomes are relevant to the evaluation. A broad societal perspective is preferred for an economic evaluation.
- Costs: requiring identification, measurement and valuation of cost components relevant to the chosen perspective.
- Outcomes: requiring identification, measurement and valuation of outcomes relevant to the evaluation. 'Final' measures of outcome focused on the individual, such as quality of life, are preferred for an economic evaluation.
- Study design: selection of a design capable of producing good quantitative evidence of both effectiveness and cost. Experimental designs, such as randomised controlled trials, are preferred for an economic evaluation.

3 Methods of economic evaluation

Five main methods of economic evaluation can be used to compare the relative efficiency of alternative social welfare interventions: cost-effectiveness analysis, cost-minimisation analysis, cost-consequences analysis, cost-utility analysis and cost-benefit analysis. Useful introductions to the different methods of economic evaluation are provided by Robinson (1993), in a series of brief papers, and Holtermann (1998). For a more detailed discussion, see Chapters 5–7 in Drummond *et al*. (1997).

All methods of economic evaluation involve the identification, measurement and comparison of all relevant costs and benefits, and they all measure costs in monetary terms, i.e. in pounds sterling or US dollars. They differ, however, in their approach to measuring the benefits of interventions under examination and also in the questions they can answer. This chapter defines and describes the main methods of economic evaluation, but a brief summary is also provided in Box 11 at the end of this chapter to aid recall and for those readers who would like to distinguish between the different methods but do not have a need for a more detailed exposition.

Cost-effectiveness analysis

Cost-effectiveness analysis (CEA) is the most commonly adopted approach to economic evaluation in health care and involves the valuation of benefits in a single 'natural' (condition-specific or service-specific) outcome measure such as level of depression or life years gained. The benefits of two or more interventions are combined with their respective costs to provide a measure of relative cost-effectiveness which can then be compared to other interventions employing the same measure of effect (described in more detail in Chapter 6). The use of natural units of outcome makes CEA easily transferable into social welfare research, where natural units would include such things as crime rates or measures of social exclusion. Some examples of cost-effectiveness analyses are summarised in Box 6.

CEA does have its weaknesses. First, it is impossible to make comparisons across a diverse spectrum of interventions competing for a share of a finite budget. Social welfare services are extremely varied and the aims and outcomes of services will differ greatly and often be multiple. Comparisons of cost-effectiveness using natural units can be made only between interventions whose outcomes can be measured on the same scale. Thus, CEA might be used to support funding decisions between two competing schemes for reducing, say, crime, but it cannot determine whether the same money would be better spent on a scheme to provide subsidised child care.

Second, it is difficult to capture all possible effects of an intervention on a single outcome scale that measures change in only one area of an individual's life. Social welfare services will often influence many areas of life, thus requiring a range of condition-specific or service-specific outcome measures. Multi-dimensional outcomes that are measured on a number of different scales can be difficult to interpret, particularly if improvements are seen on some scales but not others. To adequately explore the relationship between costs and

Box 6 Examples of cost-effectiveness analyses

A cost-effective comparison of supported employment and rehabilitative day treatment (Clark *et al*., 1996)

This study examined whether supported employment for the severely mentally ill would be more cost-effective, in terms of vocational outcomes, than day treatment. Data on vocational outcome, cost of the interventions and community resource use were obtained from 58 patients with severe mental health problems. The study found that supported employment was more cost-effective than day treatment since benefits in terms of vocational outcomes were better and costs were lower, or at worst equal to those of the day treatment programme.

Cost-effectiveness of intensive v. standard case management for severe psychotic illness: UK700 case management trial (UK700 Group, 2000)

Randomised controlled evaluation comparing intensive case management (caseloads of ten to 15) with standard case management (caseloads of 30–35) for people with severe psychotic illness. A large range of outcomes were recorded, including social, clinical and intellectual functioning, quality of life, satisfaction, drug and alcohol misuse and violence, but the primary outcome measure selected was days in hospital for psychiatric reasons. Costs included all health, local authority and voluntary sector services, as well as use of prison and police custody. The cost-effectiveness was carried out using the primary outcome measure, which found no significant difference between intensive and standard case management.

outcomes, a single number is needed to represent the effectiveness.

To illustrate, a scheme that provides home adaptations for the promotion of independent living for disabled people may have an impact on psychological, social and family functioning, as well as physical functioning. All these areas can be measured in natural units but a CEA can only be carried out with one outcome scale. Under such circumstances, a CEA will often be based on the outcome measure considered to be of greatest importance to the purpose of the evaluation (the primary outcome), requiring some judgement to be made about the relative value of the alternative outcomes of interest.

Cost-consequences analysis

Cost-consequences analysis (CCA) is one method that has been employed to overcome the problem of capturing all relevant consequences within a CEA, and a useful discussion of it is provided by Mauskopf *et al*. (1998). CCA involves the presentation of a range of outcome measures alongside the costs. No attempt is made to formally combine costs with benefits and decision makers are left to form their own opinion regarding the relative importance of the alternative outcomes presented. Where one service is found to be dominant on all measures of outcome, the relative cost-effectiveness may be obvious, but

this will not always be the case. CCA has been used to evaluate complex interventions where outcomes cannot easily be summarised in a single measure and some examples are provided in Box 7.

Although CCA is limited by the inability to rank interventions in terms of cost-effectiveness, it is perhaps more consistent with the way policy decisions are made in practice. Decisions are not necessarily made on the basis of cost-effectiveness alone but on the basis of a range of different criteria and perspectives. Explicit presentation of the impact of an intervention on all costs and outcomes provides decision makers with all relevant economic evidence, while leaving them free to incorporate other forms of evidence or other perspectives, which may include public perceptions and preferences or equity considerations.

Box 7 Examples of cost-consequences analyses

A cost-effectiveness evaluation of parent training (Thompson *et al.*, 1996)

Evaluation of the Common Sense Parenting approach to the treatment of children with behavioural problems. The intervention involved regular two-hour group sessions with two parent trainers for eight weeks, followed by in-home clinical assessment and fortnightly sessions with one parent trainer. The intervention had already been found to be effective but concern about its expense led to an evaluation comparing a cheaper approach, involving just one parent trainer, to a control group consisting of families on the waiting list for the service. The experimental intervention was found to significantly improve levels of child behaviour, parental esteem and family satisfaction, in common with the original programme, and the cost was less than half that of the original programme. Although the title suggests that this study was a cost-effectiveness analysis, the authors in fact presented the costs and a range of outcomes without selecting a primary outcome measure with which to assess cost-effectiveness.

A three-year comparative longitudinal study of a school-based social work family service to reduce truancy, delinquency and school exclusions (Pritchard and Williams, 2001)

Comparison of an intensive educational social work intervention, including teacher counsellors, with standard education social work, in terms of both costs and effects. The evaluation was carried out in four schools in areas of high socio-economic deprivation in England. A broad range of outcomes were assessed, using both qualitative and quantitative methods, including psycho-social, educational, crime and delinquency indicators. The economic component of the study was conducted from the perspective of the education authority and criminal justice system, and included the cost of the interventions and estimates of the costs avoided as a result of reductions in school exclusions and crime rates. The study found that the experimental intervention was broadly successful across the range of outcome measures and was cost saving.

In the absence of adequate measures capable of capturing all the consequences of a particular intervention, CCA is likely to be a useful tool for social welfare research. The presentation of all costs and consequences can greatly enhance the understanding gained from an economic evaluation and thus CCA should be encouraged even when a primary outcome measure has been selected and a CEA carried out.

Cost-utility analysis

An alternative solution to multiple outcomes is to condense them into one generic measure, which is the approach adopted in cost-utility analysis (CUA), a specific form of CEA. As with CEA, a measure of relative cost-effectiveness can be derived, but outcomes are measured in terms of utility (level of satisfaction, well-being, quality of life, etc.). One example of a utility-based measure is the quality-adjusted life year (QALY) (Loomes and McKenzie, 1989 or Chapter 6 in Drummond *et al*., 1997).

The calculation of QALYs involves the application of quality-adjustment weights to the length of time spent in a range of potential outcome states that may result from a particular intervention. Taking a simplistic example, evidence suggests that users of a new housing initiative for homeless people will experience one of four possible outcomes: (a) permanently accommodated in a flat of their own, or (b) provided with short-term accommodation in a hostel for homeless people, or (c) provided with temporary emergency accommodation in a refuge, or (d) remain homeless. Individuals will have different preferences for these four possible outcome states and the aim is to value each option in terms of the quality gained relative to perfect quality of life (with a score of 1) and death (with a score of 0) – the utility gained from each. On average, we would probably expect option (a) to be valued higher than (b), (b) higher than (c) and (c) higher than (d), although this cannot be assumed.

Once generated, the quality adjustment weights are multiplied by the time spent in each outcome state and summed to provide the number of quality-adjusted life years, thereby incorporating the effects of an intervention on both the quantity and quality of life. The results are expressed in terms of the additional cost per QALY gained from undertaking a particular intervention, providing a common measure of output that allows comparisons to be made between any number of diverse interventions. Resources should then be directed towards those interventions that involve the lowest cost per QALY ratio. Examples of cost-utility analysis are provided in Box 8.

The theory behind utility measures is attractive and their importance should not be dismissed, but a number of weaknesses limit the usefulness of CUA within the field of social welfare, at least in the short term. Conceptually, the idea of condensing the benefits of a scheme for urban regeneration, for example, into a single outcome measure can be hard to swallow. Such schemes are area-based, rather than focused on a specific group of individuals, and they may influence many people, in many different ways, across many dimensions. In addition, utility scales have been criticised for their conceptual foundations, for the methodology employed, for their lack of sensitivity to change, for focusing on health-related quality of life while ignoring other aspects of care and for ignoring equity

Box 8 Examples of cost-utility analyses

Cost-effectiveness of an HIV risk-reduction intervention for adults with severe mental illness (Johnson-Masotti and Pinkerton *et al.*, 2000)

Three education interventions designed to reduce the rate of HIV infection in an adult population with severe mental illness were evaluated: a one-off, individual risk-reduction programme; a seven-group-session cognitive-behavioural risk-reduction intervention focusing on behavioural change; or a seven-group-session intervention which taught participants effective communication strategies for disseminating information on HIV prevention (group advocacy). Quality-adjusted life years (QALYs) were used as an outcome measure to take account not only of the shorter life expectancy of someone with HIV, but also of the impact of the virus on quality of life. QALYs lost as a result of HIV infection were estimated from a review of studies that had calculated QALY values for HIV infections averted in the general population. Group advocacy was found to be the most cost-effective intervention for men, while, for women, only one-to-one individual education sessions were cost-effective.

Economic evaluation of a support programme for caregivers of demented elderly (Drummond *et al.*, 1991)

A programme to support caregivers looking after people with dementia, including the provision of two levels of short-term respite care, was compared with conventional community home-based care. Caregivers were randomly allocated to one of the support programmes. Carers' health outcomes were measured in terms of quality-adjusted life years (QALYs), using the Caregiver Quality of Life Instrument to describe health states.

considerations (Loomes and McKenzie, 1989; Drummond, 1991; Williams and Kind, 1992; Oliver *et al*., 2002). As research continues to address these problems, however, the usefulness of utility scales may grow.

Perhaps the main obstacle to the use of CUA in the evaluation of social welfare services is the lack of utility scales appropriate to the field. Although a significant quantity of research has been carried out into the development of utility scales for use in health economics, these measures tend to be health focused and may not be broad enough to capture the full impact of social welfare policies. To increase the usefulness of the utility approach in the social welfare context, appropriate measures must be developed, which will require investment in research. An alternative approach is to directly value relevant outcome states within an evaluation. This method, discussed in more detail in Chapter 5, is more time and resource consuming than using a pre-existing generic scale, but at least provides more relevant valuation in areas where few appropriate generic scales exist.

Cost-benefit analysis

Cost benefit analysis (CBA) is a method of economic evaluation that is used less often in health care than CEA or CUA, although it is more common in some policy areas, such as transport and environment. CBA requires both costs and benefits to be valued in monetary units. It thus becomes possible to directly compare the costs with the benefits of an individual project (i.e. calculate the net benefit) to see which is greater, without the need for a comparator. To improve the allocation of resources, an intervention should be adopted if the benefits exceed the costs, subject to any overriding budget constraint. Examples of cost-benefit analysis in the social welfare field are provided in Box 9.

Like CUA, CBA allows the comparison of any number of diverse interventions, since the benefits are always measured in the same units. In addition, it is possible to make comparisons across different sectors, such as health care, education or defence. However, difficulties arise when attempting to value benefits in monetary terms. How do you decide how much an increase in the length of a person's life is worth? How do you put a monetary value on freedom from abuse or avoidance of homelessness? Methods do exist to answer questions like these, which are discussed in more detail in Chapter 5 in the section headed 'Monetary valuation of outcomes', but they are difficult to apply and can be a time-consuming and costly addition to an evaluation. Hence, CBA in health care is relatively rare and the extent of its use in the wider social welfare field is as yet unknown. Useful references to the methods of CBA include Johannesson and Jonsson (1991) and McIntosh *et al*. (1999).

CBA should be clearly distinguished from the related technique of cost-savings analysis, which is and may continue to be, more commonly used in the evaluation of social welfare services. The analysis of cost savings involves the comparison of costs and benefits that are easily converted into monetary units, with other effects ignored. The costs of an intervention are compared to the savings that are generated through, for example, reductions in crime or school exclusions. 'Outcomes' of this type can be converted relatively easily into monetary units since they involve known and observed costs. The savings from reduced crime rates, for example, can be measured in terms of reductions in expenditure on the criminal justice system. Such analyses are less scientifically sound than CBA since they do not attempt to value *all* relevant outcomes, in particular final outcomes for the users involved.

One example of a cost-savings analysis is an evaluation of a London-based Link Worker scheme, which aims to improve access to support in the community for people with mental health problems who have come into contact with the criminal justice system (Revolving Doors Agency, 2000). This study looks at whether the costs to local agencies of providing additional services to the scheme's client group are offset by a fall in the cost of 'crisis' services, including temporary accommodation and A&E services.

Box 9 Examples of cost-benefit analyses

Cost-benefit analysis of domestic energy efficiency (Clinch and Healy, 2000)

Evaluation of the costs and benefits of a potential domestic energy efficiency programme in Ireland, which has the highest rate of fuel poverty and excess winter mortality in northern Europe. The intervention would involve the introduction of energy efficient technologies into 1.2 million homes over a ten-year period. A number of outcomes were included, such as morbidity, mortality, comfort and energy efficiency, and these were valued using a range of different techniques. One approach was willingness to pay, which involves asking respondents how much they would be willing to pay for a particular improvement in outcome (see Chapter 5, section headed 'Monetary valuation of outcomes' for more detail). The costs of the initiative were estimated from consultations with surveyors regarding the labour and resources required. Total costs of the programme were estimated to be Euro 1,601 million while the benefits were valued at Euro 4,723 million, thus the programme resulted in an overall net benefit and was considered cost-effective.

Drug abuse, crime costs and the economic benefits of treatment (Rajkumar and French, 1997)

This study sets out a proposed methodology for the valuation of the costs of crime, including the psychological costs incurred by the victims. A number of different valuation methods were used, including property valuations (used as an indicator of the willingness to pay to avoid crime by moving to a low-crime area) and jury compensation awards for different injuries sustained (exploration of the amount of compensation awarded in previous cases involving similar injuries) (see Chapter 5, section headed 'Monetary valuation of outcomes' for more detail). These methods were applied to data from a previous study of offenders who entered a drug treatment programme, which included an assessment of crimes committed in the year before and after the programme was implemented. A partial cost-benefit analysis was undertaken that included treatment costs (estimated to be $3,000 per individual) and the cost of crimes avoided ($10,000 per individual).

Cost-minimisation analysis

Where there is already good existing evidence to suggest that the interventions under consideration are equally effective, a cost-minimisation analysis (CMA) can be undertaken which involves assessment of the costs alone. Given equal outcomes, the evaluation involves the comparison of costs, to determine the least cost alternative. Where outcomes are expected to vary, alternative methods of economic evaluation should be employed. Examples of cost-minimisation analyses are provided in Box 10.

The main advantage of this approach is simplicity – an assessment of costs alone greatly reduces the burden of an evaluation in terms of time and resources, providing evidence more

cheaply and possibly in less time than would a full economic evaluation. The burden on service users is also reduced if the decision is made not to reassess patient-level outcomes. The disadvantage, however, is the loss of accuracy that may result from the collection of costs and outcomes from different individuals and/or over different time periods. For example, if the effectiveness data had been collected some years ago, the current service may differ from the original service and thus the impact on outcomes may no longer be the same. It is also worth being aware that some studies purporting to be CMAs may in fact be more limited cost studies that are not based on any prior evidence of the equality of outcomes.

