UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE, HEALTH AND LIFE SCIENCES

School of Medicine

Justice, Fairness and Equity in Health Care: Exploring the Social Value of Health Care Interventions

by

COLIN GREEN

Thesis for the degree of Doctor of Philosophy

June 2007

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE, HEALTH AND LIFE SCIENCES
SCHOOL OF MEDICINE

Doctor of Philosophy

JUSTICE, FAIRNESS AND EQUITY IN HEALTH CARE: EXPLORING THE SOCIAL VALUE OF HEALTH CARE INTERVENTIONS

Colin Green

This thesis is motivated by the need in many health care systems, but especially in the UK NHS, to make difficult choices over the use of limited resources. The starting point for the thesis is that when making difficult choices over the provision of health care, the overall value of health care interventions to society is a function not only of the total benefits available from health care, but also the distribution of health care resources across different groups in society. The thesis investigates this proposition that 'distribution matters' and presents research to consider the social value of health care interventions.

The research in the thesis is undertaken within the analytical framework of health economics, and in the context of health policy decisions over the funding of health care interventions in the UK NHS. The health technology appraisal process is used as an example of an allocation problem, and the thesis uses the UK National Institute for Health and Clinical Excellence (NICE) as an example of the health technology appraisal process.

A variety of methods are used, including an assessment of general theories of justice, a systematic review of the literature on empirical assessment of distributive preferences, an empirical study to investigate issues around the specific social value related to the severity of health condition, a discrete choice experiment (DCE) to explore a range of key social values and the relative weights placed on these social values. The research is drawn together in a policy-relevant analysis of social preferences and NHS decision-making.

The thesis makes a contribution to the health economics literature and to the health policy literature. It relates general theories of justice to the process of health technology appraisal. It draws together a broad and complex literature, and characterises the literature according to the general quality of the methods used. The thesis contributes to the empirical evidence base on severity of health as an important social value. It develops a hypothesis that the empirical evidence against the importance of severity of health may be a proxy preference for giving priority to a worst off group of patients in health care priority setting; providing empirical evidence to support this hypothesis. The DCE, in a sample of the general public, finds support for using the social values around level of health improvement, value for money, severity of health, and the availability of other treatments, to offer an insight to the societal value of health care interventions. The level of health improvement and value for money had the greatest impact, in the discrete choice analysis, with severity of health condition also shown to have an important role in distributive preferences.

The research contributes to the empirical evidence on the relative importance of social values in the context of difficult priority setting decisions, and it contributes to the literature on the use of the DCE framework to elicit social preferences. The thesis extends the current evidence base by using the results from the DCE to derive a measure of 'strength of preference' across health care interventions described using the experimental design used. The thesis demonstrates how such data may be used in a policy-relevant manner.

The research in the thesis provides a greater understanding over what may be meant by equity in the allocation of health care resources, in the framework of health technology appraisal, through consideration of equity as a balance between competing social values, amidst consideration of opportunity costs.

Table of Contents

List of Tables

List of Figures

Appendices

Conference Presentations

Green C, Gerard K. Exploring the social value of health interventions: A stated preference discrete choice experiment. International Health Economic Association (iHEA), Annual Meeting (Copenhagen), July 2007 (oral presentation).

Green C, Gerard K. Exploring the social value of health interventions: A stated preference discrete choice experiment. Paper presented at the UK Health Economists Study Group (HESG), University of York, July 2006.

Green C. Exploring the social value of health care interventions. University of Southampton, School of Medicine, Postgraduate Conference. June, 2006.

Green C. Peoples preferences for fairness: examining the empirical evidence on distributive preferences. Society for Medical Decision Making, San Francisco, USA, October 21-24, 2005 (poster presentation).

Green C. Health technology appraisal: should the NHS (e.g. NICE) be maximising health gain? University of Southampton, School of Medicine, Postgraduate Conference. June, 2005.

Green C. Justice, fairness and equity: Utilitarianism in public health decisions. UK Public Health Association, Brighton, April 2004, (oral presentation).

Green C. Health and equity: Considering health technology appraisal. Paper presented to the UK Health Economists Study Group (HESG), Brunel University (London), July 2002.

ACKNOWLEDGEMENTS

I am grateful to my supervisors Dr Karen Gerard (University of Southampton), Professor James Raftery (University of Southampton), and Professor Alan Hamlin (University of Manchester, formerly University of Southampton). Particular thanks to Alan Hamlin for his broad perspective, and general enthusiasm, in the early part of the research, and to Karen Gerard for her encouragement and support during the final two years of my study.

Thanks are also due to Professor Norman Waugh (University of Aberdeen) and Professor Emeritus John Gabbay (both formerly University of Southampton) for their support in the supervision of the thesis.

Thanks to Ipsos MORI Social Research Institute (London) for their help in data collection, and especially to Gregor Jackson for his support in the planning and organisation of the data collection element of the thesis.

Thanks are due to The Wessex Institute for Health Research and Development, University of Southampton, for sponsoring me throughout the PhD, with thanks especially due for finanical support to enable funding of the data collection process in the empirical research.

Thanks are due to Alison Price, Information Scientist, University of Southampton, for assistance in the literature searching, and to Liz Hodson, University of Southampton, for help with literature retrival.

Thanks are due to the participants in the main empirical research, and in the pilot and pre-pilot activities.

Thanks are also due to Dr Peter Davidson (NCCHTA, University of Southampton) and Professor Andrew Stevens (University of Birmingham), for their advice on the Public Health and decision-making elements of the NHS.

Thanks to colleagues in the Peninsula Medical School, University of Exeter, and especially to Professor Ken Stein, for their support and encouragement during the final six months of study.

Special thanks to my wife Sara, and my children Joseph and Esme, for their unfailing support, without which this thesis would not have been possible.

ABBREVIATIONS

BMA	British Medical Association
CBA	Cost benefit analysis
CEA	Cost effectiveness analysis
CLM	Conditional logit model
CUA	Cost utility analysis
DALY	Disability-adjusted life-year
DCE	Discrete choice experiment
DTC	Drugs and therapeutics committee
HTA	Health technology assessment
HTAP	Health technology appraisal
HYE	Healthy years equivalent
ICER	Incremental cost effectiveness ration
MRS	Marginal rate of substitution
MS	Multiple sclerosis
MND	Motor neurone disease
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
PCT	Primary care trust
PPI	Potential Pareto improvement
PTO	Person trade-off
QALY	Quality-adjusted life-year
RUT	Random utility theory
SHA	Strategic health authority
SWF	Social welfare function
UK	United Kingdom
UPF	Utility possibility frontier
VFM	Value for money
WHO	World Health Organisation

1 AIMS

The starting point for this thesis is that society's overall valuation of the benefits available from health care is a function not only of total health benefits from health care, but also of the distribution of health benefits across different individuals and groups in society. The aim of this research is to explore the social value of health care interventions. It sets out to achieve this, through a variety of methods, within the analytical framework of health economics, and in the context of health policy decisions over the funding of health interventions in the UK National Health Service (NHS).

