

'Reaching the parts other methods cannot reach':

Methods of economic evaluation of interventions that cross the health and social care divide – the example of palliative day care

Hannah-Rose Douglas, Health Services Research Unit, London School of Hygiene and Tropical Medicine

Presented at the Health Economics Study Group meeting, University of Kent at Canterbury, July 9th -11th 2003

Work in progress – not to be cited without author's permission

Introduction

This paper discusses the problem of economic evaluation of health care interventions that are complex. Complex services are defined as those that have multiple goals in relation to improving welfare, and services that have subtle outcomes that are hard to define or measure. The paper considers how welfare economic theory can help to clarify some issues and obscure others in relation to the problem of evaluating these services. These issues will be explored with reference to the evaluation of palliative day care (PDC). All methods of evaluation make value assumptions and use proxies for outcomes since individuals cannot be observed making their choices for health care commodities in a free market. All methods therefore represent a trade-off between the data that it is possible to collect and validity of the findings in terms of adherence to economic theory and consequent ability to interpret results. Methods have developed in health economic evaluation based on the idea of measuring health gain from the use of health services. These approaches may not provide useful insights for services where outcomes cannot be easily observed or understood, and may differ from the subjective value that people place on using those services. The main question posed in this paper is whether approaches that are based on the theories of choice, trade and the satisfaction of preferences represent a way forward in evaluation of these services.

The paper begins with an analysis of the nature of complexity in relation to health and social care and why the evaluation of services such as PDC represent a particular challenge for economic evaluation. A very brief review of the theoretical underpinning of health economic evaluation is presented as a way of understanding the development of the current methods of empirical analysis. In light of this theory, the cost-utility approach is contrasted with the choice experiment approach that attempts to measure the value individuals place on particular services rather than measuring improvements in health, however constructed. It builds an argument for why the choice experiment approach has a good theoretical basis as well as practical advantages for the evaluation of complex services. These ideas will be explored in an empirical study comparing both approaches.

1. The problem of complexity in economic evaluation

'Complex' services are defined here as those where the welfare benefits derived from the service is difficult to define, or contested by different stakeholders and therefore hard to value and measure. It also applies to services where the relationship between service inputs and outcomes is not easy to describe, and to services that incorporate a range of different inputs that can be accessed by individuals. The separate components of the service may not themselves have been well evaluated, and overall, there may be a general lack of understanding (or disagreement) about how services should be provided to maximise their value to patients.

Palliative day care is presented as an example of a complex intervention. It is a specialist service for people who usually have advanced life-limiting illness but are able to live at home. These people are perceived to have particular needs that can be met by attending a specialist day centre (Fisher and McDaid 1996, Tebbit 1999, Hearn, Higginson et al 2001). Palliative day care therefore crosses the somewhat artificial boundary between health and social care (Robbins 1998). Patients who access palliative care services are identified as having a spectrum of needs or one specific need that can be alleviated or reduced by a combination of health, social, psychological and spiritual activities provided in the palliative day care centre (Higginson 1993). Some people only use the service infrequently whereas others may need more intensive support. There is some difference of opinion as to how services should be organised and what they are for (Douglas, Higginson et al 2000).

The authors of descriptive studies have argued that PDC is more than the sum of its parts. "The purpose is more important than the activity. Helping individuals retain a sense of integrity and dignity in the face of progressive losses is the primary goal of hospice and hospice day care" (Thompson 1990, pp 30). However, the means by which PDC brings meaning and value to patients is not well described, and the lack of descriptive studies of PDC from the patients' point of view underscores this.

Palliative day care is therefore an example of a complex service since it is multidimensional and is designed to be responsive to the needs of individuals and services are tailored to their individual circumstances. Also, the same input (e.g. counselling, reflexology) may provide a different magnitude or type of benefit to individuals depending on their expectations, life circumstances or experience of care. Therefore the benefits are diverse and not easily defined, even by those who provide the services.

It may be the case that the usual approaches to evaluating health care based on the measurement of disease progression, life years saved or improvements in health-related quality of life are not sensitive to the kinds of benefits that may be derived from palliative day care. While these putative benefits or changes in welfare may be hard to describe (even by the people who provide these services or by the patients themselves) this does not mean they cannot be very significant changes in welfare to the people who experience them.

What this adds up to is that the value of palliative day care to those who attend may be different from the objectively measurable changes in health gain achieved. Those attending palliative day care may have strong preferences for attending a centre but objective changes in health may not be observed. Furthermore, the value of subtle improvements in quality of life when faced with a

life-limiting illness or towards the end of life may be greater than the value of the same change in quality of life at other times of life. Therefore, instead of considering the value of health care in terms of health gain, defined in terms of additional units of time adjusted for quality, it might be more meaningful in this context to consider what an individual might be willing to give up (in terms of financial sacrifice or consumption forgone) in order to access palliative day care. This argues for the return to the fundamental idea in welfare economics of value measured by the satisfaction of preferences, as discussed further on.