Box 10 Examples of cost-minimisation analyses

Developing community alternatives to group home placements for SED special education students in the Ventura County system of care (Ichinose *et al.*, 1994)

Ichinose and colleagues evaluated the impact of a new multi-agency, community-based approach to the management of children with special education needs in a county in California. The approach involved the provision of case management and additional school-based services, psychiatric and social workers, and day treatment, and an important aim was to reduce the number of children being placed in residential care, and thus the costs of residential care provision. Community-based services were assumed to provide equivalent care to residential care and outcomes were estimated to be at least as good, thus the evaluation assessed costs only. The first nine months of the programme were compared with experiences immediately prior to the programme and with estimated projections of the number of anticipated placements had the programme not been implemented. There was no change in the number of students in residential placements over the first nine months of the programme compared with that prior to implementation, but the number was significantly lower than the projections. Similarly, although actual costs did not differ greatly from costs prior to implementation, they were much lower than the projected estimates.

Estimating the cost of three case management programmes for treating people with severe mental illness (Hu and Jerrell, 1998)

Assessment of the costs (and savings) of three models of case management (assertive community treatment, clinical model and intensive broker model) for 122 adults with severe mental illness. Service utilisation and cost data were obtained for all participants who completed four time periods in the study: six months before the case management programme began and six, 12 and 18 months after the programme began. The findings indicated that all three case management programmes studied significantly reduced the costs of caring for people with severe mental illness, as compared to the period before case management was introduced.

Selection of a method of economic evaluation

Five main methods of economic evaluation can be used to compare the relative efficiency of alternative social welfare interventions, summarised in Box 11. All methods involve the identification, measurement and comparison of all relevant costs and benefits, and they all measure costs in monetary terms, i.e. in pounds sterling or US dollars. They differ, however, in their approach to measuring the benefits of interventions under examination and also in the questions they can answer.

The selection of an appropriate method of economic evaluation will thus be influenced by the method chosen to measure outcomes,

Box 11 Methods of economic evaluation

- *Cost-effectiveness analysis (CEA)*: method of economic evaluation in which the outcomes, or effects, are measured in units that would commonly be applied in the area of interest to describe either the condition of concern or the aims of the service. For example, measures of the level of depression in mental health, measures of crime rates in criminal justice research, or measures of family cohesion in social work.
- *Cost-consequences analysis (CCA)*: where interventions have an impact on multiple areas of an individual's life, such as social, psychological and family functioning, a CCA can be useful. CCA involves the presentation of a range of outcome measures alongside the costs. No attempt is made to formally combine costs with outcomes and decision makers are left to form their own opinion regarding the relative importance of the alternative outcomes.
- *Cost-utility analysis (CUA)*: an alternative solution to multiple outcomes is to condense them into one generic measure, which is the approach adopted in CUA. CUA is similar to CEA, but outcomes are measured in terms of 'utility' (quality of life, well-being, etc.). Utility-based measures are used extensively in the evaluation of health-care interventions. One example is the quality-adjusted life year (QALY), which combines gains in the quality and years of life produced by one service as compared to another.
- *Cost-benefit analysis (CBA)*: method of economic evaluation that requires all benefits, as well as the costs, to be valued in monetary units (i.e. pounds sterling). It thus becomes possible to directly compare the costs with the benefits of an individual project (calculate the net benefit) to see which is greater, or to compare the net benefit of two or more interventions.
- *Cost-minimisation analysis (CMA)*: method of economic evaluation which can be carried out when there is good existing evidence to suggest that the interventions under consideration are equally effective. Given equal effects, a cost-minimisation involves the calculation of costs alone, to determine the least costly option.

and vice versa. Where good evidence already exists to suggest that one service is more effective than another, then a cost-minimisation analysis may be the simplest way to ascertain cost-effectiveness. For the evaluation of interventions that impact on multiple areas of service users' lives, a cost-consequences analysis may be preferred. Alternatively, a cost-utility analysis could be undertaken if an adequate utility-based measure is available. Cost-effectiveness analysis is useful for the evaluation of interventions that are targeted at one specific aspect of users' lives and cost-benefit analysis may be preferred when no comparison group is available.

From a practical point of view, evaluations may be constrained by available resources and available measures of outcome. Time and resource constraints, for example, may negate the possibility of using long or complex measures of outcome, designing an appropriate generic measure where one does not exist, or attempting to value outcomes in monetary terms. In addition, it is important to be aware of the need to compare results with other studies and therefore the need to consider using the same methods as previous evaluations.

4 Cost measurement and valuation

All the methods of economic evaluation described in Chapter 3 require the identification, measurement and valuation of all costs relevant to the interventions under investigation. Some of the key issues involved were highlighted in Volume I (Sefton *et al*., 2002). This chapter provides more detailed and practical guidance on the calculation of costs, and is thus relatively 'technical'. A number of sections in this chapter have been shaded to indicate sections that can be skipped by those who do not require a detailed understanding of how to calculate costs. The unshaded areas, however, contain useful information, which should provide the reader with a relatively complete understanding of the basic terminology and important concepts involved in the calculation of costs.

Types of costs

Costs can be usefully split into the following categories.

- *Programme costs*: the direct costs of providing a service, which will include the cost of all the individual elements of the service, such as staff costs, volunteer time, buildings, equipment, transport, support services, etc.
- *Non-programme costs*: any resulting service effects, such as the costs associated with meeting uncovered need, the savings that may result because of a reduction in the need for alternative services, or the cost to informal carers (care provided by a family member or friend) as a result of an increase in the time they must spend caring.
- *User and family costs*: any costs incurred by the person using the service or their family as a result of consuming the services, such as travel to and from services, child-care arrangements or informal care requirements.
- *Productivity costs*: costs that result from an impaired ability to work as a result of illness or disability, lost economic productivity because of premature death, or economic losses incurred by informal carers who may have to reduce their work hours or give up work altogether as a result of their caregiving activities. These costs may fall to individuals (lost income) and to society (lost productivity).

Costs and savings

It should be noted that, when calculating the overall cost of a service, an evaluation should also be concerned with any savings that result. Thus, the overall economic impact of a home-based support service for disabled children that reduces the need for residential care is equivalent to the cost of the support service minus the cost of residential care that would otherwise have been used.

Transfer payments

Social security benefits are often incorrectly seen as costs. In fact, they are known to economists as transfer payments, as they are simply transfers of money from one group in society to another, which generate no productive service, unlike the costs listed above. They are a method of redistributing income within a society. Transfer payments are often excluded from micro-level economic evaluations because they

do not involve an overall cost or benefit to society as a whole, since the costs to those who pay the transfer payments are cancelled out by the benefits to those who receive them. However, there are associated administrative and other taxation-based costs that should be included in an analysis where they constitute a significant cost element (HM Treasury, 1997).

Transfer payments can have a more significant role to play when considering macro-level evaluations, such as government policies to reduce unemployment. In addition, when considering issues of equity rather than efficiency alone, the distribution of income within a society becomes an important concern. Efficiency criteria make no judgement about who gains or whether an efficient allocation of resources is 'fair'. Social security benefits are one method available to redistribute gains and losses if society believes the current distribution to be inequitable.

Identification of cost components

The first step towards the calculation of costs is the identification of all services used (or cost components). As discussed in Chapter 2, services to be included will be determined partly by the chosen perspective of the analysis and should be built up from an understanding of the client group involved. This can best be achieved through consultation with users and professionals, perhaps in interview or focus group discussions, or through pilot or feasibility studies.

Attribution of costs

Whether to include all services used, or only those directly related to the need, condition or problem to which a service is targeted, is debated and dependent on the ease with which individual service items can be attributed to the area of interest. The more complex the condition or need, the harder it becomes to accurately attribute a service, and the more arbitrary and subjective the process of attribution becomes.

The family of a young person at risk of entering local authority care, for example, may access a range of services that are directly related to the perceived problems that resulted in the 'at risk' status, including perhaps social work support, temporary foster care and family therapy. Other services are more difficult to attribute, however. Was an accident and emergency attendance the result of a fall while playing in the park, over-zealous discipline by the parent, or physical abuse? Was a contact with a solicitor related to the young person or a dispute with a neighbour? Did an appointment with an educational psychologist result from family difficulties or problems at school?

Attribution can be handled in a number of different ways that range from over-exclusion to over-inclusion.

- *Over-exclusion*: the easiest way to ensure all services included are of relevance to the condition or need of interest is to include only those services that are most directly and most obviously important. Thus, the example of an evaluation of young people at risk could be limited only to social services interventions that are focused on the young person. This method is likely to exclude a number of services of relevance, but ensures that all superfluous services are unable to influence the results in a way that may be detrimental to the purpose of the study.

- *Over-inclusion*: at the opposite end of the spectrum, an evaluation may include all services used by the young person and their family, irrespective of reason. This method avoids the problems of subjectivity and the criticism of arbitrary selection, and ensures that all relevant services are included. However, there is a danger that high-cost services of limited or no relevance will dominate the final results and mask the true impact of the intervention under evaluation (Johnston *et al*., 1999).
- *Expert attribution*: a useful compromise is to ask experts in the relevant field to attribute the services used. The accuracy of this method will be enhanced by collection of detailed information on the reasons for service use, thus an expert will be better able to attribute an accident and emergency contact if he or she is aware of the circumstances that lead up to and the reason for attendance. It must be noted, however, that the more detailed the information, the greater the burden of data collection, which is discussed in more detail in the next section.

Measurement of cost components

Once a complete list of likely services has been drawn up, there are a number of methods available to measure the quantity of these services used by study participants, including questionnaires, diaries or searches of case notes. The method chosen will depend on the time and resources available, and whether the data collection is retrospective or prospective. Prospective evaluation is generally preferred as it allows the researcher greater control over the data collected and removes many of the problems of recall and lost case notes. For more information on data collection methods see Mauskopf *et al*. (1996), Chapters 10 to 12 in Moser and Kalton (1996), Chapter 3 in Johnston *et al*. (1999), Chapter 11 in Bowling (2002).

Questionnaires

Service-use questionnaires can be self-reported or completed by researchers at interview and the choice will often depend on the practicality of carrying out interviews and the complexity of the questionnaire. Where interviews are already planned for the collection of outcome data and where service-use questionnaires are particularly long or complex, the preference would be for completion during interview. Self-report or postal questionnaires may be preferred, however, when interview time is limited or unavailable, but only for short and simple questionnaires. Although postal questionnaires may appear to be less time and resource consuming, this should be balanced against the inevitably lower response rates.

Service-use questionnaires are used in prospective evaluations (evaluations that take place at the same time as the intervention under study), but are retrospective in nature, since they involve asking respondents to recall service use over a particular period of time, such as the previous three months. Such questionnaires should not be used in retrospective studies (studies that take place at some point in time after an intervention has been undertaken), since it is unreasonable to expect respondents to be able to remember their own service use for

some period of time in the more distant past. Researchers should also be aware of the possibility of evasive responses to questions that may be considered embarrassing or incriminating to the respondent, such as contacts with drug and alcohol services or the criminal justice system.

A number of service-use questionnaires have been designed for use within the health-care field and, perhaps of more relevance to social welfare research, mental health care. One example is the Client Service Receipt Interview, which includes sections on accommodation, employment and income, service receipt, informal support and satisfaction with services (Beecham and Knapp, 2001). Such questionnaires often need to be adapted for new evaluations as some services may be used only within specific areas of health or social welfare. The disadvantage of questionnaires is the need to rely on the memory of interviewees over what can be a significant number of months.

Diaries

Service-use diaries are one method of improving recall and involve asking participants to record their use of services prospectively over the period of a study. Diaries can also be used to record time spent on different activities and are useful for exploring the processes of care or for calculating time spent on informal care. Diaries can be produced that are highly structured and involve simple tick boxes. Drug trials, for example, will often require participants to record each time medication is taken by ticking against pre-printed dates, as a check on compliance. The more complex and the broader the range of services used, however, the harder it becomes to simplify service-use diaries, and a balance must be struck between ease of completion and the comprehensiveness of the data collected. In addition, people leading chaotic or itinerant lives, those who are sceptical of authority or the research and those with limited reading abilities are unlikely to agree to or remember to complete diaries.

Case notes

An alternative method of enhancing accuracy is to collect retrospective information from case notes or electronic administrative databases. Records are likely to be more accurate than relying on user recall over a significant period of time, but record searches can be time consuming, may not record exactly the information needed and will often be hampered by poor completion, missing files and illegible entries. In addition, the data will often be limited to the use of services provided by the agency to whom the case files belong. A multi-sector picture can be built only by exploring the case files of many different agencies. Over time, the increasing use of computerised records may make this method of data collection easier.

Literature

Where retrospective and prospective data-collection strategies are not feasible, it may be possible to gather information on the services used by particular client groups from previous research, a strategy most often used in modelling studies. Although economic analysis in the area of social welfare is limited (McDaid *et al*., 2003), a systematic review of the literature available may provide some usable information.

Which cost?

The total cost of providing services for each individual over the period of an evaluation is calculated by multiplying service-use data by appropriate unit costs. A number of different types of cost exist and it is important to understand the distinctions in order to ensure that the most appropriate cost is calculated for the purpose of the evaluation.

Opportunity costs

The economist's definition of cost (opportunity cost) reflects a concern for society's well-being and can be distinguished from accounting costs, which simply reflect the amount of money spent on a service. Because total resources are limited, opportunity cost measures the true value of a resource based on its value in the next best alternative use. By choosing to devote resources to one area of social welfare, society is foregoing the benefits that would have arisen had those resources been used to fund an alternative social welfare service, or indeed a service outside the social welfare system. The opportunity cost relates to the benefits (or opportunities) that are lost by not directing those resources to their best alternative use. The opportunity cost of an individual who has volunteered to go shopping for older people with reduced mobility, for example, should be valued in terms of what would have happened otherwise. A reasonable proxy for the cost of this individual's time would be lost earnings, if that person would otherwise be in paid employment, or perhaps the cost of a home help who is paid by a local authority to help this client group with shopping and housework (discussed in more detail later in this chapter in the section headed 'Valuation of informal and voluntary care costs').

In practice, the calculation of opportunity costs for every individual item of service will often be impractical. This would involve making judgements about the next best alternative uses for a potentially large range of different services and components of services. Instead, the actual monetary cost will frequently be used to measure the opportunity cost because money is a more convenient method of measuring the cost than alternative uses. Thus, it is often assumed that the opportunity cost of a service, such as an hour of a social worker's time, approximates the actual monetary cost of all components of that service. In the case of the social worker, this would include the cost of the worker's time and travel plus any administrative, managerial and capital overheads. In fact, the money paid for a service will be an accurate measure of opportunity cost under certain conditions, in particular that of perfect competition. A detailed discussion of perfect competition and the implications of its existence or non-existence is beyond the scope of this report and readers are referred to Knapp (1993).

Fixed and variable costs

A useful distinction to make is that between fixed and variable costs. Fixed costs are those costs that, in the short term, do not vary with the level of output (i.e. the quantity of a service produced). These 'overhead' costs are borne even if no output is produced and include payment of rent on buildings and interest payments on borrowing. Variable costs are those that vary directly with the rate of output and include the cost of staff, raw materials, fuel and power. In the long run, by definition, there are no fixed costs; all costs are variable, since even the capacity of a building can be increased, given enough time.

Short-run and long-run costs

Definitions of the short run and long run are simply the antithesis of the fixed and variable cost distinction. The short run is defined as that period of time over which at least one factor of production (inputs such as land, staff, raw materials and buildings) is fixed, while the long run is that period of time over which all factors of production are variable. Since factors of production vary between service providers, the long run will also vary according to the characteristics of the individual services. Thus, a service that is significantly dependent on the location in which it is based and the equipment contained within that building, such as the provision of social security payments, will have a longer long run than a community-based service that primarily consists of staff input, such as a home-based family intervention. The importance of the long run is that it is long enough to permit the producer to choose the most efficient (cost-effective) combination of inputs to produce any given level of output.

Average and marginal costs

Also of importance is the distinction between average and marginal costs. Average costs are the cost per unit of output produced, calculated by dividing the total cost of a service by the number of units of output; for example, the cost per night spent in a residential care home. Marginal costs result from a unit increase in the rate of output of a service. For example, if a support team for children at risk currently supports 49 clients for a total cost of £200,000 and, when it increases the caseload to 50, total costs rise to £200,500, then the marginal cost is £500. The effect of a unit increase in output can vary enormously and will depend on existing spare capacity. Using the same example, if one support worker had not yet reached full caseload capacity then the additional young person could be added to the service with little addition to total costs, except perhaps extra travel expenses and administration costs. Under such circumstances, a marginal cost of £500 would not seem unreasonable. A service that has reached

Continued overleaf

Average and marginal costs continued

its caseload capacity, however, could take on a new case only if a new member of staff were employed. This would add a great deal more to the total cost since it would include salary costs, administration costs, travel costs and even extra office-space costs. The marginal cost in this case could be as high as £50,000.