A variety of methods are used: an assessment of general theories of justice and fairness; a systematic review of the literature on empirical assessment of distributive preferences; an empirical study investigating a specific distributive preference; an empirical study exploring the social preferences, and the relative weight placed on key social values, of the general public; and a policy-relevant analysis of social preferences and NHS decision-making.

The primary research questions that motivate the thesis are:

1. What are the social values that can be used to set health care priorities over the funding of health care interventions in the UK NHS?

2. Where social values are identified, for use in decisions over the funding of health interventions, what are their relative values (what are the trade-offs)?

In addressing these primary research questions the thesis also considers the following related questions:

i) Is the maximisation of health (health gain) a valid representation of the social value of health interventions?
ii) Is the social value of a health intervention dependent on factors other than health?
iii) Is the social value of a health intervention related to the characteristics of eligible patient groups?
iv) Is the social value of a health intervention related to the characteristics of the health intervention?

2 INTRODUCTION

This chapter sets the context for the research in the thesis by outlining the need for and scope of the research (sections 2.1 & 2.2), the structure of the thesis (section 2.4), and by providing a brief description of the key concepts that are central to this area of study (section 2.3).

2.1 Need for research on the social value of health interventions

In health care systems around the world decision makers are faced with competing demands and insufficient resources, even in the richest countries, and in these circumstances it is not possible to provide all available and potentially beneficial health care to those who could benefit from it (Hauck *et al*, 2003). This means priority-setting is inevitable, and difficult decisions have to be made. This is the case in the UK National Health Service (NHS), where in some cases access to effective health care is restricted or denied (New 1997, Ham & Robert 2003, Martin *et al* 2002, Newdick 2005).

In the UK NHS, health policy makers at local or regional levels, or national bodies such as the National Institute for Health and Clinical Excellence (NICE), are responsible for making difficult, and in some cases highly contentious, resource allocation decisions. The basis for these decisions is often not clear (Ham 1997, Hope *et al* 1998, Birch & Gafni 2002, Dakin *et al* 2006). The NHS more broadly does not have any clear criteria or stable principles upon which to respond to these priority-setting dilemmas (i.e. being unable to fund all treatments), (New 1997, BMA 2001, Newdick 2005).

The political nature of decisions over the availability of specific health care interventions within the NHS, and the growing media coverage of the consequences of these decisions, has raised the awareness of the general public on both the need to set priorities, and the difficulties associated with doing so. For example, there have been high profile cases such as the 'Child B' case, where there was an apparent denial of treatment for a particular individual (Pickard & Sheaff, 1999), and the media attention given to the recent issue of the availability of trastuzumab (Herceptin) for the treatment of early stage breast cancer (e.g. Barrett *et al,* 2006). These priority setting dilemmas, and the increasing pressures on the health care budget, have led to a growing demand for clearer and more transparent priority setting over health care. One of the policy objectives of the NHS is to take into account the views of the general public when setting priorities over the use of health care resources (Department of Health 1999, 2000). Therefore, there is a clear need to establish how

society values different health care interventions i.e. how the general public would wish to set priorities.

Public and patient involvement in NHS decision making, in the planning of changes to the organisation and delivery of health care, has been a prominent policy objective from the early 1990's (Department of Health 1992, 1999, 2000). For example, *The Patients Charter*, and the *Local Voices* initiative (Department of Health, 1992) have demonstrated that Government policy has been explicit over the involvement of the public in decisions over health care, and on the importance of taking into account the attitudes of the general public over health care priority-setting. The Health and Social Care Act 2002 (The Stationery Office, 2001) states that public involvement is a duty of Health Authorities and Trusts. The National Health Services Reform and Health Care Professions Act 2002 (The Stationery Office, 2002) sets out The Commission for Patient and Public Involvement in Health to promote the involvement of members of the public in health care decision making. Within the NHS, and important in the context of the research in this thesis, NICE state that "Underlying all [NICE] decisions ... is one fundamental social value judgment: that advice from NICE to the NHS should embody values that are generally held by the population of the NHS" (Rawlins & Culyer, 2004, p226).

Whilst public involvement in heath care decision making is a policy objective of the UK Government, and related health care institutions, there is an absence of empirical evidence on how the public may value different health care interventions (Sassi *et al* 2001, Schwappach 2002, Dolan *et al* 2005). Recent reviews of the literature around social values and distributive preferences have recommended that research is needed into the social values that may inform priority-setting in health care, and the relationship (i.e. relative values) between key social values (Sassi *et al* 2001, Schwappach 2002, Dolan *et al* 2005). A recent review on priority-setting for health care concludes that there is a need to make "progress in eliciting the public's views about what constitutes a 'fair' distribution of health and health care" (Hauck *et al* 2003, p35). This thesis seeks to inform health care decision-making, at a policy level, through research to identify key social values, relevant when comparing alternative heath care interventions, and the general relationship between them.

In the health care literature it is common for studies to use the term equity as a conception of social value (e.g. in discussion of the equity versus efficiency trade-off) without any specification of the values that the term is used to support (or represent). In the provision of health care, 'equity' is a widely acknowledged and important goal (Dixon *et al,* 2003).

However, the definition of equity, and what it is that is meant by equity in the context of health care provision (decision making), is often not clear. It is difficult to disagree with some simple and theoretical expositions of equity. For example, where equity is defined as the absence of inequity (Starfield *et al,* 2000), or where vertical equity[1] is presented as the unequal treatment of unequals proportionate to need. But useful definitions of equity, useful to those facing allocation problems, are uncommon. This thesis, through consideration of social values relevant for health care decision making, seeks to provide a greater understanding of the notion of equity in health care resource allocation problems (e.g. health technology appraisal).