The provision of these services may be seen as important to society as a whole. This is supported by the fact that hospices within the voluntary sector have been successful at fundraising from the public, even though it is clear that palliative care does not lengthen (and can even shorten) life, and may provide only a few days or weeks of benefits. Therefore there may be something of an anomaly between the value of palliative care services as measured using a health gain approach and the support for these services from the general public. The themes highlighted in box 1 below summarise the reasons why the evaluation of complex services such as PDC may be different from the evaluation of mainstream health services. This reasoning has led to the consideration of different ways of undertaking economic evaluation methods. It reflects the contrast between these services and others that are well-defined, one-dimensional interventions that are amenable to clinical trial type evaluation (Douglas and Normand 2001).

Box 1 Challenges to applying health care evaluation techniques to complex services

Health services are relatively straightforward to evaluate if they have the following characteristics:

- They are highly structured, specific interventions, delivered in mostly the same way for all patients. (Complex services can address social, psychological, spiritual and existential aspects of quality of life as well as physical health);¹
- They focus on improving physical health, with specific care pathways or algorithms to follow. (Complex services do not aim to provide the same pattern of care for all patients. The focus is on the patient as an individual);
- The outcomes are clearly defined in terms of improvements in physical health and length of life, and there are uncontested, accepted definitions of positive and adverse outcomes. (Positive and adverse outcomes are not well defined and may even be contested between the different professional groups who provide them. Complex services can be relatively low intensity interventions: for example patients may only access a service once a week alongside other clinical or social services they use more regularly);
- The focus is on the individual patient's health only and usually on a specific aspect of their health (Complex services focus on wider social or pastoral needs and the well-being of the whole family/ community).
- There may be specific difficulties in economic evaluation for interventions that occur at the end of life. These relate to the short amount of time a person has left to live and the (usually) worsening health state they experience, relative to an expectation of a full healthy life. Since any improvements are relatively small, a quality adjusted life year saved may not be a sensitive or meaningful unit of outcome (this is explored further in the text).

¹ Existential health refers to maintaining 'meaning' to life even at the end of life.

2. Does economic theory provide some insights into how to address these problems?

The origins of welfare economics are based on subtly different theories about the purpose of economics and the role of economics in deciding policy questions. Utilitarianism is founded on principles that have at their heart strong assumptions about the nature of human behaviour and how decisions about human welfare ought to be made (Hunt 1979, Black 1990, Bonner 1995). Pareto's gave a formal definition to the concept of social welfare and the conditions necessary to achieving a socially optimal level of welfare. These conditions assume that individuals are rational, welfare-maximising decision-makers, and who make decisions to trade their wealth/income for goods and services that improve their well-being (Winch 1971).

The overall aim of welfare economics has been the maximisation of the social welfare function derived from individual desires (Arrow 1963). Even with its theoretical challenges, economists have adhered with tenacity to the underpinning value judgements of Pareto welfarism (Edwards 2001). These are that every individual is the best judge of his own utility and each has a specified utility function. Social welfare is defined as the sum of individuals' utility. What has been adhered to with less enthusiasm is the notion that the utility of individuals cannot be compared (Arrow 1963, Williams 2001). One of the reasons for this has been the problem of what decision rule could replace Pareto optimality (Edwards *ibid.*). This problem will be returned to further on in the context of health economic evaluation.

Late twentieth century economists have argued that welfare economics should be rejected since it does not provide practical, value free decision rules. Sen has shown that Pareto optimality assumes that if a move makes everyone better off, society is automatically better off. This may be an unobjectionable value judgement for many, but it is a value judgement nonetheless (Sen 1987). There is also the problem that using Pareto optimality as a criterion to determine welfare isolates Hume's dictum that you cannot derive a 'should' from a fact. Therefore, in its pure form, Pareto optimality may have little to contribute to real world decision-making. Almost all social policy decisions will inevitably hurt some people, even if in a small indirect way, while helping others.

In the wake of Arrow's general possibility theorem, and other critiques, a new approach to the problems of welfare economics had to emerge. Social choice theory set out to examine the manner in which individual choice, preferences and welfare should enter into social judgements and decisions about economic matters. This theoretical development coincided with philosophical discussions of equity and ethics in economics in order to expand the subject matter and range of welfare economics.

Sen has been very influential in this field. Sen appeals to the common sense belief that people's experience and satisfactions are comparable and human beings act as if they could compare their satisfaction with others. Sen has analysed Arrow's theorem and has questioned the strict axioms he presents, especially the rejection of interpersonal comparisons of utility (Sen 1970a). Rarely are decisions made without some interpersonal comparisons of utility. Furthermore, he has argued that satisfaction of private preferences alone is an inadequate basis for social judgement. Therefore at the heart of Pareto's rules is a "liberal paradox." (Sen 1970b). This paradox is that Pareto decision rules are incompatible with even the most minimal interpretation of individual rights. Sen has suggested some modification of the Pareto rule so that an individual's choices for the satisfaction of personal preferences can be separated from an individual's preferences to be counted for social choice.

The difference in interpretation of how economics ought to develop has led to different areas of focus for economics. This divergence in pathways in the development of economics is still evident in modern welfare economics. Sen has proposed that modern welfare economics has dual origins in both ethics and engineering and that this explains the duality of concerns and interests (Sen 1987). One pathway has been to take economic thought (and economic evaluation) into the area of mathematical complexity, using methods developed in operation research and statistics. The other has been to maintain some political, ethical or moral dimension of economics and economic questions. It has been argued that questions of human well-being and welfare (what is 'good' for society) have been marginalised and the egalitarian element of the utilitarian doctrine has been sterilised by moving from utility to physical output as the object to be maximised in modern economics (Robinson 1964).