Long-run marginal costs

The preferred cost to calculate in an economic evaluation is the long-run marginal cost (LRMC) of a service, which can be defined as the additional cost of producing one more unit of output when all elements of a service (factors of production) are variable. The long run is preferred because only in the long run can we assume that a service is capable of taking on any number of users. An evaluation that finds a particular service being piloted in a small area of an authority to be cost-effective, for example, cannot simply be provided to all those in the authority who may need it in the short run. Only in the long run can expansion (or contraction) of all factors of production take place, allowing widespread implementation. Similarly, the marginal cost is preferred because this cost is more relevant to expansion or contraction than average costs. Expansion and contraction are issues of importance because economic evaluation is concerned with the notion of shifting resources from one area (contraction) to another (expansion).

In practice, LRMCs are difficult to estimate because they require knowledge of how factors of production can be changed in the future. Instead, unit costing often involves attempts to approximate LRMCs. Short-run average costs (SRACs) are assumed to be a good proxy for LRMCs, where SRACs include all factors of production, both those that are fixed and variable. Hence, costing requires the calculation of all the individual elements of a service, including buildings and management overheads that are less readily identifiable and difficult to allocate across appropriate units of production, but are nonetheless an important component of unit costs.

Valuation of direct costs

In many instances, local service providers may be able to supply unit cost information or the raw data needed to calculate the costs directly. However, this will not always be the case and direct calculation is often needed. There are two main approaches to the direct calculation of costs: bottom-up and top-down.

Bottom-up costing

Bottom-up (or micro) calculation, described in more detail below, involves detailing and quantifying all the individual elements of a service, then costing each element separately. Bottom-up costing is the most accurate method of unit-cost calculation but, dependent on the complexity of the service being costed, can be extremely time and resource consuming. The

more complex the service or the greater the number of component parts (e.g. staff, buildings, equipment, etc.), the more complex the costing exercise will be.

Top-down costing

A detailed bottom-up costing may not always be feasible or appropriate, for example where time and resources are limited or for particularly complex services. In such situations, a top-down approach may be preferred. Top-down (or gross or average per diem) calculation is much less resource intensive but also less accurate. This approach involves taking the total cost of a service, for example the total cost per annum (as provided by the appropriate finance department), and dividing it by an appropriate unit, such as the number of individuals receiving the service per year.

Prioritisation of unit costs

Where multiple agencies are involved in the care of a study population, it may not always be appropriate to carry out a detailed cost analysis of all services used by participants. Time and resource constraints may necessitate some prioritisation of costs, with greater effort devoted to the costing of 'key' services. Key services include those that are central to the evaluation (such as nursing staff in an evaluation of home-based nursing care) and services that contribute towards a significant proportion of the total costs (such as residential care facilities in a study of local authority care and accommodation placement types). Detailed calculation of unit costs is recommended for these key services, but top-down approaches or published unit costs could be applied to those services where expending significant effort on bottom-up costing would not be worthwhile in terms of the value added to the evaluation. Evaluations thus commonly employ more than one method of valuing costs.

Where evidence suggests that the additional cost of certain services will be extremely small in relation to total costs, it may be appropriate to exclude these services altogether since the effort expended on measuring and valuing these costs may not be worthwhile. The selection of a limited number of those services considered to be of most significance is known as reduced list costing (Knapp and Beecham, 1993). Similarly, where evidence suggests that the use of a particular service is unlikely to vary between the different groups in the evaluation, it is reasonable to exclude this service on the grounds that economic evaluation is concerned more with relative than absolute costs – the additional cost of one intervention over another, rather than the actual total cost of each comparator. The exclusion of costs that are equal across interventions will therefore not greatly impact on the purpose of the evaluation.

A decision to exclude certain cost elements, however, should not be taken lightly. It is extremely useful, for example, for the full cost results to be available for comparison with future research that adds to the body of evidence in a particular area. Cost exclusions may render such comparisons impossible. Furthermore, with a client group that has not been subject to an economic evaluation before, there will be little evidence to ensure that excluded costs are either small or equal across groups. Any cost exclusion should be noted, justified and discussed in terms of the potential impact of the exclusion.

Timing

The year chosen to calculate costs should be as up-to-date as possible to maximise accuracy, since accounting procedures tend to improve over time. For relevance, however, costs should apply to the time period in which the evaluation took place, since services change over time (Beecham, 2000). Thus, the ideal accounting year to choose for a study that took place over the period January 1998 to December 2000 is likely to be the financial year 2000/2001, since the service was still in existence throughout the majority of the period and these costs are likely to be more accurate than, say, 1998/1999 or 1999/2000 costs.

Costs (and indeed benefits) that are borne over a number of years need to be adjusted to take into consideration the notion of time preference, that is the preference for benefits earlier and costs later (Krahn and Gafni, 1993; Torgerson and Raftery, 1999). Money now is valued more highly than money in the future, so future costs must be discounted to provide a 'present value'. There is no firm consensus on the most appropriate rate to employ, although a 6 per cent annual discount rate has been approved by the UK Treasury and a 3 per cent annual discount rate is recommended by the US Panel on Cost Effectiveness (Gold *et al*., 1996). In the UK, the Department of Health recommends that costs be discounted at a rate of 6 per cent per annum, while outcomes should be discounted at a rate of 1.5–2 per cent per annum (Department of Health, 1996; National Institute for Clinical Excellence, 2001). Discounting is not necessary in studies of short duration, usually deemed to be less than one year.

The use of discounting has been criticised for biasing the outcome of the evaluation process towards specific types of intervention, particularly in relation to discounting outcomes. Examples include health promotion strategies whose benefits can fall over many years, as compared to medical interventions with more immediate health benefits (Sheldon, 1992). In general, the best approach to take is to calculate and present both discounted and non-discounted costs and benefits, and to vary the discount rate, perhaps from 0 to 10 per cent, in order to test the sensitivity of the results to changes in the discount rate.

Guide to micro-costing

Unit-cost calculation involves a number of steps. The following exposition should provide the reader with a greater understanding of the basic issues involved and the means with which to begin to calculate unit costs. The calculation of unit costs can be complex and should be done with the support of finance departments and economists, and with reference to more detailed guidelines on costing (Beecham, 2000; Beecham and Knapp, 2001, Brouwer *et al*., 2001).

Identification of service elements

To accurately cost a particular service using a bottom-up approach, a detailed description of the service is required. This involves identification and description of all the individual elements of a service, which may

include such things as buildings, staff, transport, equipment, management and administrative overheads. A detailed list of possible service elements is included in Table 1.

At the same time, it is important to identify the most appropriate unit for the calculation of costs. This unit should be relevant to the service, thus, for a home help who provides support on an hourly basis, a cost per hour would be appropriate. For a social worker whose appointments vary in length, a cost per minute may be preferable and a day-care facility that is open to clients to attend in the morning or afternoon should be costed per session.

Estimation of cost implications

Each service element identified will involve different cost implications. The cost of employing a member of staff, for example, will include not only their salary but also other associated employer costs, such as national insurance and superannuation (contributions to an employee's pension). The cost of a building will include its capital valuation (dependent on location, size, etc.), but also the cost of equipment, furniture and fittings. Table 1 summarises the cost implications of the different service elements and provides an indication of the source of cost data for each component.

Calculation of unit costs

Once all the financial information has been collected, each component must be treated appropriately since there are different conventions for calculating the cost of different service elements. Buildings, for example, are calculated on the basis of the valuation of capital, described in more detail in HM Treasury (1997). The usual method is to assume that the next best use of the resources devoted to a building is to invest the money, earning interest over the lifetime of the building, often assumed to be 60 years (Beecham, 2000). The opportunity cost of capital is therefore the stream of cash payments that would have been received over the lifetime of the building had the money been invested. The cost of staff will require salary information for the appropriate profession and

Table 1 Examples of service cost elements and sources

Service components	Information required	Source of information
Staff		
Unit head, professional staff, unqualified staff, secretarial and administrative support, cleaners, catering, drivers, ground staff and other support staff	Profession, grade, full-time equivalent, salary scale and associated employer costs (i.e. contributions to national insurance and superannuation)	Finance, salaries and wages or personnel department. Relevant trade union head office or website (e.g. UNISON, National Union of Teachers). Relevant government department or website (e.g. Department for Education and Skills, Department of Health)

Continued overleaf

Table 1 Continued

Service components	Information required	Source of information
Buildings and related costs		
Buildings	Location, size, purpose	Finance department. New-build prices (Building Cost Information Service – www.bcis.co.uk). Market resale value
Equipment, furniture, fixtures and fittings	Replacement cost	Finance department or direct from relevant providers
Power, rates	Actual expenditure	Finance department expenditure accounts or estimated percentage of salary costs
Maintenance, repairs	Actual expenditure	Finance department expenditure accounts
Contracted maintenance, cleaning, catering, etc.	Actual expenditure	Finance department expenditure accounts
Other service-related expenses		
Telephone, stationery, printing, computing, food	Actual expenditure	Finance department expenditure accounts
Transport	Number and average length of journeys or mileage, type of vehicle and replacement cost	Running costs from the finance department expenditure accounts or the Automobile Association. Replacement cost from finance department or manufacturers
Advertising, recruitment, training	Actual expenditure	Finance department expenditure accounts
Organisational overheads		
Finance, personnel, management, central information technology services, training, etc.	Appropriate proportion of these costs shared by the service in question	From finance department

Adapted from Beecham (2000).

grade as well as the costs of other salary-related expenses, usually available from the relevant accounts department. Building-related expenses, such as heating and lighting, can generally be located in the expenditure accounts of a facility.

Illustration of service-costing approaches

A study of a social work intervention for young people who have deliberately poisoned themselves provides a good illustration of the different costing methods that can be employed (Harrington *et al*., 1998; Byford *et al*., 1999). Although based in child and adolescent psychiatry, this evaluation employed social work professionals and the young people involved were found to use a wide range of different services provided by a number of sectors, including education and social services. There is a great deal of overlap between young people with mental health problems and young people in contact with social services, and thus the services used and the methods employed in this study are likely to be applicable to either group. The study employed different methods of collecting service-use data and different methods of calculating unit costs, dependent on the relative importance of each service and the information available.

Measurement of costs

Information on the use of all health, education and social services was collected from the parents at the six-month follow-up interview, using a questionnaire designed for the purpose of the study. Services were included in the questionnaire following consultation with relevant professionals and through exploration of published studies of similar child and adolescent mental health services. As little economic research had been carried out in this area before, the questionnaire was adapted through interview with users and their parents, and contained sections for 'Other services' to capture any services that had been missed. In addition to the questionnaire, an audit of medical records was carried out to verify data on hospital contacts, key services in this user population. It was assumed that the medical records would be more accurate than user recall over the six-month period of the evaluation thus, where the figures differed, data from the medical records was used.

Valuation of the cost of key services

The key service in this evaluation was the social work intervention itself. This home-based service involved an assessment session and four intensive, family-centred intervention sessions conducted by psychiatric social workers. The service was a fairly significant additional cost to be borne and thus a micro-costing approach was selected, which involved the following steps.

1. Collection of information on the time the therapists spent on face-to-face and non-face-to-face contact, including preparation, travel and supervision, using questionnaires completed by each therapist. This information was then used to calculate the total time input associated with each therapy session.
2. Collection of appropriate salary scales and employers' contributions to national insurance and superannuation from the finance department.

3 Collection of information on contractual employment conditions, including hours worked per week and holiday entitlements, from the personnel department.

4 Estimation of overhead costs (buildings, equipment, management, administration, office expenses, etc.) as a proportion of salaries provided by the finance department and also estimated from previous research into the unit costs of similar professionals (Netten and Dennett, 1996).

5 Calculation of an annual cost of each therapist, taking the midpoint of the relevant salary scales, plus employer costs and overheads.

6 Calculation of a cost per minute using the contractual information from step 3.

7 Calculation of a cost per session, using the information from therapists in step 1.

Other key services were hospital contacts, including in-patient and day-patient stays, out-patient appointments and attendance at accident and emergency departments. It was not considered feasible, in terms of time and resources, to carry out a micro-costing approach for all possible hospital contacts, of which there were many, so, instead, unit costs were collected directly from the relevant hospital departments. NHS Trusts are required to calculate annual unit costs by speciality for the services they provide, thus this information is readily available. These costs are calculated using a top-down approach, which, put rather simplistically, involves dividing the total budget of a particular speciality by the total number of out-patient attendances or overnight stays for that speciality.

Although not as accurate as micro-costing, such unit costs are likely to closely approximate the true cost and, although averaged across a large number of individuals, this makes them more representative of and generalisable to the population as a whole. By comparison, micro-costs will be more accurate for the service users involved in the study, but the generalisability of these costs will be weaker since the sample may not be representative of all users of child and adolescent hospital psychiatry services.

Valuation of the cost of other services

Other health services that were less significant, in terms of their contribution to the focus of the study or to the total costs of caring for this user group, included contacts with general practitioners, practice nurses, community psychiatric nurses and clinical psychologists. These professionals have been subject to relatively detailed costing exercises in previous research and national unit costs are available (see Box 14). Since their employment conditions and salary scales are fairly standard throughout the country, the use of national unit costs is likely to be a fairly close approximation and thus these costs were employed in the study.

Social services included contacts with social workers and time spent in foster and residential care. The unit cost of social workers has also been calculated in previous research and national unit costs were applied. The costs of foster and residential care, like hospital costs, have to be calculated annually by local authorities and thus this information is publicly

available (see Box 14) and was applied in the current study.

Services provided within schools included contacts with school doctors and nurses, education welfare officers and educational psychologists. At the time of the study, no national unit costs for these professionals were available, either from previous research or government requirements, so the only option was direct calculation. A fairly limited micro-costing was undertaken which involved collecting salary scales, information on employment conditions and employers' national insurance and superannuation contributions. This information was collected from a number of sources including the relevant trade unions, the Department for Education and health authorities, which employ school doctors and nurses. Since these services were not considered significant enough to the purpose of the study to justify detailed costing, certain shortcuts were taken and assumptions made, such as estimation of the likely grade of each professional and the likely cost of capital, managerial and administrative overheads.

Valuation of informal and voluntary care costs

Informal care is care provided by friends and family members for which no payment is made, in common with voluntary work. This does not mean that there is no cost involved, however. There is an opportunity cost of such care – time that could have been used for some alternative purpose. Valuing this 'lost opportunity' is not easy since the alternative activity may not always be obvious. In addition, it is not always straightforward with informal care to distinguish between the care provided as a result of the health or social care problems of concern and the care that would have been provided as a result of the relationship between the carer and recipient. Parents, for example, will often provide much 'caring' support to their children, irrespective of health or social needs. A number of different methods of valuing the time of informal and voluntary care have been suggested, which are summarised in Box 12, but for more detailed information readers can refer to McDaid (2001) or Brouwer *et al*. (2001).

Box 12 Methods of valuing the time of informal and voluntary carers

Market price

Time spent on informal or volunteer care is valued at the market price that would have to be paid if the work was undertaken by a formal caregiver, such as the hourly wage rate of a professional support worker. Where it is felt that the carer is less efficient than a professional, the market price could be applied to the amount of time a professional would have to spend doing a certain activity, rather than the actual amount of time taken by the informal carer.

Continued overleaf

Box 12 Continued

Reservation wage

Time is valued at the wage rate that an informal carer would have earned if their time had instead been spent in paid employment. Where the carer is already in employment, then the current wage rate can be used, although differentials in wage rates may need to be considered where, for example, a carer is sacrificing overtime for which a different rate is paid. Where the carer does not work, a wage rate of someone with similar characteristics can be used instead, or the wage rate of the carer's last place of employment. The reservation wage method more closely reflects the true opportunity cost of the carer than the market price method but is difficult to estimate.

Lost productivity

Where informal carers have to give up paid employment or reduce the number of hours worked, their time can be valued using the methods for calculating the cost of lost productivity, described in the section headed 'Valuation of productivity costs' below.

Social security benefits

The value of social security benefits paid is applied, such as unemployment benefit or income support, on the basis that these benefits can be viewed as remuneration for the caregiver's work undertaken. This approach, however, ignores the fact that benefits may only be available to a limited number of individuals and is not a measure of the true opportunity cost of the time sacrificed.

Leisure time

When leisure time is sacrificed rather than paid employment, the opportunity cost is not monetary but consists of activities that improve an individual's quality of life, such as socialising, gardening, reading, etc. Such opportunity costs may thus be better valued in terms of quality of life and can be incorporated into an economic evaluation on the outcome, rather than the cost, side of the equation. For more information on the valuation of quality of life, see Chapter 5.