2.2 Scope of research

It is beyond the scope of this thesis to consider the wide range of resource allocation and priority setting decisions in the UK NHS. Therefore the research is set against the context of the process of health technology appraisal in the UK NHS (e.g. NICE), in order to consider a specific type of allocation problem. In the process of health technology appraisal the decision maker arrives at a judgment on whether a health technology can be recommended as an appropriate use of resources. In the NHS the most prominent programme of health technology appraisal is that undertaken within NICE (NICE, 2004). This thesis will refer to, and use, the NICE health technology appraisal process as an example of the appraisal process within the UK NHS (see below), and a general example of an allocation problem in health care.

To consider such allocation problems, the thesis explores the notions of justice, fairness and equity in the context of health technology appraisal. It specifically examines equity, looking at what it is that may be meant by equity, and how it overlaps with justice and fairness.

Whereas conceptions of justice and fairness may be broadly defined, or interpreted, to cover a wide range of decisions, equity is considered here to be largely context-specific (Young 1995, Konow 2000). Furthermore, the use of the term equity in the setting of health policy decisions should be accompanied by some explanation of the social values underpinning the equity arguments involved in arriving at the decision (otherwise ambiguity and confusion will continue). These considerations lead to the core contribution of the thesis; to explore key social values that could be used to choose which health interventions should be offered and which should not. The thesis considers the way that these social values may conflict with one another, and the relative value that the general public may place on a number of these

[1] Vertical equity addresses the question of the extent to which individuals who are unequal in society should be treated differently (unequally). Horizontal equity is concerned with the equal treatment of equal need.

social values, when faced with difficult decisions over whether or not to fund a health technology (i.e. the process of health technology appraisal).

This thesis reviews common theories of justice, and conceptions of fairness, to provide support for a view that there should be a clear understanding of the social values (criteria) used to make decisions over social arrangements and specified alternative states of the world (e.g. a state of the world where a health intervention is provided versus a state where it is not provided). These resource allocation decisions are often stated to be made on the basis of trade-offs between efficiency and equity, but what are the social values that are relevant to a judgment on equity? The thesis presents a detailed review of the empirical studies in the health economics literature, to inform on the social values that may be appropriate in allocation problems. The thesis explores these social values more generally, considering which of these social values may be key criteria for the purposes of health technology appraisal, and it presents empirical work which seeks to establish the relative value that may be placed on a number of these criteria. The thesis applies both attitudinal survey methods and more advanced experimental techniques to make contributions to the empirical evidence available to inform on social values and their relative importance.

The thesis considers the validity of the use of health maximisation as a prominent decision making objective. There is a growing literature demonstrating limited societal support for the maximisation of health production (e.g. quality-adjusted life-years) from available resources, when persons are faced with difficult decision making scenarios involving different groups in society (Chapter 4). This literature indicates that it is not just the production of health benefits that matters to society. In many cases the distribution of resources across different groups in society is important, regardless of foregone efficiency gains i.e. society has distributive preferences over health care and is prepared to trade-off (sacrifice) units of health production in order to pursue distributive preferences (e.g. a preference to treat the most severely ill persons).

At this point it is appropriate to note what the thesis will not cover. There is a vast literature on this area of research, such that limits must be set. This means that it is not possible, however interesting, to investigate or debate issues such as broader notions of justice, rights, norms and rules in society that are found in the general theories of justice. Neither will the thesis consider legal issues pertinent to society and the provision of health care or process in the provision of health care. Procedural justice and procedural fairness are two aspects which get some coverage in the thesis but do not form a key element of it, and the thesis will rely on citations to relevant work by others to support the views expressed.

Finally, and importantly, whilst I would argue alongside many others for greater levels of funding for the provision of health care I do not raise such issues in this thesis. Instead the thesis assumes that whatever the level of funding available for the provision of health care, difficult allocation problems would still persist.

2.3 Framework and basic concepts

The thesis uses the analytical framework of health economics, in the context of the UK NHS. The thesis considers the resource allocation context of health technology appraisal within the NHS, which provides guidance on whether specific health technologies should be available for treatment as part of the NHS funded health care. The main component parts of the analytical framework and the NHS context are set out below, at an introductory level, to provide an introduction for the remainder of the thesis.

2.3.1 Health

Health may be described in a number of different ways as a scientific or sociological phenomenon. Whilst medicine has taken a traditionally biological approach to health, a broader view of health in terms of functional definitions and normal social functioning is now more common. The definition of health itself is not of prime concern for this thesis, and an adequate basis for the research here is the widely accepted definition of health from the World Health Organisation (WHO), who state that "health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" (WHO, 1999). The WHO modified this original statement to include the ability to lead a "socially and economically productive life" (WHO, 2002). These definitions have been criticised for being too idealistic, but the definitions do emphasise the holistic nature of health, and its positive aspects.

When considering the allocation of resources, in this thesis, the focus is on the distribution of health care, rather than the broader ideal of achieving a fair allocation or distribution of health (or good health). Importantly, whilst the research in this thesis considers health care, it is accepted that health is determined by a number of factors, including biological and genetic factors, lifestyle and behaviour, the environment, social and economic factors, and health care. In many of these factors the concepts of justice, fairness and equity, may be important, although it is only on the latter issue of health care that this thesis will focus.

2.3.2 Health care

It is helpful to call on the WHO to provide a definition of health care. The WHO state that health care embraces all the goods and services designed to promote health, including preventive, curative and palliative interventions, whether directed to individuals or to populations (WHO, 2000).

2.3.3 Health technology

Health technology is now an internationally recognised term that covers any method used by those working in health services to promote health, prevent and treat disease and improve rehabilitation and long-term care. Health technologies include the activities of the full range of health care professionals, and the use of equipment, pharmaceutical and health care processes and procedures generally. Health technologies in this context are not confined to new drugs or pieces of sophisticated equipment. In this thesis the terms health technology and health intervention are used interchangeably.

2.3.4 The UK National Health Service

The UK NHS is a publicly funded national health care system, with services (almost entirely) free at the point of delivery. It was founded in 1948 with the objectives of universal coverage and equity of access (Ham & Robert, 2003). The NHS is organised principally under the National Health Services Act 1977 (Newdick, 2005), however the NHS has been in a constant state of reform from 1980 (Webster, 2002). Given the constant state of reorganisation in the NHS, the thesis does not document in any detail the previous or current organisational arrangements for the NHS, other than to state that it is a publicly funded system, and to map out the main organisational arrangements and structures relevant for the context of health technology appraisal in the NHS, which is used for context in this thesis.