The efforts of the "extra-welfarists" in health economics have been steered towards the search for an operational definition of health gain as the unit for its cardinal measurement, thereby bypassing the problem of the impossibility of utility measurement or comparison (Hurley 1998). By attempting to replace utility with other comparative units of health, health economists in particular have moved away from traditional welfare economics. It has been suggested that this approach has made the discipline more vulnerable to "an intellectual vacuum [that] looms large in the history of economics" (Blaug 1985). A dominant focus in economic evaluation has been in the validity and reliability of measurement that has attempted to mirror physical sciences. Interest in the mathematics of human behaviour and psychology has developed within this tradition.

Consequently, it is suggested here that there is a tension between theoretical validity and practical application at the core of health economic evaluation. These tensions are reflected in the different approaches to evaluation presented further on since they are about trying to understand what is

good for society and how this can be determined. The satisfaction of individual preferences is at the core of welfare economics, but not at the core of health economic evaluation. The question of how to value complex service that aim to satisfy people's needs rather than to produce objective health benefits is especially relevant to the discussion of how to evaluate services such as palliative care.

Economic evaluation of health and social care – which economic evaluation approaches to use?

Economic evaluation of health and social care has developed in generally welfarist paradigm, within the parameters set out above. In this theoretical framework, public sector economists have had to address some general problems of finding ways to express the benefits of health and social care, and measuring and comparing different kinds of benefits across interventions that aim to improve welfare in contrasting ways. Cost-benefit analysis, as the empirical embodiment of welfare economics, would indicate what an individual or society might be willing to sacrifice (in terms of alternative consumption/ welfare forgone) in order to access a particular service or intervention. However, monetary valuation of the benefits of health care has presented a serious methodological challenge for evaluation. Instead, health economists have identified other means of expressing the benefits of care, by using notions of improved length or quality of life to conceptualise the outcomes of different interventions. Therefore the idea of value or revealed preference, central to welfare economic analysis, has been less important in the economic evaluation of health care interventions. The value of the outcomes derived from health (and this could also be applied to social care) has usually been implied rather than explicitly argued.

For some types of interventions this has not been a major problem: the definition of the outcome or purpose of the intervention has been widely accepted. The welfare or benefits derived from particular interventions are well understood by professionals and the public, and the relationship between the intervention and the outcome is straightforward. Furthermore, the value of the benefits to one person of, for example, a hip replacement operation can be assumed by most reasonable people to be comparable to the value of the operation to another. The comparison of welfare (or interpersonal comparison of utility) between individuals would be considered to be relatively unproblematic in this context.

For services such as PDC, the value to an individual may be different from the objectively measurable or observable change in health status as perceived by those undertaking the assessment. The 'value' of a health care intervention may differ depending on whose perspective is adopted, that of the individual or defined as measurable change in health status. What is seen as valuable (contributing to overall welfare) to one may not be seen as so valuable to another.

Interpersonal comparisons of utility cannot be so easily assumed away between two such individuals.

It may be the case that the dominant approaches to evaluating health care based on the measurement of life years saved and improvements in health-related quality of life are not sensitive to the nature of the benefits that may be produced by attending PDC. The kinds of benefits that may be important in PDC might be the social aspect of palliative care, bringing about a change in outlook in people who attend, or making them feel more able to cope with living with advanced disease (this idea is developed further on). While these benefits or changes in welfare may be hard to define (even by the people who provide these services) this does not mean they cannot be very significant changes in welfare to the people who experience them. The subjectively value of a health care intervention could be defined as the value to the individual receiving the intervention, which may differ between individuals.

The valuation of health states may also vary between patients who use PDC services. In an editorial on palliative care it was suggested that patients experience their objectively similar levels of illness in subjectively different ways:

“Some patients with minimal dysfunction are extremely dissatisfied while others seem quite able to tolerate severe impairment and may even feel fortunate to obtain therapy. Patients’ perceptions of their illness are extremely variable and factors other than their disability come into their perception.” (Cella 1995)

It might be more meaningful to consider what an individual might be willing to give up (in terms of financial sacrifice or consumption / benefit forgone) in order to access PDC. This will be returned to further on.

3. The ‘health related quality of life’ approach to valuing outcomes

By putting aside the problem of interpersonal comparisons of welfare, important progress has been made in economic evaluation from assuming the overall comparability of welfare between individuals. This has led to the design and validation of generic measures of outcome that are comparable across individuals and across different types of interventions. Preference-based (utility-based) measures of quality of life used in economic evaluation have focused on describing and valuing a range of health outcomes or states (Brazier, Deverill et al 1999, Blumenschein and Johannesson 1996) and now includes some measure of social concepts of well-being (Netten, Smith et al 2002). But QALYs are in essence health-related quality of life outcome units and do not incorporate the wider definitions of quality of life. The quality weights have focussed on aspects of

health that have clearly defined effects on living (such as being able to dress oneself, being able to walk unaided or to be free of depression and anxiety) (Loomes and Mackenzie 1989, Gafni 1997, Kopec and Willison 2003)

The EQ-5D is a preference-based measure of health and has been developed for the economic evaluation of health technologies (Dolan and Roberts 2002). It is a simple, five-item questionnaire that asks respondents to rank their current health status on three levels (no problems, some problems, acute problems). The health status measurement can be transformed into a preference-based weighting scale in order to derive quality adjusted life years.