Benefits to volunteering

The valuation methods described above all assume that the caregiver or volunteer is sacrificing something in order to provide care – they assume there is a cost involved. This ignores the possibility that caring brings benefits to the caregiver. Voluntary work does, after all, involve willingly sacrificing time and there is good reason to believe that volunteers get pleasure and satisfaction from the work that they do. This perhaps is another argument for valuing the time of informal carers and volunteers in terms of quality of life, where the outcomes can be positive (benefits) or negative (costs), rather than a monetary valuation that implicitly assumes a cost.

Valuation of productivity costs

Productivity costs are the costs of lost output to the economy as a whole that result from premature death or an impaired ability to work as a result of illness or disability. Productivity costs are not a direct cost of a service, and, in fact, a service that is successful in returning people to work will produce productivity savings, rather than costs, but they are still an important cost (or saving) from an economic perspective. It should be noted, however, that double counting might arise in evaluations that consider employment to be an outcome. Evaluators should ensure that issues of productivity are incorporated in terms of either costs or outcomes, but not both.

Although many guidelines to economic evaluation recommend the inclusion of productivity costs in economic evaluations, criticisms of the valuation methods used have led some authors to suggest that such costs should be excluded unless inclusion is likely to have a large impact on the results of a study (Luce and Elixhauser, 1990). Many areas of social welfare, however, are directly concerned with a person's ability to work, thus the exclusion of these costs should be considered carefully. Examples include interventions that are aimed at improving the education or employment status of individuals, such as training schemes or supported workshops, services aimed at promoting independent living, or the provision of day care for young children, freeing up parental time to engage in other activities.

Two main methods of valuing productivity costs have been suggested: the human capital and the friction cost approaches. These approaches are summarised in Box 13 but, for more information, useful guides include Koopmanschap and Rutten (1996), Goeree *et al*. (1999), Pritchard and Sculpher (2000).

Box 13 Methods of valuing productivity costs

Human capital approach

Productivity costs are traditionally valued on the basis of an individual's gross wage rate, to reflect the actual loss of productivity resulting from premature death or disability, or the gains that result from improvements in a person's well-being. This 'human capital' approach involves calculating productivity costs (or benefits) on the basis of the present value of the stream of lifetime income that an individual would otherwise have earned (or the additional lifetime income resulting from a successful intervention). Taken literally, the human capital approach implicitly values the time of children, housewives, pensioners and the unemployed at zero, which has important equity implications and indeed the time of these groups is often excluded from productivity cost calculations. The value of changes in leisure or non-paid working time, however, can be taken into consideration, for example using the reservation wage approach

Continued overleaf

Box 13 Continued

described in the previous section of this chapter. Alternatively, all individuals can be valued at the same rate, perhaps using a national average wage rate, implying that the value of one life is equal to that of all others.

Friction cost approach

The main criticism of the human capital approach is that it ignores the fact that workers absent for short periods of time can often 'catch up' on return to work or their work can be covered by other staff members. For longer periods of absence, the existence of unemployment allows workers to be replaced at little cost, thus reducing the impact of the productivity losses. For this reason, attention has turned towards a new method of calculating productivity losses – the friction cost approach. This method attempts to account for the level of scarcity in the labour market and involves adjustments to human capital estimates. Productivity losses are valued on the basis of the average time that unfilled positions are left vacant plus the costs of training and recruitment, so the loss in productivity as a result of premature death is calculated over the time it takes to replace and train a new worker. It should also be noted that workers in some positions can be replaced by an investment in capital equipment.

Sources of unit costs

As was highlighted in the above section, for some services, detailed calculations of unit costs have already been carried out in previous research and it may be appropriate to use these costs, rather than attempting an intensive, detailed costing exercise. In addition, a number of public bodies are required by the government to provide unit costs of their services and these are often publicly available. Some useful sources of unit costs are provided in Box 14.

Box 14 Sources of published unit costs

Unit Costs of Health and Social Care **(Netten and Curtis, 2002)**

This annual publication, which is now available on the Personal Social Services Research Unit website (http://www.kent.ac.uk/PSSRU), is perhaps the most comprehensive list of unit costs available and provides detailed costs in the following categories:

- services for elderly people
- services for people with mental health problems

Continued

Box 14 Continued

- services for people who misuse drugs/alcohol
- services for people with learning disabilities
- services for children and their families
- hospital and other services
- community-based health-care staff
- community-based social-care and educational-support staff
- hospital-based health-care staff
- care packages.

Personal Social Services Statistics **(CIPFA, 2002)**

This annual publication provides details of actual expenditure by local authorities on services in the following categories:

- generic services
- children and families
- elderly people
- physical or sensory disabilities
- learning disabilities
- mental health needs
- other adult clients.

Health Service Financial Database **(CIPFA, 2001)**

This database, available annually on CD-ROM, contains information from the annual financial returns of NHS Trusts and includes unit costs by hospital speciality for the following services:

- in-patient day
- in-patient episode
- out-patient attendance
- day-care attendance.

Summary of steps to follow when costing services

Unit costing is not a straightforward process and the methods used to measure and value the costs of services can vary depending on the type of cost of interest (i.e. programme costs, informal care, productivity losses, etc.), the resources available to the evaluation and the importance of the cost element to the purpose of the evaluation. Box 15 provides a brief summary of the steps to follow and the choices to be made when undertaking a costing exercise.

Box 15 Summary of steps to follow when costing services

Identify and describe cost components through:

- consultation with users
- consultation with professionals
- pilot or feasibility studies
- previous research/literature.

Attribute cost components through:

- over-inclusion
- over-exclusion
- expert opinion.

Measure cost components using:

- questionnaires
- service-user diaries
- case notes
- previous research/literature.

Value cost components:

- select an accounting year
- prioritise cost components as key to the evaluation or less so
- bottom-up calculation for key components
- top-down calculation or published unit costs for less important components
- select a method of valuing informal or voluntary care if appropriate
- select a method of valuing productivity losses if appropriate
- discount costs if necessary.

5 Outcome measurement and valuation

Outcome can be defined as 'the effectiveness of an activity in relation to the achievement of the intended goal' (Bowling, 2002). Thus, the types of outcome that an evaluation chooses to focus on will depend to a large extent on the objectives of the service and the question to be answered by the research. In relation to social care services, the primary objective will often be to have a positive impact on the service user's life, either in terms of a specific dimension of life, such as social functioning, or in terms of broader concepts of the overall quality of a person's life. Other outcome types may also be appropriate, however.

The process of outcome measurement follows a similar pattern to that of cost calculation, described in Chapter 4. The first step is the identification of outcomes that are relevant to the evaluation being undertaken, which involves clarification of a programme's objectives through, for example, discussions and focus groups with policy makers, funders, service providers and users. Volume I of this report (Sefton *et al*., 2002) provides a detailed discussion of the broad issues involved in outcome measurement, covering such things as clarification of a programme's objectives, choosing the right outcomes, dealing with multiple outcomes and defining outcomes. Such issues will not be discussed further here.

Types of outcomes

Process and intermediate outcomes

Process outcomes focus on services, rather than users, and are expressed in terms of level of provision, throughput and quality of care. They are concerned with service performance or activity, such as days in local authority care, contact with social workers, or truancy levels at school. Intermediate outcomes are more user-focused measures of easily observed changes, such as employment status, educational attainment, acquisition of new skills, or days of disability. Both process and intermediate outcomes are commonly used in social welfare and beyond, often because they are relatively straightforward to observe and to measure. In addition, they provide valuable information on whether or not a service is working in the way in which it was intended and can help to determine which service characteristics are driving good or bad user outcomes. Volume I of this report (Sefton *et al*., 2002) provides a more detailed discussion of the importance of exploring how and why a programme works, and the usefulness of intermediate and process measures of outcome.

Many evaluations are purposely focused on intermediate or process outcomes because the funding body of a service will often wish to be convinced that a service is being carried out in the way it was intended. One example is the ongoing evaluation of healthy living centres (HLCs). HLCs use a variety of different approaches to reduce social exclusion and isolation, and improve the health of the most deprived members of society. In this study, intermediate and process outcomes are an important focus of the initial evaluation, such as contact time with HLC staff, individual participation in programmes, or smoking cessation rates. HLCs have been established only recently and, initially at least, it is difficult to measure their overall impact on health, health inequalities, social exclusion or long-term quality of life, since it may take a number of years to observe such benefits. Such longer-term

outcomes of HLCs will be evaluated in future research.

Intermediate and process outcomes are, however, proxy indicators of user outcome that involve an assumption regarding the benefit to the user. Although, from the perspective of a service provider, these indicators may seem to provide a measure of success, changes in these outcomes may be of only limited benefit to users and may in fact be detrimental. For example, an evaluation of early hospital discharge schemes for older people may be concerned primarily with enabling people to return home to a more independent lifestyle. Although a majority of older people may value highly the ability to live in their own homes, others, perhaps those who have few relatives to care for them or a greater fear of isolation, may prefer to remain within an institutional environment and thus their quality of life may fall once in the community. In addition, some process measures of outcome are inputs into a service and therefore a cost of that service. Including them as outcomes could thus involve double-counting, which should be avoided.

Final outcomes

By contrast, final outcomes are concerned with changes to an individual's life rather than changes in the service or other proxy indicators of outcome. Examples include the quality of life or the quality of individual aspects, such as social or family functioning. The focus of final outcomes is on the ability of a service to meet its final objectives, such as improvements in physical and mental health, social functioning, well-being or quality of life. In addition, the measurement of final outcomes enables much broader comparisons to be made across different types of intervention, thus better supporting resource allocation decisions.

An important objective of all social welfare interventions is to improve the quality of people's lives and, since economics is concerned with maximising well-being, economic evaluations should ideally include measures of overall well-being. In practice, such an ideal may not always be possible. Outcomes may not be observable for many years. For example, the full effects of Sure Start, an initiative that aims to improve the health and well-being of families and children under four (Sure Start, 2002), may not be seen until adulthood. In addition, the impact on quality of life may sometimes be expected to be small and diluted across a large population, such as the impact of community centres or other small-scale community initiatives to enhance urban regeneration. Where quality of life is unlikely to be an adequate indicator of change, other outcomes will need to be measured.

The use of final measures of outcome should not preclude the use of intermediate and process measures. Since intermediate and final outcome measures serve different purposes, more can be learnt from studies that include a range of these measures than from studies that focus exclusively on intermediate and process measures or studies that focus exclusively on final measures of outcome.

Condition-specific outcomes

Two main categories of final outcome exist: condition-specific and generic. In health care, condition-specific or disease-specific scales are used to measure, in natural units, the outcomes of treatment that are specific to a particular disease or health problem. They are used mainly

to classify a disease and monitor its progression over time. For example, depression scales will generally measure severity of depression and change over time. Similar condition-specific, client-specific or need-specific outcomes, such as social exclusion or social functioning, are measured in social welfare research. Such measures are generally sensitive to changes in the problem under investigation and can be used in cost-effectiveness analyses. They are limited, however, by their inability to capture all the potential impacts of an intervention, of particular concern in areas where outcomes can be multiple, and the impossibility of comparing diverse interventions, required for a societal perspective on resource allocation.

Generic outcomes

Although it is possible to explore a range of condition-specific outcomes in an attempt to capture all possible impacts of an intervention, it will still be difficult to determine a single, overall impact of an intervention if users perform well on some outcome dimensions, but not on others. Generic outcomes, used in cost-utility analysis, attempt to capture all aspects of the quality of a person's life and are thus more widely applicable than condition or service-specific outcomes and can be more broadly compared across services and conditions. Generic outcomes are commonly measured in health care, but are much rarer in other areas of social welfare. They fulfil more of the criteria for an economic evaluation and, since improving an individual's quality of life is central to both social welfare and health care, efforts to measure such well-being should be encouraged. Generic outcome measures will not necessarily be as sensitive as condition-specific measures to small or specific changes in the condition of interest, but, used alongside more detailed condition-specific scales, assessments of their accuracy and sensitivity can be made.

Measurement of outcomes

The next step is the measurement of outcomes selected and, as with the measurement of costs described in Chapter 4 in the section headed 'Measurement of cost components', a number of different methods exist, including routine service records, diaries, qualitative interviews and quantitative methods. The method chosen will depend on the type of outcome measure of interest, the time and resources available, and the methodological perspective of the evaluator. As with the measurement of service use, studies can record such information prospectively or retrospectively, with prospective measurement being preferred since it allows the researcher greater control over the data collected and removes problems of recall and lost or poorly kept service records or case notes.

Service records

Process and intermediate outcomes tend to be objective rather than subjective in nature (with a few notable exceptions, such as quality of care) and are thus relatively straightforward to measure. For this reason, they are sometimes described as 'hard' measures. The measurement of days in residential care or job interviews attended, for example, is a simple task that is easily incorporated into the everyday recording systems of services. As with service-use data, records are likely to be more accurate than relying on user recall over a significant period of time, but record searches can be time

consuming, may not record exactly the information needed and will often be hampered by poor completion, missing files and illegible entries.

Diaries

More user-focused intermediate measures, such as activities of daily living, can also be recorded relatively easily, through user diaries, for example. Diaries are a prospective method of data collection that users can complete on a daily basis, recording perhaps their activities or the hours of support provided by family members. They are reliant on user co-operation and need to be relatively simple to complete, as discussed in Chapter 4 in the section headed 'Measurement of cost components', but are a potentially rich source of information.

Qualitative interviews

Final outcomes are more subjective in nature than intermediate outcomes and thus are less amenable to direct measurement. To gain more understanding of subjective final outcomes, qualitative interviews can be undertaken to explore an individual's perceptions of the changes that have resulted from an intervention. Social welfare research is dominated by qualitative measurement of outcomes, for many good reasons. Qualitative analysis, discussed in more detail in Volume I (Sefton *et al*., 2002), provides valuable insights into the context in which a programme is delivered and the mechanisms through which a programme is able to deliver good or bad outcomes. Compared to quantitative research, which dominates economic evaluation, qualitative research provides a more detailed picture of how, why and for whom a programme works.

Quantitative interviews

The results of qualitative analyses cannot easily be combined with costs in order to undertake an economic evaluation, so there is a strong preference for quantitative evaluation. This preference is also driven by the desire for statistical validity – the ability to subject differences in costs and outcomes to statistical tests to determine the probability that the differences found in the sample exist in the general population. To convert subjective final outcomes into quantitative units within an evaluation, pre-existing outcome scales can be used (see the following section of this chapter) and these are often selected for use in a cost-effectiveness analysis. To undertake either a cost-utility analysis or a cost-benefit analysis, outcomes need to be valued in terms of utility (see the section headed 'Measurement and valuation of quality of life' later in this chapter) or in monetary terms (see the section headed 'Monetary valuation of outcomes' later in this chapter), respectively.

Outcome scales

Outcome scales (or research instruments) are tools for measuring the variables of interest and will generally take the form of a questionnaire. As with service-use questionnaires, outcome scales can be self-reported or completed by researchers at interview. There exist a daunting array of standardised outcome scales developed in previous research to answer different questions of different populations in different contexts and countries, which may be appropriate for use in current research. It should be remembered, however, that, despite the large number of existing scales available, a detailed

search of the literature might still conclude that none is appropriate for the present research and instead a new scale must be developed. In particular, there are far fewer generic than condition-specific scales and many of those that do exist were developed for use in health-care research. Thus, they are often health focused and may not be broad enough to capture the full range of impacts of a social welfare intervention.

Appropriateness and acceptability of outcome scales

To locate appropriate scales for measurement of particular variables, an initial search of the literature should be undertaken and each scale located must be assessed to determine its appropriateness and acceptability to the research questions of interest. A number of questions must be asked of the research instrument and these are summarised in Box 16. Both pre-existing scales and newly developed scales should be tested to ensure they fulfil the criteria listed.

An outcome scale must measure all the dimensions of interest to the research and must be appropriate to the study population. A scale that has been developed for use with children, for example, may not be appropriate when applied to adults. Similarly, a scale that was originally developed for use in developing countries may not be appropriate for use in developed countries. This leads us to the issue of translation of outcome scales – if an instrument has been translated from one language to another, researchers must ensure that the conceptualisation and meaning of the scale are not lost, and that language and terminology are appropriate to the different population. An example of the processes involved in the translation of outcome scales is provided by Knudsen *et al*. (2000).