The current organisation of services within the NHS is shaped by the modernisation program set out in the *NHS Plan* of July 2000 (Department of Health, 2000). The *NHS Plan* set out a full-scale modernisation of the NHS, to transform the way it cares for patients. It states that it moves the NHS to a patient-centred health service, retaining the commitment to the founding principles of the NHS, summarised as[2]:

[2] The stated NHS core principles are set out by the NHS in a more detail at www.nhs.uk/England/AboutTheNhs/CorePrinciples (last accessed March 2007).

The provision of quality care that:

- Meets the needs of everyone
- Is free at the point of need
- And is based on a patient's clinical need not their ability to pay

The total budget for the NHS in 2006-07 was £104 billion,[3] with over 80% of the budget (for England) controlled by Primary Care Trusts (PCTs). PCTs are local organisations at the centre of the NHS, responsible, together with other local agencies (e.g. local authorities), for the provision of health and social care locally, to meet the needs of the local community. PCTs have been in place since 2002, reporting to a local Strategic Health Authority. PCTs are responsible for ensuring adequate provision of hospital care, dentists, opticians, mental health services, NHS Walk-in Centres, NHS direct, patient transport (e.g. A&E services), population screening, and pharmacies. As well as purchasing and monitoring NHS services, they also support GP practices, NHS Acute Trusts, and other parts of the NHS in the delivery of care to their local patient communities (www.nhs.uk).

Strategic Health Authorities (SHAs) are responsible for managing and setting the strategic direction of the NHS in their region, they support PCTs and other NHS organisations and make sure they are working effectively. One of the responsibilities of the SHAs is to ensure national priorities are fully reflected in local health service plans. Recent developments in the NHS (July 2006) have led to a reduction in the number of SHAs in England, with 10 SHAs now providing a link between the NHS in England and the Department of Health.

The Department of Health is a Department of State, a Government organisation. It is accountable to the public and the Government, and is responsible for health and health care across the NHS, and for social care, as well as the public health role. Its aim is to improve the health and wellbeing of the people of England (www.dh.gov.uk). It sets national standards (e.g. through National Service Frameworks) and is responsible for the policy and direction of the NHS and social care services. The Department of Health does not directly run the NHS (and social care) but it does act as a form of central HQ, setting out and communicating the overall strategic direction of the NHS. Importantly, for the current thesis, the Department of Health has to ensure that public money is spent wisely and efficiently. It sets out to get the highest quality health and social care for patients and service users, at the lowest possible cost to taxpayers, and it seeks to ensure the expenditure on health and social care represents value for money (www.dh.gov.uk).

[3] Source: http://budget2007.treasury.gov.uk, accessed 30/03/07

2.3.5 Health technology 'assessment' in the UK NHS

Health technology assessment is the process of evaluating the clinical, economic, and other evidence relating to the use of a health technology. It primarily provides research information about the clinical effectiveness and cost-effectiveness of health technologies, compared to the relevant alternative course of action, for those who plan, provide or receive care in the UK NHS. It may also provide specific information on costs, epidemiology, the characteristics of the relevant patient group, and the broader impact of the technology. Health technology assessment, as practiced in the UK, provides information for use in decision-making, but does not lead to any decisions directly.

2.3.6 Health technology 'appraisal' in the UK NHS

Health technology appraisal is a common process in the UK, and it is directly involved in policy decisions and guidance. It is undertaken, either implicitly or explicitly, at all decision making levels within the UK NHS (e.g. NICE, Primary Care Trusts, drug therapeutic committees). Limited health care resources must be assigned to competing demands, and decision-makers are continually in a position whereby they must come to some judgment on whether health technologies can be recommended as a cost-effective use of NHS resources in general, or for specific indications, or for defined subgroups of patients. The process of health technology appraisal informs these judgments.

Health technology appraisal is a broader process than health technology assessment, the latter contributing to the evidence base available for the former. Health technology appraisal, unlike assessment, does not predominantly constrain itself to the direct assessment of a technology based on clinical and cost-effectiveness, although these are central elements in any health technology appraisal undertaken. Health technology appraisal is only appropriate where technologies have been shown to be safe (efficacious) and effective, i.e. they offer some health benefit (e.g. via results from clinical trials). Where this thesis uses the process of health technology appraisal as an example of a resource allocation problem in health care, it assumes a position whereby health technologies have demonstrated an acceptable safety profile, and been shown to be effective (if this is not the case health technologies should not be subject to appraisal) (Culyer, 2001). It is thereafter necessary to consider the benefits available in the context of health policy.

Health technology appraisal involves both scientific and social value judgments in the appraisal of a health technology, and it results in guidance as to the overall value of a technology to a particular population, subject to the objectives of the organisation or the

health care system it is servicing. In the process of appraisal there may be considerations related to institutional, political or social priorities.

Health technology appraisal may mirror the narrower process of technology assessment if: (i) resources are infinite, whereby all effective treatments would be universally available, or (ii) efficiency is the sole objective of the health care system, whereby cost-effectiveness would be the only input to a 'funding decision'.
However, neither of these positions adequately describes the prevailing health policy environment in the UK.

In the UK the NICE health technology appraisal programme is the most public and prominent process for health technology appraisal. Whilst there will be some variation in the process of health technology appraisal within the NHS, and it is accepted here that the methodology of health technology appraisal continues to develop, the NICE health technology appraisal programme is used in this thesis to demonstrate the appraisal process, as a resource allocation problem in the delivery of health care.

NICE was set up as a special health authority for England and Wales in 1999. It has three main functions: to appraise new technologies, to produce and approve guidelines, and to encourage improvement in quality (Raftery, 2001). As a special health authority it is part of the Department of Health. It's technology appraisal programme is charged with the task of making recommendations on use in the NHS of particular health technologies (selected new and established technologies).

In reaching its appraisal recommendations (judgments) NICE have to take into account the factors listed in the Secretary of State and National Assembly for Wales' Directions (NICE, 2000), namely:

- the Secretary of State and National Assembly for Wales' broad clinical priorities (as set out for instance in National Priorities Guidance and in National Service Frameworks, or any specific guidance on individual referrals);
- the degree of clinical need of the patients with the condition under consideration;
- the broad balance of benefits and costs;
- any guidance from the Secretary of State and National Assembly for Wales on resources likely to be available and on such other matters as they may think fit;
- the effective use of available resources.