There are *a priori* reasons for suggesting that the EQ-5D instrument and other similar preference-based outcome instruments (such as the Health Utilities Index) may not be appropriate tools for enumerating the outcomes of palliative care, as measured in quality adjusted life-years. These reasons relate to the notion of using a time as the measurement of benefit, and to the sensitivity of the instrument to measure changes in quality of life in domains that are important to patients in this phase of illness.

The assumption in cost-utility analysis is that the value of an additional QALY is the same regardless of when it is lived (thereby side-stepping a major problem in welfare economics of the impossibility of interpersonal comparisons of utility mentioned earlier). For palliative care interventions, that there may be reasons why people may value health state differently at the end of life. An improvement in the quality of time at the end of life may be more valuable (even if it is only a few days) than the same absolute improvement earlier in life. For example increased quality of life may be more valued in order for the dying person to be able to take part in special activities with loved ones. People may be less (or more) frightened of pain or some discomfort at the end of life than at other times. This is because people may not expect to have any improvements in quality of life at the end of their life. Also they may value more highly any additional quality of life in order to be able to do particular tasks that are important to them at the end of their lives. It may also be extremely valuable to their family and loved ones. This subjective preference cannot be captured by this approach without also undertaking the exercise to re-evaluate the QALY weightings for this group of patients.

QALYs are generated by measuring additional years of life weighted by the quality of life in a particular health state. However good (or bad) a health state is, the endpoint for the outcome is a time-based measure. Since the majority of patients who require palliative care are not expected to live long, the fact that the outcome measure is dominated in units of time is problematic. If an intervention that improves life expectancy by only a fraction (say from 6 months to a year) were to

be compared with a palliative intervention that improves quality of life, the palliative intervention would have to have an impact on quality of life equivalent of 0.5 to be equivalent (say from a health state valued as 0.2 to one valued as 0.7). An assessment of how likely this might be in a PDC setting can illustrate this point. Table 2. shows the calculation for estimating quality of life weighting for each EQ-5D health profile. Full health is given the value of 1.0. Any state of health less than full health is estimated by subtracting from 1.0. The constant term, for any state less than perfect health, is 0.081. If level 3 occurs in any domain, an additional parameter of 0.269 is subtracted. The values to be subtracted for level 2 and 3 for each domain are given in table below.

Table 2. Values for estimating weighting for each state of health. Values for the UK general population (source: Kind, P., P. Dolan, C. Gudex, A. Williams, 1998 Variation in population health status: results from a United Kingdom questionnaire survey)

EQ-5D dimension	Level 2	Level 3
Full health	1.0	1.0
Mobility	-0.069	-0.314
Self care	-0.104	-0.214
Usual activity	-0.036	-0.094
Pain/ discomfort	-0.123	-0.386
Anxiety/ depression	-0.071	-0.236
	Constant for any level two or higher, subtract 0.081	Any level 3, subtract an additional 0.269

A hypothetical PDC patient at first attendance would be expected to be able to get out of bed (with help from a carer), be able to travel, be able to eat with others and socialise to a limited extent. Very few patients will be in either acute pain/ discomfort, unable to perform any usual activities, or be completely unable to wash and dress. They may have acute anxiety or depression (level 3) and have some problems with other domains (level 2). This could be represented as a score of 2 2 2 2 3, calculated as a weighting of 0.082. Marginal improvements in this score, say in anxiety and depression (moving from level 3 to level 2) and in pain discomfort (level 2 to 1) and usual activity (level 2 to 1) could be represented as the profile 2 2 1 1 2, which has a weighting 0.675. This would represent a difference of about 0.6 QALYs if a patient was in this improved state for a year. If the patient lives for only 6 months in this health state, then the intervention would have produced around 0.3 additional QALYs. This level of improvement might be expected in PDC, given the aims and objectives of the service. This is equivalent of an intervention that increases life expectancy (or reduces premature death) by about three and a half months. The analysis demonstrates that interventions that increase length of life by only a few months will dominate life improving

interventions such as PDC even where the change in quality of life is dramatic to the individual concerned (and their family).

People who attend PDC are not expected to regain full health and may not live for a full year. They would be expected to continue to have some problems and deteriorate until death. As there are only three levels of health states for each domain in EQ-5D, the likelihood is that patients in a palliative phase of illness may improve but continue to tick the box 'some problems' even though they may have marginally more or fewer problems. This distinction would not be detectable using this three-level approach.

There are several issues to consider with regard to the sensitivity of the instrument to distinguish different states of health. If the instrument were insensitive, then it would not show changes in quality of life even when respondents reported that a PDC intervention has made an important difference to their quality of life. Also, the valuation of health states should reflect the values of patients using PDC. There has been extensive work undertaken to identify population level valuations for all 243 health states (for example, Kind, Dolan et al 1998, Dolan and Roberts 2002). Work has not yet been undertaken to assess the value of health states for people facing the end of their lives, although some work has been undertaken to assess EuroQoL values among elderly patients, mainly for the practical aspects of doing this (Coast 1998). As described above, there may be good reasons why people facing a life-limiting illness would value a slight change in their circumstances more than an otherwise healthy individual expecting to regain full health after an illness.