The scale must also be acceptable both to the population of interest and to the research. A self-completed postal questionnaire, for example, may not be well received by a frail elderly population, nor will complex questionnaires or those with small print. Instead, a questionnaire given in interview would be more appropriate. On the other hand, respondents of questionnaires exploring certain sensitive subjects, such as sexuality or drug-taking behaviour, may give more honest responses in self-completed questionnaires than to an interviewer. The acceptability of the scale to the research will depend on the resources available to the research team. Evaluators must assess whether they have adequate resources to cover such costs as printing, administration, interviewer time, data entry and data analysis.

Level of measurement and statistical analysis

As mentioned earlier in this chapter in the section headed 'Measurement of outcomes', there is a preference among economists for quantitative data that are capable of being subjected to statistical tests. There are four levels of data measurement – nominal, ordinal, interval and ratio – and each has different implications for statistical analysis, with the more sophisticated level of data allowing the more rigorous statistical techniques. Each level of data is briefly summarised below and readers are referred to Chapter 6 in Bowling (2002) for more detailed information on the statistical techniques available.

- *Nominal data*: with nominal (or categorical) data, numbers are simply used to classify different responses into discrete categories, such as 1 = white, 2 = Asian, 3 = black, 4 = other ethnic group. There is no statistical relationship between the different responses, so, for example, they cannot be added together, multiplied or averaged and they cannot be ranked with one item being 'better' or 'higher' or 'worse', etc. For this reason, nominal data are the least sophisticated and, thus, so are the statistical tests available to the researcher.
- *Ordinal data*: ordinal data are also categorical but there is some relationship (or scaling effect) between the individual items. In other words, the items can be grouped or ranked and more powerful statistical tests appropriate for ranked data can be applied. For example, when asked how helpful a keyworker has been, responses may be categorised as: 1 = always helpful, 2 = mostly helpful, 3 = sometimes helpful, 4 = mostly unhelpful, 5 = always unhelpful.
- *Interval data*: an interval scale is equivalent to an ordinal scale but, in this case, the distances between any two numbers on the scale correspond to equal and known differences in the underlying dimension that the scale is meant to measure. In other words, the distances between two points are meaningful. The scale on a thermometer, for example, has interval properties. For interval data, more powerful statistical tests that make fuller use of the data available can be applied.
- *Ratio data*: a ratio scale is equivalent to an interval scale but it additionally has a 'true' zero point, rather than an arbitrary one. Weight, for example, can be measured on a ratio scale. As the most powerful level of data, the most powerful statistical tests can be applied to ratio scales.

Quality of outcome scales

When selecting an outcome scale, researchers should also explore whether the scale has been tested in previous research to ensure that the scale fulfils certain 'quality' criteria, namely reliability and validity. These criteria are described below and more detailed information can be found in Bland and Altman (2002) and Chapter 6 in Bowling (2002). Newly developed scales and existing scales of uncertain quality must always be tested against these criteria to ensure they are adequate for the purpose of the study.

Reliability measures the degree to which the results of a test are reproducible if repeated. There are three main measures of reliability: inter-rater reliability, test-retest reliability and internal consistency:

- *Inter-rater reliability*: refers to the level of agreement obtained when the same sample is assessed under the same conditions by two different 'raters'. Inter-rater reliability tests involve an assessment of the correlation (association, relationship) between the scores obtained

from two or more raters. A number of statistical tests to measure correlation exist, such as Pearson's correlation, and, generally, a score of 0.8 or more (where 0 = no correlation and 1 = perfect correlation) would be considered an indication of inter-rater reliability (Bowling, 2002).

- *Test-retest reliability*: refers to the stability of the results when carried out by the same rater at different time points. Again, tests of correlation are used to assess the level of association achieved.
- *Internal consistency*: refers to the level of homogeneity of the items in an outcome scale. In other words, the extent to which an individual item records information on the dimension of interest and only the dimension of interest.

Validity, specifically internal validity, refers to the degree to which an instrument measures what it is intended to measure in the target populations. Outcome scales need to be tested for a number of different types of validity, of which the main ones are as follows.

- *Face validity*: refers to the subjective assessment of the relevance of the outcome scale to the purpose that it is designed for.
- *Criterion validity*: when the outcome scale is found to correctly predict certain criteria in comparison to another measure which is accepted as valid.

Box 16 Summary of criteria to assess outcome scales

- Does the outcome scale measure dimensions relevant to the current study?
- Does the outcome scale measure dimensions that are important to the respondents?
- Is the outcome scale responsive to change within the chosen study period?
- Is the outcome scale appropriate for the population of interest?
- Is the outcome scale acceptable to the study population?
- Is the administrative burden of the outcome scale acceptable, given the resource limits of the research?
- Do general population scores exist for comparison?
- What level of data will the instruments relate to and is this appropriate for the planned statistical analyses?
- Have the reliability and validity of the outcome scale been tested? If yes, have they been tested on appropriate populations?

Adapted from Bowling (2002).

- *Construct validity*: refers to the extent to which the outcome measure consistently tests the hypothesis it is designed to measure and demonstrates appropriate relationships with other variables.
- *Content validity*: relates to the extent to which the coverage of the outcome measure appears to incorporate the full scope of the dimension it is intended to measure.

Existing outcome scales

Boxes 17 to 19 provide examples of the range of outcomes that can, and have been, measured in different areas of social welfare and an idea of some of the outcome scales that have been used. This list is for illustration only. It is by no means an exhaustive list and readers are reminded that decisions regarding which outcomes to measure and which outcome scales to include in an evaluation should be made only after

Box 17 Examples of outcomes in the criminal justice sector

Process or intermediate outcomes

- Reoffending rates, i.e. number of contacts with criminal courts or reconviction rates (Miers *et al*., 2001; Trulson *et al*., 2001).
- Reparation/compensation rates, i.e. assessment of the proportion of agreed financial compensation paid (Presser and Van Voorhis, 2002).
- Sociometric checklists of participation in and level of group dialogue in restorative justice, i.e. programmes that attempt to repair the damage caused by crime, often involving meetings between victim and perpetrator (Berg and Rounds, 1992; Presser and Van Voorhis, 2002).

Final outcomes

- Level of victim's fear of being revictimised by the same offender (Umbreit and Coates, 1993).
- Satisfaction with restorative justice or other criminal justice programmes (Umbreit and Fercello, 1997).
- Change in victim and offender attitudes (Coates and Gehm, 1989; Umbreit, 1994).
- Social disorganisation (Warner and Pierce, 1993).
- Sense of community or 'collective efficacy' (Sampson *et al*., 1997).
- Defining Issues Test, a measure of the development of empathy/moral responsibility for harms done (Rest, 1979).
- Criminal Sentiments Scale and Pride in Delinquency Scale, to assess change in attitudes in criminal behaviour (Simourd, 1997).

discussions with key parties to the research and extensive review of the literature.

Researchers must also carefully review the reliability and validity of any existing outcome scales proposed. To ensure the responses gained are useful, it should be remembered that existing scales are designed for a particular purpose and many have been tested to ensure that the criteria for an instrument, discussed above, have been met. It is, therefore, inappropriate to amend the scales (e.g. alter the wording or add or remove items) without re-testing the validity and reliability of the scales.

Box 18 Examples of outcomes in housing/urban regeneration

Process and intermediate measures

- Crime and vandalism rates (Power and Bergin, 1999; Lupton, 2001; Richardson 1999).
- Occupancy rates (Power and Bergin, 1999; Richardson, 1999).
- Proportion of derelict land (Lupton, 2001).
- Unemployment and employment rates (Lupton, 2001; Richardson, 1999).
- Level of community involvement in decisions and priority setting (Power and Bergin, 1999).
- School absence and exclusion rates, levels of literacy, qualifications, etc. (Lupton, 2001; Richardson, 1999).

Final outcomes

- Sense of pride and commitment to the area (Power and Bergin, 1999).
- Tenant sense of security (Power and Bergin, 1999).
- Tenant satisfaction (Power and Bergin, 1999).
- Mortality rates (Lupton, 2001).

Box 19 Examples of outcomes in social care

Process and intermediate measures

- Educational attainment and employment status (Minty, 1999).
- Admissions to local authority care, return home, placement breakdowns (Minty, 1999).
- Home Conditions Scale, a measure of the condition of the home environment (Davie *et al*., 1984; Cox and Bentovim, 2000).
- Family Activity Scale, a measure of the environment provided for children (Cox and Bentovim, 2000).

Continued overleaf

Box 19 Continued

- Recent Life Events Questionnaire (Brugha *et al*., 1985; Cox and Bentovim, 2000).
- Family Assessment Device, a measure of family functioning (Miller *et al*., 1985).
- Eyberg Inventory, a measure of parental report of their child's behaviour (Eyberg and Ross, 1978).
- Parenting Scale, a measure of dysfunctional parenting (Arnold *et al*., 1993).
- Barthel Activities of Daily Living Index (Wade, 1992); Activities of Daily Living Index (Katz *et al*., 1963; Challis and Davies, 1986).
- Interview Schedule for Social Interaction (Henderson *et al*., 1980).

Final outcomes

- Philadelphia Geriatric Center Morale Scale (Lawton, 1975; Challis and Davies, 1986).
- The Strengths and Difficulties Questionnaire, a measure of emotional and behavioural problems in children and adolescents (Goodman, 1997; Cox and Bentovim, 2000).
- The Parental Daily Hassles Scale, a measure of parental stress (Cox and Bentovim, 2000).
- Adult Wellbeing Scale, a measure of depression, anxiety and irritability (Snaith *et al*., 1978; Cox and Bentovim, 2000).
- General Health Questionnaire, a measure of psychological well-being (Goldberg, 1972).
- Adolescent Wellbeing Scale, a measure of depression in young people (Birleson, 1980; Cox and Bentovim, 2000).
- Needs and unmet needs (Challis and Davies, 1986).

Measurement and valuation of quality of life

As discussed in Chapter 3, cost-utility analysis is often advocated for economic evaluations because of the ability to capture multi-dimensional outcomes on one scale and to allow comparisons across a broad spectrum of interventions. However, appropriate utility-based measures of outcome are lacking in the social welfare field. Thus, to undertake a cost-utility analysis, at least in the short term, researchers may need to directly measure and value such generic outcomes. This section explores the measurement and valuation of utility or quality of life, while the following section describes methods of measuring outcomes in monetary terms, for use in a cost-benefit analysis.

To generate any outcome measure, the first step is to decide on the relevant dimensions to incorporate. Dimensions that may be considered important to the development of a quality of life scale capable of capturing all elements of relevance to social welfare are listed in Box 20.

The next step is to describe the different 'states' on each dimension (e.g. different levels of dependence on informal carers, different degrees of child–parent attachment or different levels of disability) encompassing the range of possible outcome states that are relevant to the client group of interest. A hypothetical outcome state, covering the dimensions listed in Box 20, is contained in Box 21 for illustrative purposes. Developing the description system for an

Box 20 Dimensions of quality of life

- Social functioning (e.g. relationship with friends, ability to work).
- Psychological functioning (e.g. mental health status, mood and feelings).
- Physical functioning (e.g. activities of daily living, disability).
- Family functioning (e.g. relationships with family members).
- Dependence level (e.g. informal care and service support needs).
- User satisfaction (e.g. with services).
- Carer burden (e.g. quality of life of carers).
- Community development (e.g. social exclusion, deprivation, poverty).

Box 21 Hypothetical description of an outcome state

- You get involved in social activities once or twice a month.
- You are currently employed but only on a temporary basis.
- You feel depressed and/or anxious most days.
- You suffer from moderate pain and/or discomfort most days.
- You have no physical difficulties in relation to getting around, self-care or undertaking your usual activities (e.g. work, housework, social activities).
- You see your immediate family two or three times a year.
- You are not dependent on others for help in getting around or undertaking your usual activities (e.g. work, housework, social activities).
- You feel moderately secure in your own home.
- You feel no commitment to the community in which you live.

outcome measure will often begin with theoretical hypotheses that are refined with the help of professionals, users and carers, or the wider general public, through focus groups and interviews. Thus, qualitative research has a large part to play in the development of such classifications.

Once the outcome states have been described, they must be graded according to how good or bad they are considered to be. The form of this grading will depend on the data type selected, i.e. nominal, ordinal, interval or ratio (described in the previous section). An ordinal ranking, for example, would simply involve ranking the different outcome scales from best to worst. Generic outcome scales suitable for use in cost-utility analyses are generally interval scales, such that the distance between any two points on the scale represent differences in the underlying dimension the scale is intended to measure that are of equal and known size, and a number of different valuation methods capable of generating an interval scale have been employed. Three commonly used methods of valuation are described in Box 22, although this list is not exhaustive. More detailed information can be found in Kind (1988), Chapter 2 of McGuire *et al*., (1992), Dolan (2001) or Chapter 4 of Bowling (2002).

There are a number of possible groups of respondents to provide the valuations for outcome states, including users, carers, professionals or the general public and the choice will often depend on the purpose of the evaluation. An outcome scale that is designed for use by service providers to monitor changes in a condition over time as a result of the service provided would be better to focus on valuations provided by service users. Broader scales that

Box 22 Methods of valuing outcomes

Standard gamble

Standard gamble valuation involves asking respondents to choose between the certainty of a particular outcome state that is less than perfect or a gamble that could produce either the best or the worst imaginable state. For example, respondents may be asked to choose between a guaranteed ten years with limited mobility or a gamble that may produce ten years with perfect mobility or immediate death. Probabilities are attached to the gamble and these are varied until the respondent is unable to choose between the gamble and the certainty of remaining in the intermediate state. Thus, a respondent may choose to take the gamble when the probability of death is very low, perhaps 5 per cent, but is more likely to choose the certainty of the intermediate outcome state when the probability of death in the gamble is high, say 90 per cent. At some point between these two probabilities, the respondent will become 'indifferent' between the two options and this probability is the value of the intermediate outcome state.

Continued

Box 22 Continued

Time trade-off

Time trade-off involves asking respondents to choose between an outcome state that is less than perfect for a specified length of time or being in the best imaginable state for a shorter period of time, followed immediately by death. For example, 20 years with limited mobility or 15 years with perfect mobility. The period of time spent in the best imaginable state is varied until the respondent is indifferent between the two options – in other words, the respondent considers the two options to be equal in terms of the utility gained.

Rating scale

Respondents are asked to rank outcome states in order of preference and relative to each other and to two anchor states: perfect quality of life (score = 1) and death (score = 0). They are provided with a visual 'thermometer' graded from 1 to 0 where 1 represents the most preferred outcome state (usually perfect health or perfect quality of life) and 0 represents the least preferred outcome state (often death). They are then asked to locate a number of different intermediate outcome states along the length of the thermometer in order of preference and relative to each other and to the two anchor states, 0 and 1. In other words, outcome states should be placed in such a way that the spaces between the different states reflect the differences in the strength of preference.

aim to measure quality of life for use in evaluations to support resource allocation decisions, however, may prefer to focus on the valuations of the general public since it is this group that may benefit from the services that receive government funding and it is this group who pay for the services, through taxation. There is some concern, however, that a public perspective may discriminate against minority groups in society (Sassi *et al*., 2001).

A useful illustration of the process involved in developing a utility-based outcome measure is provided by Bennett *et al*. (2000) who describe the development and testing of a utility measure for depression, the McSad. Among other things, the authors describe the process of outcome state identification and classification, pilot testing, survey methods, scaling methods, statistical analysis, acceptability and quality, outcome state valuation and reliability of the scale.

Monetary valuation of outcomes

An alternative approach is to value outcomes in monetary terms, the approach taken in cost-benefit analysis. There has been comparatively little use of monetary outcomes in most areas of social welfare including health, but the techniques are commonly used in other areas of appraisal, notably transport (e.g. assessing the value of building a new bypass) and

environment (e.g. putting a value on an anti-pollution programme). Despite the limited use of this approach, its principal advantage is that outcomes across all sectors can be compared, which can help to determine whether it is better to invest in programmes in very different areas of public policy. For example, it may be difficult to identify a common, non-monetary measure of outcome for comparing a neighbourhood regeneration programme, where the principle impact may be on community development, with an early years intervention for children, where the principle long-term outcomes may be educational attainment.

There are two principle approaches to valuing outcomes in monetary terms: identification of actual examples where individuals have placed a monetary value on the benefits associated with a service (revealed preferences); or, in the absence of such information, presenting individuals with hypothetical situations in order to obtain monetary values (contingent valuation). These two approaches are discussed briefly in Box 23. For a more detailed discussion of monetary valuation of outcomes see Perkins (1994), Chapter 7 in Drummond *et al*. (1997), McIntosh *et al*. (1999) and Ryan *et al*. (2001).