A further factor, which the Institute will take into account in its appraisal, is the wish to be sympathetic to the longer-term interests of the NHS in encouraging innovation of good value to patients (NICE, 2004).

NICE have published reports detailing the appraisal process, and methodological guidance for submissions made to NICE (NICE, 2004). NICE have also issued guidelines to its Advisory bodies on scientific and social value judgments within the appraisal process (NICE, 2005). However, whilst the mandatory guidance on health technologies issued by NICE is said to be based on clinical evidence, cost-effectiveness and other considerations (NICE, 2004), the exact factors considered, their relative importance and trade-offs between them are not made explicit (at the present time) (e.g. Dakin *et al* 2006, p352).

2.3.7 Health economics

Economics is the social science that studies the production, distribution, and consumption of goods and services. It is commonly referred to as being concerned with scarcity, and the distribution of scarce resources. Economics can be divided into a number of streams of activity, including microeconomics, macroeconomics, positive economics ("what is"), and normative economics ("what ought to be"), although all areas are interlinked. Economics seeks to explain how economies operate and the relations between economic agents in economies, and society at large. Economics takes an explicit analytical approach, and is often 'reductionist', abstracting from wider issues to focus on specific questions. Methods of economic analysis have been increasingly applied to specific fields of interest that involve choices in a social context (e.g. education, crime, health).

Health economics is the branch of economics concerned with issues related to scarcity in the allocation of resources for health care. Health economics has both theoretical and applied strands, and when concerned with the allocation of scarce resources, it is largely founded on the microeconomic theory of consumer choice and welfare economic theory.

2.3.7.1 Economic theory of consumer choice

Microeconomics consists of a set of theories. These theories aim to provide an understanding of the process by which scarce resources are allocated across alternative uses in the economy. Microeconomics considers 'commodities' (goods and services), 'prices', 'markets', and 'economic agents'. Economic agents are classified as consumers or firms, reflecting a distinction between activities of production and consumption (although this distinction can be blurred in many cases e.g. consumer as a producer). The theory of

consumer choice considers consumers as individuals, with individuals as rational utility maximising consumers.

Consumer choice theory provides an axiomatic approach to individual preferences over the goods they choose, with axioms of rational choice used to represent preferences in an ordering that meets the form of a utility function. Preferences are examined using indifference curves, exchange across goods, and marginal rates of substitution (MRS) between goods. Individuals are assumed to behave in a way that maximises utility subject to budget constraints.

Preferences: When an individual reports that 'A' is preferred to 'B', it is taken to mean that all things considered, they feel better off under position A than under B. Such preferences are assumed to be consistent with the following basic axioms (properties) of consumer choice:

(1) *Completeness*: Under the assumption of preference completeness, when individuals are presented with any choice, they are assumed to be able to fully understand and make up their minds about the desirability of the alternatives (however alike or unlike the alternatives may be). If A and B are any two situations (e.g. goods or services), the individual can always specify exactly one of the following possibilities:

(i) A is preferred to B
(ii) B is preferred to A, or
(iii) A and B are equally attractive

The assumption rules out the possibility that the individual can report both that A is preferred to B and that B is preferred to A. It also leads to the assumption that individuals are not prone to indecision.

(2) *Transitivity*: Under the assumption of transitivity the consumer choice, or preference, is assumed to be internally consistent. That is, if 'A is preferred to B' and 'B is preferred to C', then the individual must also report that 'A is preferred to C'. If this third statement did not hold there would be an inconsistency in preferences (preferences would be intransitive).

(3) *Continuity*: This assumption of continuous preferences dictates that alternatives in a choice set (e.g. different goods or services) are substitutable (tradeable). That is, where one alternative is altered (e.g. a different mix of characteristics) there is a means of compensating for this by a gain in the alternative in the choice set. This assumption requires

that there is unlimited substitutability between commodities (alternatives). This is referred to as 'compensatory decision-making' and it involves the economic notion of the marginal rate of substitution of one economic good for another. The discrete choice experiment assumes that when an individual makes a choice between alternatives, because preferences are continuous, they trade between the attribute levels on offer (described for each option).

(4) *Non-satiation*: This assumption dictates that "more is preferred to less", i.e. that consumption bundle A will be preferred to bundle B if A contains more of at least one good (attribute) and no less of any other. This is easily understandable in the context of a DCE choice set, with common attributes set at different levels, where one option will be preferred to another if at least one of the attributes contains more of a good and there is no difference between any of the other attributes. This assumption establishes a relationship between the quantities of goods (attributes) in a bundle (an alternative) and its place in a preference ordering, where more of each good or attribute it contains the better. The consumer is assumed never to be satiated with goods (attributes).

Other more general technical assumptions of consumer theory are that preferences are reflexive (a weak assumption that each commodity is preferred or indifferent to itself e.g. one apple and one orange is at least as desirable as one apple and one orange), and that preferences are strictly convex (a more technical assumption, related to indifference curve analysis).

Indifference curves: Indifference curve analysis is used in microeconomic theory to reflect preferences, and preference ordering. An indifference curve (or surface) shows a set of consumption bundles among which an individual is indifferent. That is, the consumption bundles on the indifference surface all provide the same level of utility. In Figure 1 the indifference curves U_1 to U_3 each show a combination of X and Y that provide certain levels of utility; with movements in a northeast direction representing movements to a higher level of utility. The slope of the indifference curve provides the marginal rate of substitution between goods (the rate at which the individual is willing to trade X for Y while retaining the same level of utility overall).

Utility Function: A utility function is a way of attaching numbers to the consumer's indifference curve surfaces, where the utility increases as higher or more preferred indifference curves are reached. It reflects the ordering of consumption bundles. Given the assumptions of completeness, transitivity, and continuity of preferences, it is possible to show formally that people are able to rank in order all possible situations

(alternatives) from the least desirable to the most. Using the notion of utility, it also follows that the more desirable situation offers more utility than the less desirable ones. Therefore, if a person prefers A to B, the utility assigned to A exceeds the utility assigned to B. Individual preferences are assumed to be represented by a utility function of the form:

$$U(X_1, X_2, \ldots Xn),$$

where $X_1, X_2, \ldots Xn$ are the quantities of each of n goods that might be consumed. In Figure 1 the indifference curves represent a utility function of a Cobb-Douglas form (Utility $=f[X,Y]$, $= X^{\alpha} Y^{\beta}$, where α and β are positive constants), which is a commonly used utility function in microeconomic theory (Gravelle & Rees, 2004), however other forms are often used.