Also, the HRQoL approach assumes that a year of life is an additively separable unit of outcome. It has the same value regardless of when it is received or who receives it and is the same regardless of how much total life a person may be expected to live. It does not take into account the additional benefits that might arise from health care, for example, of younger people knowing that they will be cared for later in life, or knowing that less healthy people in society have access to the services they require now. In the context of services that are provided to people towards the end of their lives, the value of these other kinds of benefits may be considerable.

EQ-5D is a standardised instrument for use as a measure of health outcome designed to be applicable to a wide range of health conditions and treatments. Domains such as pain/ symptom control and anxiety/depression that are relevant to palliative care are covered by the EQ-5D instrument but it clear that other domains such as existential or spiritual health are not. These domains are not clearly understood or even accepted as relevant domains of health-related quality of life (Warr 1996, Kaasa 2002). The EQ-5D instrument should not be criticised for not achieving

something it did not set out to achieve. But the limitations of its use in PDC population are that it might not capture the domains of health-related quality of life (or quality of life more generally) that are important to those who attend a centre or are responsible for the care that it provides. However, the use of this instrument in a palliative care population has not been explored to date.

4. The choice experiment approach

The interest in choice experiments in health economics literature appears to have been driven by two areas of economic research: valuing aspects of care “beyond health outcomes”, and as a means of deriving willingness-to-pay valuations without using lengthy and cognitively challenging stated preference techniques (Ryan and Farrar 1995). Another potential use of technique is as a way of evaluating interventions and services that have multiple attributes or properties (Ryan, Scott et al 2001).. It can also be used where the research or policy problem is to value one or more attributes or where a specific combination of attributes does not currently exist in the market or new attributes are being considered (Green, Tull et al 1988, Haaijer and Wedel 2000).

A choice experiment is a stated preference technique for establishing the importance of individual attributes in the overall utility of a good or service (Cave et al 1993, Ryan 1996). Stated preference methods are “a family of techniques which use individual respondents’ statements about their preferences in a set of options to estimate utility functions.” (Kroes and Sheldon 1988). All forms of the approach are concerned with the valuation of attributes of a good or a service and transforming respondents’ subjective preferences for attributes of a service into a numerical valuation. These numerical valuations have different properties and can be interpreted in different ways depending on the approach and underlying theory. For example, marketing researchers have been concerned with disaggregated values of attributes and in designing commodities with optimum value attributes, that is, the highest overall value to customers (Green and Wind 1975) By contrast, economists have been mostly interested in the aggregate implications of multiattribute utility structures in terms of how they helped to describe the aggregate demand function for a particular good or how the results might be used in cost-benefit analysis (Cave et al 1993). Public sector economists have interpreted this approach as being a way to understand and measure the utility of a good or service where a market value may not exist (Adamovicz 2002). In other words, it may be used as a proxy for measuring monetary value directly (i.e. by individuals’ revealed behaviour in the marketplace).

Ryan and Hughes (1997) have argued that choice experiments are another method of estimating an individual’s utility in a way that is similar to standard gamble and time trade-off techniques. The approach starts from the premise that a service (or product) can be defined in terms of a few important characteristics of that service (Ryan, Bate et al 2001). An individual decides which

characteristics (or attributes) are more important to them and which are less important within constrained choices to maximise his or her utility from that service. Unlike other techniques for estimating individual utility, choice experiments can estimate the utility of specific attributes of an intervention and establish the relative importance of different attributes by estimating the marginal rate of substitution between them. Another argument for the choice experiment approach in evaluation is that it resembles “real life” choices and is therefore superior to other contingent valuation methods.

Like all current methods of estimating utility, choice experiment theory makes a number of strong assumptions about the functional form that can transform the attributes into additively separable utilities for individual attributes and sums the utility of individual attributes to arrive at the overall utility of a product. However, the conceptual framework appeals to health economists working on public policy issues because it is grounded in Lancaster’s microeconomic theory that recognises that the utility of a good or a service can be decomposed into separable characteristics (Lancaster 1966, 1971). Furthermore, it has been argued (although not yet irrefutably demonstrated through experimentation) that choice experiments are theoretically closer to cost-benefit analysis since willingness-to-pay values can be derived from the experiments (Ryan, Scott et al 2001).

On the feasibility side, one particular property of choice experiments is that they can provide a large amount of information from a relatively small amount of data from respondents (Cattin and Wittink 1982). In this sense it is a highly efficient research tool. Another property of choice experiments is that they do not rely on comparative data, either over time, or between patient groups. This makes it an attractive method for research in contexts where there is a particular problem of patients being too frail to continue in a study, or dying before they can be followed up (McWhinney 1994, Jarvis, Burge et al 1996, Grande, Todd et al 2000). One of the other advantages of this method is that it allows experimentation on data which may not exist in the real world. This means it has important advantages over other methods that use data from observation, either of outcomes, or revealed demand (Cave et al 1993). Where these data are difficult to collect, it allows information to be gathered where otherwise none might be available for decision-making purposes. Related to this, it also allows for the valuation of particular (secondary) aspects of a product or service that might be “swamped” by other primary attributes (Cave et al 1993). These properties make choice experiment techniques potentially attractive for evaluating complex health services.