Box 23 Methods of valuing outcomes in monetary terms

Revealed preference approaches

- *Hedonic pricing*: widely used in environmental appraisal, hedonic pricing uses the price of land and property to identify the additional amount of money individuals are willing to pay to have access to a specific intervention or to live in a certain area. For example, house prices may be higher in an area that has a good supply of nursery schools or a good crime prevention rate. Statistical techniques are used to identify other factors that contribute to house prices and any remaining unexplained difference in prices between, for example, an area with a neighbourhood watch scheme and one without, can be assumed to reflect the value that society places on such schemes. In practice, this approach can be very data intensive, requiring detailed information on a large number of properties, in order to arrive at a plausible monetary valuation.
- *Time travel method*: this method explores how far people are willing to travel to use a service. This has typically been used to value national parks or museums, but it could also be applied to social welfare interventions and has, for example, been used to place a value on community-based health-screening programmes. (Clarke, 1998). The total estimated value would include any direct charges for using the service, direct travel costs (e.g. petrol, bus fares, etc.) and the value of time spent travelling (see the section headed 'Valuation of informal and voluntary care costs' in Chapter 4 for a discussion of methods used to value time).

Continued

Box 23 Continued

- *Jury compensation method*: awards made by juries to compensate for injury, criminal damage and negligence have also been used to reflect societal monetary values. This method has predominantly been used in the USA, for example to value the costs of criminal injury in a recent cost-benefit analysis of a community-based drug-treatment intervention (Rajkumar and French, 1997).

Contingent valuation methods

- *Willingness to pay/accept*: revealed monetary valuations for an outcome may not be available, or they may be considered too limited to incorporate all factors associated with a service, for instance their psychological impact. In this situation, the willingness to pay approach can be used to estimate monetary outcomes. Typically, an intervention is first described and detailed information on effectiveness and consequences collected. For example, a local community-based crime-prevention programme may involve improved street lighting and increased police patrols, and it may have been shown to reduce the likelihood of residents being mugged or burgled. Individuals are then asked what would be the maximum amount of money they would be willing to pay to introduce the crime prevention programme, or, conversely, how much financial compensation they would require in order to live in an area without such a programme. Theoretically, willingness to pay and willingness to accept should produce the same value; in practice, however, valuations will often differ, in part because individuals are influenced by their budget when deciding how much to pay, but have no such constraint when deciding how much to accept in compensation. Individuals can have difficulty estimating their willingness to pay using open-ended questions. Some contingent valuation surveys have dealt with this issue by asking whether or not an individual would be willing to pay £*X* for an intervention and getting yes/no responses. Different groups are given different monetary values and an overall willingness to pay is calculated using statistical techniques on the basis of these responses. Another approach is to conduct an auction to estimate maximum willingness to pay. For example, if an individual indicated a willingness to pay £500 for an intervention, she (or he) would then be asked if she would be willing to pay a higher amount e.g. £1,000. Similarly if she had not been willing to pay £500, the price would have dropped and the question would have been asked again.

- *Conjoint analysis*: conjoint analysis has long been used in marketing to show how individuals are willing to trade between different aspects of an intervention based on their individual tastes and preferences. Recently, this has been used in economic evaluations in health care (including community-based health-promotion schemes) as an alternative way of estimating

Continued overleaf

Box 23 Continued

either utility or willingness to pay. Essentially, conjoint analysis involves presenting individuals with a series of scenarios related to an intervention, which they then rank in order of preference. Each scenario will contain a number of different attributes associated with an intervention, including, in the case of cost-benefit analysis, a monetary cost that varies according to the scenario. Statistical techniques are then used to calculate willingness to pay. For a fuller discussion of conjoint analysis see Ryan (1999), Ratcliffe (2000) and Ryan and Farrar (2000).

Summary of steps to follow when measuring outcomes

The methods used to measure and value the impact of services can vary depending on the purpose of the intervention, the evaluation question, the availability of suitable existing measures and the resources available to the evaluation. Box 24 provides a brief summary of the steps to follow and the choices to be made, once the service objectives and evaluation question have been clarified.

Box 24 Summary of steps to follow when measuring outcomes

Types of outcomes

- Process: service-focused outcomes.
- Intermediate: user-focused, readily observed outcomes.
- Final – user-focused, subjective outcomes – preferred for economic evaluation.

Measurement of outcomes

- Service records.
- Service-user or service-provider diaries.
- Qualitative interviews.
- Quantitative interviews – preferred for economic evaluation.

Selection of existing outcome scales

- Literature search for relevant scales.
- Assess the appropriateness and acceptability of the scales.
- Assess the appropriateness of the level of measurement for statistical analysis.
- Assess the quality of the scales – reliability and validity.

Continued

Box 24 Continued

Development of new scales

- Identify and describe the relevant dimensions of the outcome of interest.
- Describe the alternative outcome states.
- Select a method for valuing the alternative outcome states.
- Select a method for valuing the outcome states in monetary terms, if relevant.
- Select respondents to provide valuations – users, carers, professionals, public.

6 Presentation and implementation of results

Once the evaluation has ended and adequate data have been collected, the relative cost-effectiveness of the alternative interventions must be determined before the results can be presented to decision makers. This chapter describes the decision rules for assessing relative cost-effectiveness, dependent on the method of economic evaluation chosen. Then it discusses how to present the results in a way that ensures clarity and enables readers to accurately assess the implications of the results. Finally, it discusses the implementation of the results of economic evaluations.

Decision rules

To determine the relative cost-effectiveness of the alternative interventions, and hence inform policy decisions, rules are needed to enable decision makers to clearly understand under what circumstances a service can be considered more cost-effective than an alternative. The processes involved will be determined by the method of economic evaluation selected. The decision rules in cost-consequences, cost-benefit and cost-minimisation analyses are, in principle, straightforward. For cost-effectiveness and cost-utility, however, the synthesis of costs and outcomes becomes more complex and these will be dealt with last.

Decision rule for cost-consequences analysis
The evaluator undertaking a cost-consequences analysis is faced with a range of outcome measures to explore in relation to cost. In such a situation, it is not necessary to formulate the decision rule, but the evaluator can present information to help guide the decision maker towards a decision. Both the costs and the range of consequences measured are presented separately and the decision maker is allowed to make resource allocation decisions on the basis of their own preferences or weighting system for the different outcomes measured, taking into consideration all relevant evidence. These preferences might be based on government or other priorities, the objectives of the providing agency, etc. and may take into consideration evidence other than economic evidence, such as issues of equity which are discussed in more detail later in this chapter in the section headed 'Presentation of the results'. The evaluator may provide guidance on the relative strengths and weaknesses of the alternative effectiveness measures, or attempt to prioritise the measures, but cannot come to any firm conclusions regarding cost-effectiveness unless better outcomes are demonstrated by one group on all measures.

Decision rule for cost-benefit analysis
In cost-benefit analysis, if the benefits of an individual intervention exceed the costs of providing it then the intervention should be funded since there is a net gain to society, subject to existing budget constraints. Where two or more interventions are compared, resources should be directed towards the intervention that produces the greatest net benefit (benefits minus costs). No complex comparisons are required since both costs and outcomes are valued in the same units. The

simplicity of this decision rule makes cost-benefit attractive and, in the absence of appropriate quality of life scales, the difficulties of valuing outcomes in monetary terms should be weighed carefully against the difficulties of direct quality of life valuation and the development of new quality of life scales. The use of cost-benefit analysis in social welfare research should not be dismissed.

Decision rule for cost-minimisation analysis
Given evidence of equal outcomes, a cost-minimisation analysis involves assessment of the relevant costs alone and thus the decision rule simply involves selection of the least cost alternative. The cheapest option is the most efficient use of resources since fewer resources are needed to bring about the same improvement in outcomes.

Decision rule for cost-effectiveness and cost-utility analysis
Combining costs with outcomes in cost-effectiveness or cost-utility analysis is more complex. Assuming a single or primary outcome measure in a CEA, eyeballing the data from a CEA or CUA will tell you whether or not one intervention is cheaper and whether or not it is more effective than the alternative. Four possible scenarios can arise.

- The new intervention is more effective and less expensive than the existing intervention, thus the new intervention is relatively more cost-effective and is said to *dominate* the existing intervention.
- The new intervention is more expensive and less effective than the existing intervention, thus the existing intervention is relatively more cost-effective and is said to *dominate* the new intervention.
- The new intervention is more effective but also more expensive than the existing intervention and there is a need to assess whether the additional cost of the intervention is justified by its additional effectiveness – there is a *trade-off* between the two interventions.
- The new intervention is less expensive but also less effective than the existing intervention and there is a need to assess whether the loss in effectiveness can be justified by the costs saved – there is a *trade-off* between the two interventions.

When the results of an evaluation involve a trade-off between costs and effects (one intervention is more effective and more costly), an incremental cost-effectiveness ratio (ICER) is calculated. An ICER is the ratio of the additional costs to the additional effects produced by one intervention in comparison to another, as illustrated in the equation below. Higher cost-effectiveness or cost-utility ratios indicate lower cost-effectiveness. Incremental analysis is required, as opposed to a simple ratio of total costs to total effects, because economic analysis is concerned with how much we are paying for each extra unit of effectiveness by undertaking the new intervention.

But what does this ICER tell us? The new intervention has been found to be more effective but without further information it is still not possible to determine whether or not society

$$ICER = (C_N - C_C) / (E_N - E_C)$$

Where C_N = cost of the new service
C_C = cost of the comparison group
E_N = effectiveness of the new service
E_C = effectiveness of the comparison group

considers this additional benefit to be worth the additional costs that would have to be expended. To do this, an external value system is needed – something to compare the ICER with. In health economics, the idea of a 'cut-off point' or 'ceiling value' for the ICER (also referred to as the Greek symbol lambda – λ) has been introduced, which represents the maximum amount society is willing to pay for a unit increase in effects, such as an additional quality-adjusted life year (QALY). Any services found in an evaluation to produce a QALY for a

The ICER decision rule is thus:

ICER decision rule $\Rightarrow$ $(C_N - C_C) / (E_N - E_C) < \lambda$

There are, however, some statistical problems with the use of ratios that have resulted in a recent movement away from ICERs. Instead, attention has turned to the incremental net benefit (INB) approach, which involves reworking the ICER decision rule, noted above, in the following way:

ICER decision rule $\Rightarrow$ $(C_N - C_C) / (E_N - E_C) < \lambda$

INB decision rule $\Rightarrow$ $(C_N - C_C) < \lambda (E_N - E_C)$

$\Rightarrow$ $0 < \lambda (E_N - E_C) - (C_N - C_C)$

In the INB decision rule, the ceiling ratio has been internalised. In other words, it is now internal to the equation, rather than an external figure against which the ICER must be compared. Thus, to calculate the INB of one intervention in comparison to another, a value for the ceiling ratio is required. Since this value does not actually exist, and selection of a value is often arbitrary, it is common in economic evaluations to present INBs for a range of different ceiling ratios, which are then presented to decision makers who can decide on the basis of their own preferences for a ceiling ratio.

More detailed information on ICER and INB decision rules and related issues can be found in the following publications: Johannesson and Meltzer (1998), Stinnett and Mullahy (1998), Briggs and Gray (2000) and Briggs (2001).

cost below this ceiling ratio should be funded, within the constraints of existing budgets.

Presentation of the results

The results of an economic evaluation are an aid to decision making, which will be considered alongside other evidence, preferences and pressures. The aim of any decision-making tool should be to clearly and objectively present results in a way that is understandable by the reader, including those readers who may be unfamiliar with the methods used. Clarity will enable readers to accurately assess the implications of the results and their applicability to alternative settings, gain an adequate understanding of the limitations and uncertainty associated with the results, and enable comparisons with other similar studies. Aspects of the evaluation that should be clearly described in the presentation of the results of an economic evaluation are listed in Box 25.

Generalisability

Generalisability is concerned with how representative the study sample is of the real-world situation (Mason, 1997). While it is important for an evaluation to demonstrate how effective an intervention is in a particular setting or for a particular user group, it is also important to assess how effective the same intervention would be in other settings – a

Box 25 Aspects of economic evaluations that should be reported

- Background information to explain the reason for the evaluation.
- The evaluation question posed.
- Description of the interventions being evaluated to enable replication.
- Design and setting of the evaluation.
- Description of the sample and methods of sample selection.
- Perspective of the evaluation.
- Outcomes measured and methods of measurement and valuation.
- Costs included and methods of measurement and valuation.
- Year and currency of cost data.
- Description and explanation of statistical methods employed.
- Description of the resources used by participants in each group.
- Description of the total costs incurred, disaggregated by different interest groups.
- Description of the benefits produced by each intervention.
- The incremental costs and benefits generated.
- The conclusions reached on the basis of the data generated.
- The policy relevance of the study.
- The generalisability of the results to other settings.
- Limitations of the study.
- Distributional implications.

different area, population or time period, for example. The representativeness of the results of an evaluation is of particular importance to decision makers who must interpret the results of studies conducted in settings different to their own.

There is much that researchers can do when designing and reporting the results of an economic evaluation to aid the generalisability of studies, including the following.

- *Pragmatic study design*: pragmatism in the design of an evaluation means attempting to evaluate services that are provided in the way that they would be delivered in everyday practice. Pragmatism thus helps to ensure that results are likely to be applicable to similar 'real-world' services. A pragmatic design means limiting the influence and interference of the research on the delivery of services, and evaluating a service that is already up and running, rather than a brand new service that has not yet had time to develop or establish itself in the real world.
- *Clear description of the interventions*: providing an adequate description of the intervention and comparison services can help decision makers in other settings to assess how similar their own services are to the ones being evaluated. The more similar they are, the more likely the results will be applicable to the new setting.
- *Clear description of the sample*: researchers should provide detailed descriptions of the study sample, including basic socio-demographic information (age, gender, ethnicity, social class, employment status, etc.) and the criteria used to include or exclude participants (for example, a study may have limited its population to adolescents).
- *Presentation of service-use data, unit costs and total costs*: reporting the physical quantities of services used by study participants and the unit costs of each service, as well as the total costs, will enable decision makers to make useful comparisons and perhaps apply their own locally applicable unit costs.

Uncertainty and limitations

The accuracy, usefulness and generalisability of results can be further enhanced by an exploration of the uncertainty surrounding the results and the limitations of the study. When dealing with cost data, there will almost inevitably be an element of uncertainty involved since service-use data will often be subject to recording errors or recall bias, and unit cost data may have been calculated using a general top-down approach, rather than a detailed bottom-up method. In addition, certain assumptions may have been made, such as assumptions regarding the average length of an appointment with a service professional, where more detailed data was unavailable.

One method commonly employed to test the robustness of the conclusions of an economic evaluation to the uncertainty and underlying assumptions made is sensitivity analysis. Sensitivity analysis involves systematic assessment of the impact of changes in the assumptions made. In other words, certain

variables in the analysis, such as individual cost items, are varied and cost-effectiveness reassessed on the basis of new values. For example, a study that applied local unit costs to service-use data may benefit from testing the impact of changing these costs to nationally applicable ones, in order to test the generalisability of the results to a wider setting. Alternatively, a unit cost based on an assumption regarding the length of time participants spent with a particular service professional could be varied to explore the impact of increasing or decreasing the chosen time period. When variation of a particular variable in sensitivity analysis has no impact on the final results of an evaluation, the researcher can have greater confidence in the accuracy of the results than in the absence of such evidence. More detailed information on the approaches to sensitivity analysis can be found in Briggs *et al*. (1994), HM Treasury (1997) or Johnston *et al*. (1999).

Statistical analysis can also be used to help quantify and explore uncertainty. For example, where an evaluation employs statistical analysis to test for differences in cost or outcome between two groups, a *p*-value is generated that indicates the presence or absence of a statistically significant result. Such estimates, however, say nothing about the uncertainty involved. To do this, evaluators should additionally report the confidence intervals around the point estimates (the range of values within which the true difference between the two interventions is likely to lie with, say, 95 per cent confidence) and some measure of variability, such as the standard deviation or variance (measures of individual deviations from the mean). Readers should refer to standard statistical texts for more detailed information on such statistical analysis, or see Johnston *et al*. (1999) for a simple introduction.

Presentation of the limitations of the study is essential to enable readers to assess the likely accuracy and applicability of the results produced. Limitations will vary from study to study but may include a small, biased or unrepresentative sample, a narrow cost perspective, missing data or a lack of a measure of final outcomes. Each limitation should be alluded to and discussed in terms of its likely impact on the results reported. For example, an observational study comparing relatively intensive levels of home-care support for older people to a less intensive service may discover that the participants receiving intensive home care are older and more frail than those receiving the less intensive service. Such a situation reduces the validity of the results and is likely to result in an underestimate of the true difference in outcome between the two groups, given the advantage held by users of the less intensive service.

Comparisons with other study results

There are a number of benefits to comparing study results with evaluations of similar services. First, comparisons enable assessment of the likely accuracy of the current and existing results by asking whether similar studies produce similar results and, if not, why not? Second, comparisons help to assess the generalisability of the results by asking whether studies set in different locations or populations produce similar results and, if not, why not? In addition, comparisons with other studies can help to place the intervention in question on a broader comparative footing.