Budget constraint: The budget constraint reflects the limit on available resources. For example, in a simple two good case for an individual consumer, assume there is an income of Z to allocate between goods X and Y. If P_x is the price of good X and P_y is the price of good Y, then the individual is constrained by:

$$P_xX + P_yY \leq Z$$

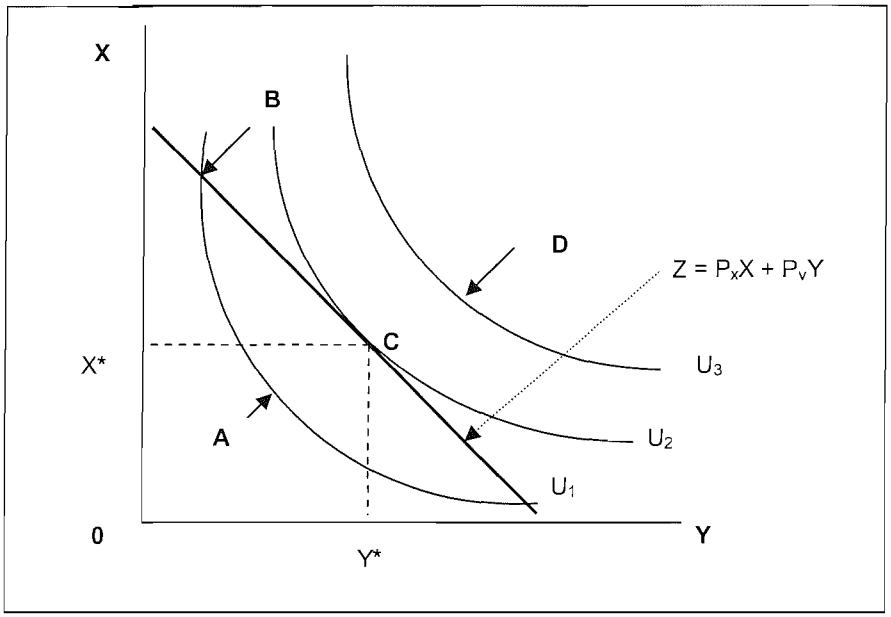
where no more than Z can be spent on the two goods in question. A budget constraint is shown graphically in Figure 1.

Utility maximising behaviour of the individual consumer

As above, the theory of consumer choice represents consumers as individuals, with individuals as rational utility maximising consumers, and consumers as the best judge of their individual preferences. Figure 1 presents a graphical demonstration of the approach of the economic theory of consumer choice, depicting utility maximization subject to a budget constraint.

Point C represents the highest utility level that can be reached by the individual, subject to the budget constraint (Z). The combination of X^* and Y^* is therefore the rational way for the individual to allocate their resources; only for this combination of goods will the two important conditions of consumer theory hold i.e. budget is consumed (assumption of non-satiation), and the MRS = ratio of P_x and P_y (P_x/P_y). In the figure, indifference curves U_1 to U_3, give greater levels of utility from combinations of X and Y; $U_3>U_2>U_1$. In terms of the utilitly maximizing combination of X and Y, points A and B can be improved upon, by choosing C, and point D is not achievable given the budget constraint Z.

Figure 1. A Graphical Representation of Consumer Theory of Individual Level Utility Maximisation.



2.3.7.2 Welfare economics

Whereas the consumer choice theory above shows how an individual may choose between alternative bundles of goods, in a utility maximizing manner, welfare economics involves comparisons of welfare (utility) across different individuals. Welfare economics is a framework of assumptions and normative propositions, and is that part of economics which studies the possible effects of various policies on the welfare (or utility) of society.

Welfare economics retains an analytical framework at an individual person level, as it states that where one thing is judged to be better than another this judgment is based solely on the welfare of individuals, with group welfare defined in terms of the sum of total welfare attained by all individuals within a group. Whenever one situation (state of the world) is said to be better than another this assessment of alternative situations must be based on a certain set of value judgments. Economists have applied the tools of welfare economics, in both theoretical and applied contexts, to consider the issues related to, and the evaluation of, such value judgments.

Welfare economics introduces value judgments into economic analysis using the concept of the social welfare function (SWF), (Bergson, 1938). Bergson (1938) introduced the concept of the SWF in a strict sense, with a relation between social welfare and its determinants in the form of a dependent variable, i.e. social welfare, and a number of independent variables

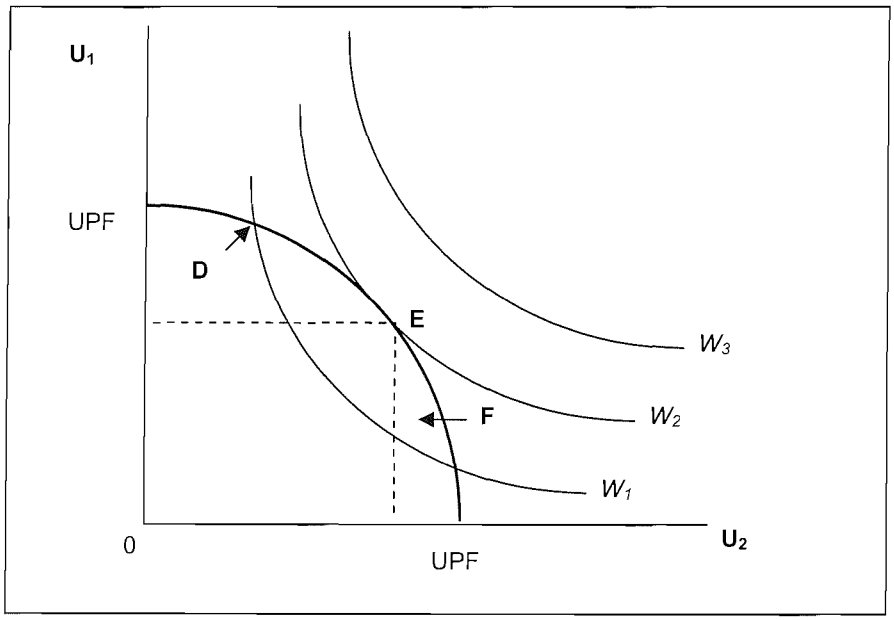
which determine social welfare, with the relation taken to be a well-behaved, continuous and differentiable function, which in principle could be given a well-defined form. But the well defined SWF is not the only form of the SWF. Nath (1973) has highlighted that for practical purposes, a more mundane view of the social welfare function may be taken, where the concept may sometimes be taken to simply imply any general statement of the objectives of a society, with some "rough and ready idea of the relative weights of these objectives" (Nath 1973, p25). With respect to the assessment of alternative arrangements for health care, both of these approaches to the SWF can be taken. It is via the SWF that economists have introduced aspects of welfare economics to the study of different allocations of health care resources, in both a strict theoretical sense, using a defined form of the SWF, and in a more general application of the SWF with some general statements surrounding potential objectives of a society with respect to health. The most important aspect of the SWF is that it allows value judgments to be introduced into economic analysis in a systematic and objective way, and it allows, in principle, an objective analysis of implications of different sets of value judgments.