The choice experiment method has a number of advantages for examining the value of complex health and social care services. Choice modelling can be used to estimate the structure of preferences for these services, that is, what is contributing most and least to the decision to access care. This re-directs the question away from one that is very difficult to answer, like “is a service

cost-effective overall?” towards a more answerable question that is valuable to decision-makers, “what particular aspects of care are valued by this group of patients?”.

An analysis of the demand for complex areas of health and social care using choice experiments may reveal the characteristics of the service that are important to patients. In a service where external measurements of health gain may not be as important as a respondent’s assessment of quality of life (welfare/utility), this approach to measuring outcomes could be enlightening . It can make progress in understanding the nature and value of the outcomes of a multidimensional service in ways may not be amenable using other economic evaluation methodologies. This hypothesis will be tested in a study using both the EQ-5D and choice experiment methods in a study of palliative day care.

The choice experiment methodology allows the question of effectiveness to be approached from a different perspective. Rather than measuring the impact of a service on an individual’s health as a proxy for utility, choice experiments can evaluate the services directly by establishing how much individuals would be willing to forego (in terms of other services or in terms of some proxy for money) in order to have these services.

There is also another potential use of this approach that is of specific interest in the context of PDC and is probably equally as important in other complex services as well. There is constant debate in the PDC field about the relative importance of the various activities it offers. Choice experiments can be used to elicit patients’ preferences for particular component activities of care that are difficult to obtain by other methods. They can by-pass the objective measurement of health status and consider individuals’ preferences for particular characteristics of a service directly. Measurement of preference for inputs or outcomes amounts to measuring the same change in welfare. For example, the attributes of the service (or arguments in patients’ demand function for PDC) may be described either as the objective characteristic of the service (information provided, presence of a doctor) or as the subjective perception of benefits from a specific attribute of a service (better clinical information or less time in pain).

For PDC research, this reasoning addresses a fundamental problem: that the measurement of outcome may not be the appropriate way to judge the value of the service to society. If, using choice experiment methods, the value (or lack of value) of PDC services can be established, then this will be a useful contribution to knowledge and decision-making. If the choice experiment includes a monetary valuation of PDC as an attribute, the results could be used in a cost-benefit analysis of PDC.

The purpose of the choice modeling experiment is to begin to systematically assess the value of specific components of PDC by considering patients' preferences for these attributes. The research question may not be framed in such a way as to determine whether PDC should be offered to all patients (its overall effectiveness) but how it could be organised to reflect patients' expressed preferences for particular attributes of the service that providers could influence. Analysis measurement of these preferences may be the closest approximation of the value of these benefits where measurement of outcome are not readily conceivable or meaningful and where (at the end of life) giving people what they want may be a more important service than giving them what is effective or "good" for them, as determined by others. So long as the aspects of care that are valued by patients who attend as contributing to utility (by whatever means), it is hard to argue that these services should not be provided to people near the end of their lives. Since these services have traditionally been provided outside the statutory health sector, by voluntary organisations, this is evidence that services for people at the end of life are also valued by people who are not experiencing this state of illness themselves.

Aspects of the provision of PDC such as its contribution to "existential health" or to aspects of psychosocial well-being are seen as important objectives of PDC by providers. These attributes are also contested by different stakeholders in the service, some of whom believe may believe in them strongly, some of whom are more sceptical about the contribution of PDC to these goals. The extent to which, for example, public funds should be used to provide spiritual well-being and contribute to wider social goals related to quality of life is not straightforward. Whether or not PDC services are valued by those who attend for these specific attributes (in terms of whether and how much people might make sacrifices in order to have these benefits) is also not known. The challenge for empirical research is to capture these subjective and highly context-specific attributes of care in a choice experiment that is, by its nature, a simplification of the real choice that users face when deciding whether to attend a centre. Health-related quality of life instruments have been developed that measure some aspects of existential health in the palliative care field (Massaro and Macmillan 2000), but no simple, overall form of words that would capture this domain in one attribute has yet been validated. This may be an important research topic in this field.

This choice experiment approach has a number of attractive practical qualities for the evaluation of PDC. First, the methodology may obtain relative values of different attributes of a service from individuals, in a way that mirrored real decision-making. Second, choice modelling is proposed as a way of estimating indirect willingness to pay values that could be incorporated into CBA analysis. Third, many of the problems of obtaining longitudinal health gain data can be avoided since the impact of a service on quality of life over time is not estimated. Fourth, the study can be designed

to look relatively simple, and, if the data could be collected, could produce sophisticated results from a few simple questions.

Conclusion

The argument that has been made in this paper is that a preference-based, welfarist approach is appropriate in the context of valuing the benefits of complex services. It is not the approach that has been used most commonly in health economics because it is difficult to elicit preferences and often unnecessary when other simpler measures as proxies for welfare will suffice (following Occam's razor²). But in the context of services that aim to meet a range of needs of individuals, a value-based approach would seem to be the most appropriate.