To illustrate, if an energy efficiency education campaign were compared to a neighbouring area where no such campaign existed, a great deal of useful information would be produced, but the study would be unable to say anything about the relative cost-effectiveness of the campaign as compared to, perhaps, the direct provision of winter payments to the elderly. In order to try and build up a more informative picture to guide resource allocation, the results should also be compared with existing evaluations that employ alternative comparators. If such studies do not exist, the search for comparison evaluations will help to highlight areas where future research would be of benefit.

Economic evaluations that use standard generic outcome scales, such as quality-adjusted life years or monetary valuation of outcomes, can additionally be compared to a broader range of evaluations in social welfare and beyond. Such comparisons can help to locate the intervention in question along a spectrum of cost-effectiveness, which may better support its implementation.

In many areas of social welfare, there will currently be little with which to compare the results of an economic evaluation but this will not be true of all areas. A number of economic evaluations have been carried out in this field (McDaid *et al*., 2003) and this number will increase over time.

Distribution and equity considerations

Establishing the cost-effectiveness of an intervention in relation to alternative interventions provides decision makers with information on how to maximise user outcomes within a given budget. As Sassi and colleagues (2001) note, however, this disregards issues of who benefits from any gains in outcome and who loses. In other words, the distributional, or equity, implications of resource allocation decisions are often ignored. Focusing exclusively on efficiency, to the detriment of equity considerations, could, for example, result in the allocation of funds towards services that benefit the rich rather than the poor, the old rather than the young, the moderately disabled rather than the severely disabled, white populations rather than those from other ethnic backgrounds, developed countries rather than underdeveloped countries, etc. Although such resource allocations may indeed maximise the benefits gained to society as a whole, many would argue that such distributions are 'unfair'.

A number of steps can be taken within an economic evaluation to ensure that equity considerations are given a higher priority and these are summarised in Box 26. A more detailed discussion of distributional issues in economic evaluation is provided by Sassi *et al*. (2001).

Box 26 Methods for incorporating equity issues into economic evaluation

Equity weights

Equity weights are the most commonly applied and advocated method of dealing with issues of equity. They are numbers that express the relative importance of different groups in society, which are used to adjust improvements in outcome. Equity weights represent the extent to which society is willing to sacrifice improvements in outcome in order to ensure a more equitable distribution of resources and, thus, the greater the weight, the more society is willing to sacrifice. They can be derived from a representative sample of the population.

Person trade-off method

The person trade-off approach involves eliciting 'social' rather than 'individual' health-outcome valuations from respondents. It involves asking respondents to decide how many people with a particular characteristic would be equivalent to a given number of individuals who differ on that characteristic. For example, respondents may be asked how many lives of people aged 60 would need to be saved to be equivalent to ten lives saved of people aged 15.

Descriptive approach

Because of a number of limitations of the above methods, however, Sassi and colleagues (2001) advocate a descriptive approach, involving the presentation of information on the effects of interventions on different sub-populations. In other words, researchers are advised to present the results of their evaluations in terms not just of overall cost-effectiveness but also of the cost-effectiveness for sub-groups within the main group separated by certain characteristics such as gender, age range, ethnic background, socio-economic group, etc. This would provide the decision maker with, for example, evidence on whether males or females benefit the most.

Implementation of the results

Once the relative cost-effectiveness of an intervention has been determined, it does not necessarily follow that the intervention will be implemented, for a number of reasons. First, decision makers may take other issues into account, alongside the economic evidence, including population and media preferences, equity considerations or current government priorities. They may choose to fund a less efficient service if, for example, the benefits generated are received by the poorest members of society, or if public, political or media pressure is strong. Economic evaluation should not be seen as *the* decision-making tool, but as an additional tool to help support decisions alongside other forms of evidence and other preferences.

Second is the possibility of uncontrolled expenditure growth if evaluations find that many new services are more cost-effective than existing ones; for example, if many new services produce QALYs at less than the value of the ceiling ratio selected by the government. In such a situation, the resources available to funders might be depleted, making it impossible to fund further interventions, even those that are found to be an efficient use of resources. Rather than fund all efficient services, an alternative approach is to rank interventions in a 'league table' in order perhaps of their incremental cost per QALY gained. Resources should then be directed towards those interventions with the lowest cost per QALY gained, starting at the top of the list and moving down it until funding runs out. The cost of the QALY gained from the last intervention to be funded is the implicit value of the ceiling ratio in a society.

There are a number of obstacles to the successful use of such league tables, however. The main criticisms relate to the lack of comparability between the methods used in different evaluations, the use of inappropriate comparison groups that can produce a more favourable ranking than would otherwise have been found and the limited generalisability of results to all service users. For more discussion on resource allocation and league tables, see for example Williams (1985), Birch and Gafni (1992), Gerard and Mooney (1993) and Weinstein (1995).

7 Summary and conclusions

As long as there is a mismatch between the demand for social welfare services and the resources available to provide such care, the need for evaluation of both effectiveness and cost-effectiveness will remain. Given the philosophy of welfare economics – the desire to improve on total societal well-being by maximising the benefits yielded from limited resources – the inclusion of economics should not be seen as a threat or a replacement for traditional social welfare research methods, but as an additional tool to aid resource-allocation decisions in a complex organisation. By adding to the strengths of social welfare research, economic analysis can provide valuable information to increase the benefits to users and carers, and enhance the overall efficiency with which resources are allocated in the social welfare field.

Economic evaluation involves a number of evaluation designs and techniques that may be unfamiliar to many social welfare researchers. This report provides a relatively simplistic guide to economic evaluation that aims to give researchers a deeper understanding of the theory and purpose behind it, and enough information to be able to critically appraise economic evidence and design an economic evaluation capable of producing meaningful results. To support these aims, Box 27 contains a checklist designed to support critical appraisal of economic evaluations, which acts as a useful guide to the processes involved in undertaking an economic evaluation.

Many elements of economic evaluation are more complex than this guide would suggest. Thus, to carry out a full economic evaluation, readers are strongly advised to seek the support of economists with practical experience of applying economic techniques to evaluations in social welfare or similar fields and to refer to the more detailed references provided throughout this guide.

There are a number of areas in economic evaluation that are not clear-cut, such as the methods used to value informal care or productivity losses, or the means by which costs and outcomes should be combined. In addition, a number of obstacles to a successful evaluation exist, including the lack of appropriate outcome measures. Economics is not an exact science, but methods continue to be improved over time and, as research into these obstacles grows, current limitations will become less significant. In particular, greater multi-disciplinary collaboration between economists and social welfare researchers and other professionals will help to enhance the development of methods appropriate to the field.

Box 27 A checklist for assessing economic evaluation

1 Was a well-defined question posed in answerable form?

1.1 Did the study examine both costs and effects of the service(s) or programme(s)?

1.2 Did the study involve a comparison of alternatives?

1.3 Was a viewpoint (perspective) for the analysis stated and was the study placed in any particular decision-making context?

2 Was a comprehensive description of the competing alternatives given (i.e. can you tell who did what to whom, where and how often)?

2.1 Were any important alternatives omitted?

2.2 Was (Should) a *do nothing* alternative (be) considered?

3 Was the effectiveness of the programmes or services established?

3.1 Was this done through a randomised, controlled clinical trial? If so, did the study protocol reflect what would happen in regular practice?

3.2 Was effectiveness established through an overview of published studies?

3.3 Were observational data or assumptions used to establish effectiveness? If so, what are the potential biases in results?

4 Were all the important and relevant costs and outcomes for each alternative identified?

4.1 Was the range wide enough for the research question at hand?

4.2 Did it cover all relevant viewpoints? (Possible viewpoints include the community or social viewpoint, and those of users and carers.)

4.3 Were capital costs, as well as operating costs, included?

5 Were the costs and outcomes measured accurately in appropriate physical units (e.g. hours of social work time, number of home-help visits, lost workdays)?

5.1 Were any of the identified items omitted from measurement? If so, does this mean that they carried no weight in the subsequent analysis?

5.2 Were there any special circumstances (e.g. joint use of resources such as shared facilities) that made measurement difficult? Were these circumstances handled appropriately?

Continued

Box 27 Continued

6 Were costs and consequences valued credibly?

6.1 Were the sources of all values clearly identified? (Possible sources include market values, users' or clients' preferences and views, policymakers' views and social welfare professionals' judgements.)

6.2 Were market values employed for changes involving resources gained or depleted?

6.3 Where market values were absent (e.g. volunteer labour), or market values did not reflect actual values (such as a therapy room donated at a reduced rate), were adjustments made to approximate market values?

6.4 Was the valuation of outcomes appropriate for the question posed (i.e. has the appropriate type (or types) of analysis – cost-effectiveness, cost-benefit, cost-utility – been selected)?

7 Were costs and outcomes adjusted for differential timing?

7.1 Were costs and outcomes that occur in the future 'discounted' to their present value?

7.2 Was any justification given for the discount rate used?

8 Was an incremental analysis of costs and outcomes of alternatives performed?

8.1 Were the additional (incremental) costs generated by one alternative over another compared to the additional effects, benefits or utilities generated?

9 Was allowance made for uncertainty in the estimates of costs and consequences?

9.1 Were appropriate statistical analyses performed?

9.2 If a sensitivity analysis was employed, was justification provided for the ranges of values (for key study parameters)?

9.3 Were study results sensitive to changes in the values (within the assumed range for sensitivity analysis, or within the confidence interval around the ratio of costs to outcomes)?

10 Did the presentation and discussion of study results include all issues of concern to users?

10.1 Were the conclusions of the analysis based on some overall index or ratio of costs to outcomes (e.g. cost-effectiveness ratio)? If so, was the index interpreted intelligently or in a mechanistic fashion?

Continued overleaf

Box 27 Continued

10.2 Were the results compared with those of others who have investigated the same question? If so, were allowances made for potential differences in study methodology?

10.3 Did the study discuss the generalisability of the results to other settings and client groups?

10.4 Did the study allude to, or take account of, other important factors in the choice or decision under consideration (e.g. distribution of costs and outcomes, or relevant ethical issues)?

10.5 Did the study discuss issues of implementation, such as the feasibility of adopting the 'preferred' programme given existing financial or other constraints, and whether any freed resources could be redeployed to other worthwhile programmes?

Adapted from Drummond *et al*. (1997).

References

Arnold, D., O'Leary, S., Wolf, L. and Acker, M. (1993) 'The parenting scale: a measure of dysfunctional parenting in discipline situations', *Psychological Assessment*, Vol. 5, pp. 137–44

Arrow, K.J. (1963) 'Uncertainty and welfare economics of medical care', *American Economic Review*, Vol. 53, pp. 941–73

Beecham, J. (2000) *Unit Costs – Not Exactly Child's Play. A Guide to Estimating Unit Costs for Children's Social Care*. London: Department of Health

Beecham, J. and Knapp, M. (2001) 'Costing psychiatric interventions', in G. Thornicroft, C. Brewins and J. Wing (eds) *Measuring Mental Health Needs*, 2nd edn. London: Gaskell

Bennett, K.J., Torrance, G.W., Boyle, M.H., Guscott, R. and Moran, L.A. (2000) 'Development and testing of a utility measure for major, unipolar depression (McSad)', *Quality of Life Research*, Vol. 9, pp. 109–20

Berg, B. and Rounds, D. (1992) 'Sociometric applications in criminology and other settings: a re-examination of a traditional method', *Sociological Practice Review*, Vol. 3, pp. 26–132

Birch, S. and Gafni, A. (1992) 'Cost effectiveness/utility analyses: do current decision rules lead us to where we want to be?', *Journal of Health Economics*, Vol. 11, pp. 279–96

Birleson, P. (1980) 'The validity of depressive disorder in childhood and the development of a self-rating scale: a research report', *Journal of Child Psychology and Psychiatry*, Vol. 22, pp. 73–88

Black, N. (1996) 'Why we need observational studies to evaluate the effectiveness of health care', *British Medical Journal*, Vol. 312, pp. 1215–18

Bland, J.M. and Altman, D.G. (2002) 'Statistics notes: validating scales and indexes', *British Medical Journal*, Vol. 324, pp. 606–7

Bloom, H.S., Orr, L.L., Bell, S.H., Cave, G., Doolittle, F., Lin, W. and Bos, J.M. (1997) 'The benefits and costs of JTPA title II-A programs: key findings from the National Job Training Partnership Act study', *Journal of Human Resources*, Vol. 32, No. 3, pp. 549–76

Bowling, A. (2002) *Research Methods in Health*. Buckingham: Open University Press

Briggs, A. (2001) 'Handling uncertainty in economic evaluation and presenting the results', in M. Drummond and A. McGuire (eds) *Economic Evaluation in Health Care: Merging Theory with Practice*. New York: Oxford University Press

Briggs, A. and Gray, A. (2000) 'Using cost effectiveness information', *British Medical Journal*, Vol. 320, p. 246

Briggs, A., Sculpher, M. and Buxton, M. (1994) 'Uncertainty in the economic evaluation of health care technologies: the role of sensitivity analysis', *Health Economics*, Vol. 3, pp. 95–104

Brouwer, W., Rutten, F. and Koopmanschap, M. (2001) 'Costing in economic evaluations', in M. Drummond and A. McGuire (eds) *Economic Evaluation in Health Care: Merging Theory with Practice*. New York: Oxford University Press

Brugha, T., Bebington, P., Tennant, C. and Hurry, J. (1985) 'The list of threatening experiences: a subset of 12 life event categories with considerable long-term contextual threat', *Psychological Medicine*, Vol. 15, pp. 189–94

Byford, S. and Palmer, S. (1998) 'Common errors and controversies in pharmacoeconomic analyses', *PharmacoEconomics*, Vol. 13, No. 6, pp. 659–66

Byford, S. and Sefton, T. (2002) *First Aid: Lessons from Health Economics for Economic Evaluation in Social Welfare*. London: LSE Health and Social Care Discussion Paper

Byford, S., Harrington, R., Torgerson, D., Kerfoot, M., Dyer, E., Harrington, V., Woodham, A., Gill, J. and McNiven, F. (1999) 'Cost-effectiveness analysis of a home-based social work intervention for children and adolescents who have deliberately poisoned themselves: the results of a randomised controlled trial', *British Journal of Psychiatry*, Vol. 174, pp. 56–62

Challis, D. and Davies, B (1986) *Case Management in Community Care*. Aldershot: Gower

CIPFA (2001) *Health Service Financial Database*. London: Chartered Institute of Public Finance and Accountancy

CIPFA (2002) *Personal Social Services Statistics*. London: Chartered Institute of Public Finance and Accountancy

Clark, R.E., Bush, P.W., Becker, D.R. and Drake, R.E. (1996) 'A cost-effective comparison of supported employment and rehabilitative day treatment', *Administration and Policy in Mental Health*, Vol. 24, No. 1, pp. 63–77

Clarke, P.M. (1998) 'Cost benefit analysis and mammographic screening: a travel cost approach', *Journal of Health Economics*, Vol. 17, pp. 767–87

Clinch, P. and Healy, J. (2000) 'Cost-benefit analysis of domestic energy efficiency', *Energy Policy,* Vol. 29, pp. 113–24

Coast, J. (1999) 'The appropriate uses of qualitative methods in health economics', *Health Economics*, Vol. 8, pp. 345–53

Coates, R.B. and Gehm, J. (1989) 'An empirical assessment', in M. Wright and B. Galaway (eds) *Meditation and Criminal Justice*. London: Sage

Cox, A. and Bentovim, A. (2000) *Framework for the Assessment of Children in Need and their Families: The Family Pack of Questionnaires and Scales*. London: Department of Health

Davie, C.E., Hutt, S.J., Vincent, E. and Mason, M. (1984) *The Young Child at Home*. Windsor: NFER-Nelson

Department of Health (1996) *Policy Appraisal and Health*. London: Department of Health

Dolan, P. (2001) 'Output measures and valuation in health', in M. Drummond and A. McGuire (eds) *Economic Evaluation in Health Care: Merging Theory with Practice*. New York: Oxford University Press

Drummond, M.F. (1991) 'Output measurement for resource-allocation decisions in health care', in A. McGuire, P. Fenn and K. Mayhew (eds) *Providing Health Care. The Economics of Alternative Systems of Finance and Delivery.* Oxford: Oxford University Press

Drummond M.F. and Jefferson, T.O. (1996) 'Guidelines for authors and peer reviewers of economic submissions to the *BMJ*', *British Medical Journal*, Vol. 313, pp. 275–83