However, the SWF has been almost entirely used in a theoretical context, and the examination of the SWF has been via an analytical approach concerned with the consideration of production possibilities and the choices available on the production possibilities frontier, with the SWF offering a hypothetical means of identifying a preferred distribution of goods (often in a hypothetical two-person world). Such contributions have largely been against the Pareto criterion for assessment of social welfare changes, and Pareto optimality. A Pareto improvement is defined as a situation where it is possible to make one person better-off without making some other person worse-off. Theoretical welfare economics derives the necessary conditions for the achievement of a 'Pareto-optimum' (Ng, 1979). A Pareto optimal allocation of resources is said to be technically and allocatively efficient. Technical efficiency is achieved through an allocation of resources to minimise the inputs required to produce a given output (efficiency in production); and allocative efficiency is achieved when resources are allocated in a way that prices of each of the outputs are proportional to the utilities consumers derive from them (efficiency in consumption). Paretian welfare economics allows consideration, in theory, of the efficient allocation of goods and services between individuals (e.g. use of Edgeworth box),[4] with Pareto optimal allocations (efficiency in consumption and efficiency in production) plotted on a utility possibility frontier.

[4] It is not possible here to outline all of the elements of microeconomic theory of exchange, and readers are encouraged to consult an entry level text on microeconomics such as that of H.Gravelle & R.Rees 2004.

Figure 2 presents a simple graphical representation of a SWF. In this figure the possible welfare distributions in a two-person world (u_1,u_2), are shown by the utility possibility frontier (UPF). The UPF shows the points that are possible and efficient according to the Pareto criterion. However, the use of the SWF shows that some points on this frontier are more socially desirable than others. Point E on the SWF W_2 is presented here as the optimal point of social welfare, i.e. the highest level of utility from the given UPF. However point D represents an efficient point on the UPF, but an inefficient point like F may be more socially preferred. It is such scenarios that raise the proposition of a social preference to forego efficiency gains in order to deliver against conflicting distributional value judgments over the distribution of resources in society.

Figure 2. A Graphical Representation of the Conceptual Value of the Social Welfare Function.

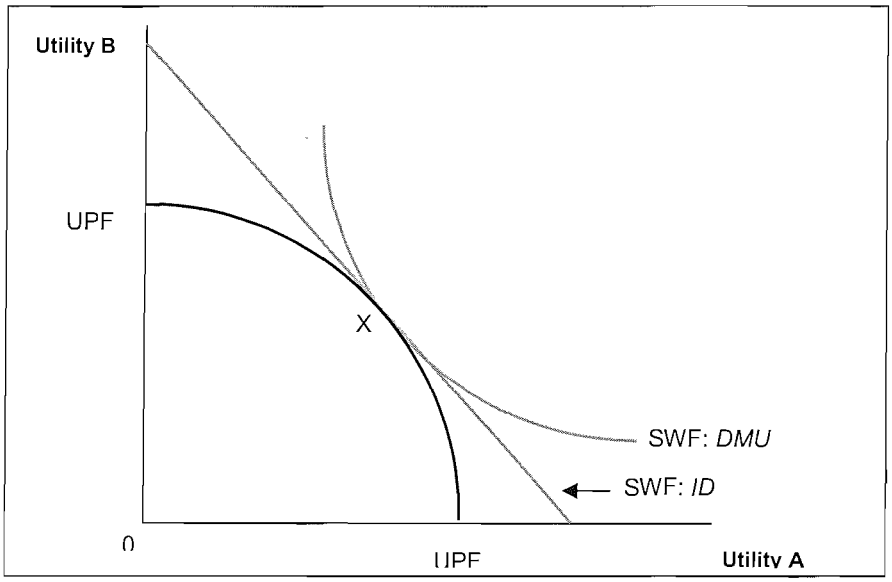


The SWF does allow in principle, albeit at a conceptual level, economic analysis to move beyond efficiency concerns and to consider social (distributive) preferences. This is achieved through the use of indifference curve analysis, characterising specific social and distributive preferences, that indicate how much utility from one individual (or group) society would be willing to trade-off in order to increase the utility for another individual (or group of individuals).

Where social preferences are indifferent over the allocation of utility, or welfare, across individuals or groups, a SWF like *ID* shown in Figure 3 will reflect a view that one unit of utility to one group is traded for an equal unit of utility for an alternative group (utility valued

equally across groups, i.e. no preference). A SWF like *DMU* shown in Figure 3 will reflect a preference to allocate differently across groups, with a decreasing marginal utility across groups represented by the curve *DMU*. However, these conceptual and analytical examples do not provide an indication of how society may wish to set priorities across different individuals or groups (i.e. no indication of specific social values).

Figure 3. A Graphical Representation of the Different Approaches to the Social Welfare Function.



The Pareto criterion does not address the analysts concern for value judgments in a practical way. The Pareto criterion is a strict specification of a weak value judgment i.e. it ensures that there are no losers. However, it is a theoretical ideal in most policy making problems and does not offer practical assistance (Reinhardt 1992, 1998). To overcome the limitations of the Pareto improvement, the 'potential Pareto improvement' was developed (Kaldor 1939, Hicks 1939). A potential Pareto improvement (PPI) is defined as a situation where an action provides an improvement if its benefits are large enough that gainers can (albeit hypothetically) compensate the losers. In such a case it assumes that the losers are (hypothetically) no worse off than before (the action) and the gainers are better off, i.e. they have a net benefit after considering the hypothetical compensation for the losers. The PPI criterion has been much criticised for its implicit distribution arrangements and hypothetical transfers (e.g. Reinhardt,1992). However, the PPI and the compensating variation associated with hypothetical compensation, are used to underpin cost-benefit-analysis as an applied element of welfare economics. Cost-benefit analysis weighs up the total expected costs and total expected monetary benefits of one or more actions in order to choose the better action, via an assessment of net benefit. It compares actions (states of affairs) with

and without the action (or project) that is being analysed. It is applied welfare economics, and comprises almost all that is of interest in applying welfare economics. However, distributional judgments are required to the effect that the benefits outweigh the losses in terms of welfare, and in many instances CBA (and other forms of economic evaluation) departs from the strict principles embodied within welfare economics (McGuire, 2001).