The danger of comparison across sectors is that the benefits gained from PDC interventions may not be measurable in objective ways, so that PDC will not score well in any league table of benefits, however defined. An important empirical question to test in the general public is whether health and quality of life at the end of life is seen as important in a 'good' society. This moves the debate back towards the normative constructs of economics – to how society ought to be. The contention is that palliative care is seen as good service by the general public (who give charitably to hospices and other voluntary organisations who care for people with advanced illness) even though its benefits cannot easily be defined or measured. The problems of evaluation reviewed here have illustrated the point that all approaches to the measurement of human experience are fallible, and that the focus on the scientific measurement and quantification of health outputs can lead to policy conclusions that may not reflect society's wishes.

This paper has situated the problem of evaluating complex services within more general debates about how to evaluate health care. It has been argued that these debates have been driven by a tension in health economic evaluation between theory and practical research. On the one hand, it is important to base research on a theory that can determine whether a policy is good for society (and to specify clearly what this means in terms of maximising social welfare), and on the other to undertake research that can answer real world problems in ways that people can understand and that answers the important questions for policy. If health gain can be seen as a proxy for the value of benefit derived from health gain, then the relationship between the satisfaction of preferences and objectively measurable health is not problematic. In areas where outcomes are not easy to define, the relationship between health gain and preferences/ welfare may not be so direct. The consequence is that services that provide benefits in ways that cannot be measured using the

² The principle states that one should not make more assumptions than the minimum needed. This principle is often called the principle of parsimony. In his writings, Occam stressed the Aristotelian principle that entities must not be multiplied beyond what is necessary, to choose the simplest solution to a problem.

quality of life instruments currently available may lose out in the competition for resources if they cannot demonstrate their value in other ways.

The paper has also critiqued the use of health gain approaches for services that have a range of goals and where the main aim is to meet the needs of the individual. As the evaluation culture spreads to more services that cross the divide between health and social care, the health gain/cost-effectiveness/cost-utility paradigm may not provide insightful results in other complex care settings. Other methods such as choice modelling may need to be incorporated into the group of acceptable methods of economic evaluation by institutions such as NICE that currently favour cost-effectiveness and cost-utility methods (Birch and Gafni 2002). The debate will continue about when and how values should be incorporated into economic analyses, and how important they are. The empirical study will contribute to this debate by exploring how PDC and similar services that aims to provide small but potentially valuable benefits can compete for resources if a strictly health gain approach is adopted. An approach to evaluation that adopts a preference-based approach may be a more useful way of obtaining evidence of the value of specific aspects of a complex service.

References

- Adamowicz, V. (2002). Design and analysis of choice experiments: lessons from environmental economics. In *Discrete Choice/ Conjoint Analysis workshop*: Odense, Denmark.
- Arrow, K.J. (1963). Kenneth Arrow on welfare economics. *Am Econ Rev*, **53**, 941-973.
- Birch, S. & Gafni, A. (2002). On being NICE in the UK: guidelines for technology appraisal for the NHS in England and Wales. *Health Econ*, **11**, 185-91.
- Black, R.D.C. (1990). Jevons, Marshall and the utilitarian tradition. *Scottish Journal of Political Economy*, **37**, 5-17.
- Blaug, M. (1985). *Economic theory in retrospect*. Cambridge University Press: Cambridge.
- Blumenschein, K. & Johannesson, M. (1996). Incorporating quality of life changes into economic evaluations of health care: an overview. *Health Policy*, **36**, 155-66.
- Bonner, J. (1995). *Economic efficiency and social justice: the development of utilitarian ideas in economics from Betham to Edgeworth*. Edward Elgar: Aldershot, Hants.
- Brazier, J., Deverill, M., Green, C., Harper, R. & Booth, A. (1999). A review of the use of health status measures in economic evaluation. *Health Technol Assess*, **3**, i-iv, 1-164.
- Cave, M., Cunningham, D., Buxton, M., Hanney, S., Pollit, C., Scanlan, M. & Shurmer, M. (1993). *The valuation of changes in quality in the public services*. Vol. H.M. Treasury. HMSO: London.
- Cella, D.F. (1995). Measuring quality of life in palliative care. *Semin Oncol*, **22**, 73-81.
- Coast, J., Peters, T.J., Richards, S.H. & Gunnell, D.J. (1998). Use of the EuroQoL among elderly acute care patients. *Qual Life Res*, **7**, 1-10.
- Dolan, P. & Roberts, J. (2002). Modelling valuations for Eq-5d health states: an alternative model using differences in valuations. *Med Care*, **40**, 442-6.
- Douglas, H.R., Higginson, I.J., Myers, K. & Normand, C. (2000). Assessing structure, process and outcome in palliative day care: a pilot study for a multicentre trial. *Health Soc Care Community*, **8**, 336-344.
- Douglas, H.-R. & Normand, C.E. (2001). The economics of palliative day care. In *Palliative Day Care in Practice*, Hearn, J. & Myers, K. (eds). Oxford University Press: Oxford.
- Edwards, R.T. (2001). Paradigms and research programmes: is it time to move from health care economics to health economics? *Health Econ*, **10**, 635-49.
- Fisher, R.A. & McDaid, P. (1996). *Palliative day care*. Arnold: London.
- Gafni, A. (1997). Alternatives to the QALY measure for economic evaluations. *Support Care Cancer*, **5**, 105-11.
- Grande, G.E. & Todd, C.J. (2000). Why are trials in palliative care so difficult? *Palliat Med*, **14**, 69-74.
- Green, P.E., Tull, D.S. & Albaum, G. (1988). *Research for marketing decisions*. Prentice Hall: New Jersey.
- Green, P.E. & Wind, Y. (1975). New ways of measuring consumer judgement. *Harvard Business Review*, **53**, 107-117.
- Haaijer, R. & Wedel, M. (2000). Conjoint choice experiments: general characteristics and alternative model specification. In *Conjoint measurement - methods and applications*, Gustafsson, A., Herrmann, A. & Huber, F. (eds). Springer: Berlin.
- Hearn, J. & Higginson, I.J. (1999). Development and validation of a core outcome measure for palliative care: the palliative care outcome scale. Palliative Care Core Audit Project Advisory Group. *Qual Health Care*, **8**, 219-27.
- Higginson, I. (1993). Palliative care: a review of past changes and future trends. *J Public Health Med*, **15**, 3-8.
- Hunt, E.K. (1979). Pure versus eclectic utilitarianism: the writing of Bastiat and Mill. In *History of economic thought: a critical perspective*. Wadsworth Publishing Company: Belmont, CA.
- Hurley, J. (1998). Welfarism, extra-welfarism and evaluative economic analysis in the health sector. In *Health, health care and health economics*, Barer, M.L., Getzen, T.E. & Stoddart, G.L. (eds). John Wiley: Chichester.
- Jarvis, H., Burge, F.I. & Scott, C.A. (1996). Evaluating a palliative care program: methodology and limitations. *J Palliat Care*, **12**, 23-33.