Drummond, M.F., O'Brien, B., Stoddart, G.L. and Torrance, G.W. (1997) *Methods for the Economic Evaluation of Health Care Programmes*, 2nd edn. Oxford: Oxford Medical Publications

Drummond, M.F., Mohide, E.A., Tew, M., Streiner, D.L., Pringle, D.M. and Gilbert, J.R. (1991) 'Economic evaluation of a support program for caregivers of demented elderly', *International Journal of Technology Assessment in Health Care*, Vol. 7, No. 2, pp. 209–19

Edwards, R.T. and Thalanany, M. (2001) 'Trade offs in the conduct of economic evaluations of child mental health services', *Mental Health Services Research*, Vol. 3, No. 2, pp. 99–105

Eyberg, S.M. and Ross, A.W. (1978) 'Assessment of child behavior problems: the validation of a new inventory', *Journal of Clinical Child Psychology*, Vol. 7, pp. 113–16

Gerard, K. and Mooney, G. (1993) 'QALY league tables: handle with care', *Health Economics*, Vol. 2, pp. 59–64

Glick, H.A., Polsky, D.P. and Schulman, K.A. (2001) 'Trial-based economic evaluation: an overview of design and analysis', in M. Drummond and A. McGuire (eds) *Economic Evaluation in Health Care: Merging Theory with Practice*. New York: Oxford University Press

Goeree, R., O'Brien, B., Blackhouse, G., Agro, K. and Goering, P. (1999) 'The valuation of productivity costs due to premature mortality: a comparison of the human-capital and friction-cost methods for schizophrenia', *Canadian Journal of Psychiatry*, Vol. 44, pp. 455–63

Gold, M.R., Siegel, J.E. and Russell, L.B. *et al*. (1996) *Cost-effectiveness in Health and Medicine*. Oxford: Oxford University Press

Goldberg, D.P. (1972) *The Detection of Psychiatric Illness by Questionnaire*. Oxford: Oxford University Press

Goodman, R. (1997) 'The Strengths and Difficulties Questionnaire: a research note', *Journal of Child Psychology and Psychiatry*, Vol. 38, pp. 581–6

Harrington, R., Kerfoot, M., Dyer, E., McNiven, F., Gill, J., Harrington, V., Woodham, A. and Byford, S. (1998) 'Randomized trial of a home-based family intervention for children who have deliberately poisoned themselves', *Journal of the American Academy of Child and Adolescent Psychiatry*, Vol. 37, No. 5, pp. 512–18

Henderson, S., Duncan-Jones, P., Byrne, D.G. and Scott, R. (1980) 'Measuring social relationships: the interview schedule for social interaction', *Psychological Medicine*, Vol. 10, pp. 723–34

HM Treasury (1997) *Appraisal and Evaluation in Central Government: The Green Book*. London: The Stationery Office

Holtermann, S. (1998) *Weighing it up: Applying Economic Evaluations to Social Welfare Programmes*. York: York Publishing Services for the Joseph Rowntree Foundation

Hu, T.W. and Jerrell, J.M. (1998) 'Estimating the cost of three case management programmes for treating people with severe mental illness', *British Journal of Psychiatry*, Vol. 173, Suppl. 36, pp. 26–32

Ichinose, C.K., Kingdon, D.W. and Hernandez, M. (1994) 'Developing community alternatives to group home placement for SED special education students in the Ventura County system of care', *Journal of Child and Family Studies*, Vol. 3, No. 2, pp. 193–210

Jan, S. (1998) 'A holistic approach to the economic evaluation of health programs using institutionalist methodology', *Social Science and Medicine*, Vol. 47, No. 10, pp. 1565–72

Johannesson, M. (1995) 'A note on the depreciation of the societal perspective in economic evaluation of health care', *Health Policy*, Vol. 33, pp. 59–66

Johannesson, M. and Jonsson, B. (1991) 'Economic evaluation in health care. Is there a role for cost-benefit analysis?', *Health Policy*, Vol. 17, No. 1, pp. 1–23

Johannesson, M, and Meltzer, D. (1998) 'Some reflections on cost-effectiveness analysis', *Health Economics*, Vol. 7, pp. 1–7

Johnson-Masotti, A. and S. Pinkerton *et al*. (2000) 'Cost-effectiveness of an HIV risk reduction intervention for adults with severe mental illness', *Aids Care*, Vol. 12, No. 3, pp. 321–32

Johnston, K., Buxton, M.J., Jones, D.R. and Fitzpatrick, R. (1999) 'Assessing the costs of healthcare technologies in clinical trials', *Health Technology Assessment*, Vol. 3, No. 6, pp. 1–76

Katz, S., Ford, A.B., Moskowitz, R.W., Jackson, B.A. and Jaffe, M.W. (1963) 'Studies of illness in the aged', *Journal of the American Medical Association*, Vol. 185, pp. 914–19

Kielhorn, A. and von der Schulenburg, J.M.G. (2000) *The Health Economics Handbook*. Chester: Adis International

Kind, P. (1988) *The Design and Construction of Quality of Life Measures*. York: Centre for Health Economics Discussion Paper 43

Knapp, M. (1984) *The Economics of Social Care*. Basingstoke: Macmillan

Knapp, M. (1993) 'The costing process: background theory', in A. Netten and J. Beecham (eds) *Costing Community Care*. Aldershot: Ashgate Publishing Ltd

Knapp, M. (1997) 'Economic evaluations and interventions for children and adolescents with mental health problems', *Journal of Child Psychology and Psychiatry*, Vol. 38, pp. 3–25

Knapp, M. and Beecham, J. (1993) 'Reduced list costings: examination of an informed short cut in mental health research', *Health Economics*, Vol. 2, pp. 313–22

Knudsen, H.C., Vazquez-Barquero, J.L., Welcher, B., Gaite, L., Becker, T., Chisholm, D.A., Ruggeri, M., Schene, A.H. and Thornicroft, G. (2000) 'Translation and cross-cultural adaptation of outcome measurements for schizophrenia: EPSILON Study 2', *British Journal of Psychiatry*, Vol. 177, pp. 8–14

Koopmanschap, M. and Rutten, F. (1996) 'A practical guide for calculating indirect costs of disease', *PharmacoEconomics*, Vol. 10, pp. 460–6

Krahn, M. and Gafni, A. (1993) 'Discounting in the economic evaluation of health care interventions', *Medical Care*, Vol. 31, No. 5, pp. 403–18

Lawton, M.P. (1975) 'The Philadelphia Geriatric Center Morale Scale: a revision', *Journal of Gerontology*, Vol. 30, pp. 85–9

Loomes, G. and McKenzie, L. (1989) 'The use of QALYs in health care decision making', *Social Science and Medicine*, Vol. 28, pp. 299–308

Luce, B.R. and Elixhauser, A. (1990) 'Estimating costs in the economic evaluation of medical technologies', *International Journal of Technology Assessment in Health Care*, Vol. 6, pp. 57–75

Lupton, R. (2001) *Places Apart? The Initial Report of CASE's Areas Study*. London: Centre for the Analysis of Social Exclusion, London School of Economics

McDaid, D. (2001) 'Estimating the costs of informal care for people with Alzheimer's disease: methodological and practical challenges', *International Journal of Geriatric Psychiatry*, Vol. 16, No. 4, pp. 400–5

McDaid, D., Sefton, T. and Byford, S. (2003) 'Identifying and assessing the quality of economic evaluations in social welfare' (in progress)

McGuire, A., Henderson, J. and Mooney, G. (1992) *The Economics of Health Care*. London: Routledge

McIntosh, E., Donaldson, C. and Ryan, M. (1999) 'Recent advances in the methods of cost-benefit analysis in healthcare. Matching the art to the science', *PharmacoEconomics*, Vol. 15, No. 4, pp. 357–67

Mason, J. (1997) 'The generalisability of pharmacoeconomic studies', *PharmacoEconomics*, Vol. 11, pp. 503–14

Mauskopf, J., Schulman, K., Bell, L. and Glick, H. (1996) 'A strategy for collecting pharmacoeconomic data during phase II/III clinical trials', *PharmacoEconomics*, Vol. 9, No. 3, pp. 264–77

Mauskopf, J.A., Paul, J.E., Grant, D.M. and Stergachis, A. (1998) 'The role of cost-consequences analysis in healthcare decision-making', *PharmacoEconomics*, Vol. 13, No. 3, pp. 277–88

Miers, D., Maguire, M., Goldie, S., Sharpe, K., Hale, C., Netten, A., Uglow, S., Doolin, K., Hallam, A., Enterkin, J. and Newburn, T. (2001) *An Exploratory Evaluation of Restorative Justice Schemes*. Crime Reduction Research Series Paper 9. London: Home Office

Miller, I.V., Epstein, N.B. and Bishop, D.S. *et al*. (1985) 'The McMaster Family Assessment Device: reliability and validity', *Journal of Marital and Family Therapy*, Vol. 11, pp. 345–56

Minty, B. (1999) 'Annotation: outcomes in long-term foster family care', *Journal of Child Psychology and Psychiatry*, Vol. 40, No. 7, pp. 991–99

Moser, C.A. and Kalton, G (1996) *Survey Methods in Social Investigation*. Aldershot: Dartmouth Publishing Company

National Institute for Clinical Excellence (2001) *Technical Guidance for Manufacturers and Sponsors on Making a Submission to a Technology Appraisal*. London: National Institute for Clinical Excellence

Netten, A. and Curtis, L. (2002) *Unit Costs of Health and Social Care 2002*. University of Kent at Canterbury: Personal Social Services Research Unit

Netten, A. and Dennett, J. (1996) *Unit Costs of Health and Social Care 1996*. University of Kent at Canterbury: Personal Social Services Research Unit

Oliver, A., Healey, A. and Donaldson, C. (2002) 'Choosing the method to match the perspective: economic assessment and its implications for health-services efficiency', *Lancet*, Vol. 359, pp. 1771–4

Perkins, F. (1994) *Practical Cost Benefit Analysis*. Melbourne: Macmillan Education Press

Power, A. and Bergin, E. (1999) *Neighbourhood Management*. CASE Paper 31. London: Centre for the Analysis of Social Exclusion, London School of Economics

Presser, L. and Van Voorhis, P. (2002) 'Values and evaluation: assessing processes and outcomes of restorative justice programs', *Crime and Delinquency*, Vol. 48, No. 1, pp. 162–88

Pritchard, C. and Sculpher, M. (2000) *Productivity Costs: Principles and Practice in Economic Evaluation*. London: Office of Health Economics

Pritchard, C. and Williams, R. (2001) 'A three year comparative longitudinal study of a school-based social work family service to reduce truancy, delinquency and school exclusions', *Journal of Social Welfare and Family Law*, Vol. 23, No. 1, pp. 23–43

Rajkumar, A.S. and French, M.T. (1997) 'Drug abuse, crime costs, and the economic benefits of treatment', *Journal of Quantitative Criminology*, Vol. 13, No. 3, pp. 291–323

Ratcliffe, J. (2000) 'The use of conjoint analysis to elicit willingness-to-pay values. Proceed with caution?', *International Journal of Technology Assessment in Health Care*, Vol. 16, No. 1, pp. 270–5

Rest, J.R. (1979) *Moral Development Advances in Research and Theory*. Westport, CT: Praeger

Revolving Doors Agency (2000) *Mental Health, Multiple Needs, and the Police*. London: Revolving Doors Agency

Richardson, L. (1999) *Tackling Difficult Estates*. CASE/Social Exclusion Unit Seminar held at the National Tenants' Resource Centre, Trafford Hall. London: Centre for the Analysis of Social Exclusion, London School of Economics

Robinson, R. (1993) 'Economic evaluation and health care', *British Medical Journal*, Vol. 307, pp. 728–9, 793–5, 859–62, 924–6, 994–6

Ryan, M. (1999) 'Using conjoint analysis to take account of patient preferences and go beyond health outcomes: an application to in vitro fertilisation', *Social Science and Medicine*, Vol. 48, No. 4, pp. 535–46

Ryan, M. and Farrar, S. (2000) 'Using conjoint analysis to elicit preferences for health care', *British Medical Journal*, Vol. 320, pp. 1530–3

Ryan, M., Scott, D.A., Reeves, C., Bate, A., van Teijlingen, E.R., Russell, E.M., Napper, M. and Rob, C.M. (2001) 'Eliciting public preferences for healthcare: a systematic review of techniques', *Health Technology Assessment*, Vol. 5, No. 5, pp. 1–186

Sampson, R.J., Raudenbush, S.W. and Earls, F. (1997) 'Neighborhoods and violent crime: a multilevel study of collective efficacy', *Science*, Vol. 277, pp. 916–24

Sassi, F., Archard, L. and Le Grand, J. (2001) 'Equity and the economic evaluation of healthcare', *Health Technology Assessment*, Vol. 5, No. 3, pp. 1–138

Sassi, F., Archard, L. and McDaid, D. (2002) 'Searching literature databases for health care economic evaluations: how systematic can we afford to be?', *Medical Care*, Vol. 40, No. 5, pp. 387–94

Schoenwald, S.K., Ward, D.M., Henggeler, S.W., Pickrel, S.G. and Patel, H. (1996) 'Multisystemic therapy treatment of substance abusing or dependent adolescent offenders: costs of reducing incarceration, inpatient, and residential placement', *Journal of Child and Family Studies*, Vol. 5, No. 4, pp. 431–44

Sefton, T., Byford, S., McDaid, D., Hills, J. and Knapp, M. (2002) *Making the Most of it: Economic Evaluation in the Social Welfare Field*. York: York Publishing Services for the Joseph Rowntree Foundation

Sheldon, T.S. (1992) 'Discounting in health care decision making: time for a change?', *Journal of Public Health Medicine*, Vol. 14, No. 3, pp. 250–6

Simourd, D.J. (1997) 'The criminal justice sentiments scale-modified and pride in delinquency scale: psychometric properties and construct validity of two measures of criminal attitudes', *Criminal Justice and Behavior*, Vol. 24, pp. 52–70

Snaith, R.P., Constantopoulos, A.A., Jardine, M.Y. and McGuffin, P. (1978) 'A clinical scale for the assessment of irritability', *British Journal of Psychiatry*, Vol. 132, pp. 164–71

Stinnett, A.A. and Mullahy, J. (1998) 'A new framework for the analysis of uncertainty in cost-effectiveness analysis', *Medical Decision Making*, Vol. 18, suppl., pp. S68–S80

Sure Start (2002) *Making a Difference for Children and Families*. London: Sure Start

Thompson, R., Ruma, P.R., Schuchmann, L.F. and Burke, R.V. (1996) 'A cost-effectiveness evaluation of parent training', *Journal of Child and Family Studies*, Vol. 5, No. 4, pp. 415–29

Torgerson, D.J. and Raftery, J. (1999) 'Discounting', *British Medical Journal*, Vol. 319, pp. 914–15

Trulson, C., Triplett, R. and Snell, C. (2001) 'Social control in a school setting: evaluating a school-based boot camp', *Crime and Delinquency*, Vol. 47, pp. 573–609

UK700 Group (Byford, S., Fiander, M., Torgerson, D.J., Barber, J.A., Thompson, S.G., Burns, T., van Horne, E., Gilvarry, C. and Creed, F.) (2000) 'Cost-effectiveness of intensive v. standard case management for severe psychotic illness: UK700 case management trial', *British Journal of Psychiatry,* Vol. 176, pp. 537–43

Umbreit, M.S. (1994) *Victim Meets Offender: The Impact of Restorative Justice and Mediation*. New York: Criminal Justice Press

Umbreit, M.S. and Coates, R.B. (1993) 'Cross-site analysis of victim–offender mediation in four states', *Crime and Delinquency*, Vol. 39, pp. 565–85

Umbreit, M.S. and Fercello, C. (1997) *Woodbury Police Department's Restorative Justice Community Conferencing Program: An Initial Assessment of Client Satisfaction*. University of Minnesota, MN: Center for Restorative Justice and Mediation

Wade, D.T. (1992) 'Evaluating outcome in stroke rehabilitation (quality control and clinical audit)', *Scandinavian Journal of Rehabilitation Medicine*, Vol. 26, pp. 97–104

Warner, B.D. and Pierce, G.L. (1993) 'Reexamining social disorganization theory using calls to the police as a measure of crime', *Criminology*, Vol. 31, pp. 493–517

Weinstein, M.C. (1995) 'From cost-effectiveness ratios to resource allocation: where to draw the line?', in F.A. Sloan (ed.) *Valuing Health Care*. Cambridge: Cambridge University Press

Williams, A. (1985) 'Economics of coronary artery bypass grafting', *British Medical Journal*, Vol. 291, pp. 326–9

Williams, A. and Kind, P. (1992) 'The present state of play about QALYs', in A. Hopkins (ed.) *Measures of Quality of Life*. London: Royal College of Physicians