- Kaasa, S. & Loge, J.H. (2002). Quality-of-life assessment in palliative care. *Lancet Oncol*, **3**, 175-82.
- Kind, P., Dolan, P., Gudex, C. & Williams, A. (1998). Variations in population health status: results from a United Kingdom national questionnaire survey. *Bmj*, **316**, 736-41.
- Kopec, J.A. & Willison, K.D. (2003). A comparative review of four preference-weighted measures of health-related quality of life. *J Clin Epidemiol*, **56**, 317-25.
- Kroes, E.P. & Sheldon, R.J. (1988). Stated preference methods: an introduction. *Journal of transport economics*, **22**, 11-25.
- Lancaster, K. (1966). A new approach to consumer theory. *Journal of Political Economy*, **74**, 134-157.
- Lancaster, K. (1971). *Consumer demand: a new approach*. Columbia University Press: London.
- Loomes, G. & McKenzie, L. (1989). The use of QALYs in health care decision making. *Soc Sci Med*, **28**, 299-308.
- Massaro, T. & McMillan, S. (2000). Instruments for assessing quality of life in palliative care settings. *Int J Palliat Nurs*, **6**, 429-33.
- McWhinney, I.R., Bass, M.J. & Donner, A. (1994). Evaluation of a palliative care service: problems and pitfalls. *Bmj*, **309**, 1340-2.
- Netten, A., Smith, P., Healey, A. & Knapp, M. (2002). OPUS: A measure of social care outcome for older people. Discussion Paper 1690. Personal and Social Services Research Unit: University of Kent at Canterbury.
- Robbins, M.A. (1998). The economics of palliative care. In *Oxford Textbook of Palliative Medicine*, Doyle, D., Hanks, G. & MacDonald, N. (eds) pp. 55-65. Oxford University press: Oxford.
- Robinson. (1964). *Economic Philosophy*. Doubleday: Garden City, New York.
- Ryan, M. (1996). *Using consumer preferences in health care decision making: the application of conjoint analysis*. Office of Health Economics: London.
- Ryan, M., Bate, A., Eastmond, C.J. & Ludbrook, A. (2001a). Use of discrete choice experiments to elicit preferences. *Qual Health Care*, **10 Suppl 1**, i55-60.
- Ryan, M. & Farrar, S. (1995). A pilot study using conjoint analysis to establish the views of users in the provision of orthodontic services in Grampian. In *Health Economists' Study Group Meeting*: University of Bristol.
- Ryan, M. & Hughes, J. (1997). Using conjoint analysis to assess women's preferences for miscarriage management. *Health Econ*, **6**, 261-73.
- Ryan, M., Scott, D.A., Reeves, C., Bate, A., van Teijlingen, E.R., Russell, E.M., Napper, M. & Robb, C.M. (2001b). Eliciting public preferences for healthcare: a systematic review of techniques. *Health Technol Assess*, **5**, 1-186.
- Sen, A. (1970a). *Collective choice and social welfare*. Oliver and B.: London.
- Sen, A. (1970b). The impossibility of the Paretian liberal. *Journal of Political Economy*, **78**, 152-157.
- Sen, A. (1987). *On ethics and economics*. Blackwell: Oxford.
- Tebbit, P. (1999). *Palliative Care 2000. Commissioning through partnership*. National Council for Hospice and Specialist Palliative Care: London.
- Thompson, B. (1990). Hospice day care. *The American Journal of Hospice Care*, **Jan/Feb**, 28-30.
- Warr, D. & Feld, R. (1996). The role of supportive care in quality of life. *Curr Opin Oncol*, **8**, 279-83.
- Williams, A. (2001). The 'fair innings argument' deserves a fairer hearing! Comments by Alan Williams on Nord and Johannesson. *Health Econ*, **10**, 583-5.
- Winch, D. (1971). *Analytical welfare economics*. Penguin: London.