2.3.8 Economic evaluation

In moving from theoretical to applied welfare economics it is necessary to move from analysis of the necessary conditions for optimum states of affairs, and how such optima may be achieved, to problems that involve deciding whether one actual outcome (state of affairs) is better or worse than another. As above, cost-benefit-analysis (CBA) is a technique for applied welfare economic analysis. The theoretical basis for CBA relies directly on the notion of the PPI, and the compensating variation derived in welfare economics (McGuire, 2001). It is a formal discipline used to appraise, or assess, the case for a project or proposed action. CBA involves monetary calculations of expected costs and expected benefits associated with an action. Sugden and Williams (1979) provide a detailed description of CBA. It has mainly been used to assess the value for money of large public and private sector projects, although there has also been wide use of CBA in the areas of transport and environmental projects. However, CBA has not been widely used in the consideration of the acceptability of health care technologies (McIntosh, 2006), due to both practical and conceptual difficulties around valuing the benefits from health technologies. Other forms of economic evaluation have been used in the assessment of health care technologies. These are the techniques of cost-effectiveness analysis (CEA). CEA does not use monetary measure of benefits, and avoids the problems related to estimating the monetary value of health benefits. CEA provides a comparison between alternative options (e.g. health technologies), and results are presented as an incremental cost-effectiveness ratio.

CEA takes a more limited, or applied, view of the benefits of health technologies than does the CBA framework. It measures benefits in terms of natural units (e.g. units of effect, or life years and life-expectancy), or using summary measures of health gain. Where a single index summary measure of health is used to reflect the benefits from health care, i.e. overall health-related quality-of-life, the form of CEA undertaken is referred to as cost-utility analysis (CUA). In its most common form CUA adopts the quality-adjusted life-year (QALY) as a summary measure of health (QALYs are introduced below). CEA and CUA present an estimate of the incremental cost per additional unit of benefit, for CUA this is the estimated cost per QALY.

The practical use of economic evaluation in health care has almost entirely involved the production of summary statistics (e.g. cost per event avoided, cost per life-year saved, cost per QALY gained), focused on informing an assumed objective function of health maximisation (Olsen 1997, Oliver *et al* 2004). This objective function is not directly linked to the welfare economic approach of maximising utility (unless these units of measurement are adequate proxies for utility and well-being), and Sugden and Williams (1979) have described one possible approach for economic evaluation as one which adopts a decision-maker perspective, with the decision-maker applying value judgments deemed appropriate for the decision problem and the community involved in the decision problem. The decision-maker perspective moves economic analysis away from the theory of welfare economics, and the compensating variation derived from the potential Pareto improvement (PPI), (as it may involve replacing the welfarists sole use of individual utilities, and the maximisiation of utility). Sugden and Williams (1979) advocate that a decision-maker perspective use the 'decision-makers objectives', which should be clear, and which will no doubt be context-specific; noting that *"the words 'cost' and 'benefit' [will] lose any meaning that is independent of [the decision-maker]"* (p237).

This decision-maker perspective, or viewpoint, has not been clearly explored in the context of health care, but it has been accepted as a prominent approach in the conduct of economic evaluation, and related health care decision-making (Olsen 1997, Hauck *et al* 2003, Smith *et al* 2005, Oliver *et al* 2004, Rawlins & Culyer 2004). The maximisation of some partial or summary measure of clinical-effectiveness or health is the basis for almost all economic evaluations (Oliver, 2004), and the economic evaluation of health technologies has not used the maximand of social welfare or utility. The literature on economic evaluation for health care indicates that differences in the distribution of benefits, or health gains, across patient groups has been ignored, and economic evaluation has been used to address efficiency concerns, presenting findings against the incremental cost per incremental gain in outcome (Olsen 2000, Sassi *et al* 2001). As such the literature on economic evaluation says very little about the social value of health, other than covering the issue of efficiency (Sassi *et al,* 2001).

2.3.9 Quality-adjusted life-years

The quality-adjusted life year (QALY) is a measure of health that combines both quality-of-life (morbidity) and length of life (mortality). QALYs are calculated by multiplying life-years by a numeric value that reflects the level of quality-of-life during those life years. A year in full health counts as one QALY. Where health problems are experienced during a year it will

count as less than one QALY. A quality-of-life weight of one corresponds to full health and zero corresponds to a health state judged equivalent to death; weights below zero are regarded as worse than death. Where pain and mobility problems, for example, reduce the weight attached to quality-of-life from 1.0 to 0.5, then each year in that health state counts as half of one QALY (0.5 QALY).

QALYs are widely used in health economic evaluation, and are used by NICE in the UK as the preferred summary measure of health-related quality-of-life (NICE, 2004). To operationalise the QALY approach it is necessary to elicit weights for the quality of life of health states. Methods to elicit health state valuations vary, but commonly used techniques are the standard-gamble, time trade-off, and visual analogue techniques. For further information on health state valuation techniques see an earlier systematic review of the literature on health state valuation techniques by Green *et al* (2000).

QALYs are not the only available summary measure of health, for example there is also the disability-adjusted life year (DALY), (Murray & Lopez, 1997), and the healthy-years equivalent (HYE) (Weinstein & Pliskin, 2006), but the QALY approach remains the most commonly used and discussed summary measure of health in health economics. All of these measures are regarded here as a measure of health output, and the points raised in the current research with respect to the maximisation of health gain (QALY gain), are valid against all of these summary measures of health production.

2.3.10 QALY maximisation

The traditional approach in economics, and economic evaluation, is to maximise some objective function, for example the benefits from an action or programme, subject to relevant constraints (e.g. budget constraints). Therefore the basic problem becomes:

$$\text{Maximise:} \quad \sum_{i=1}^{M} n_i \times Bi$$

$$\text{subject to:} \quad \sum_{i=1}^{M} n_i \times C_i \leq \overline{X}$$

where i is an intervention, n_i is the number of people involved in the intervention i, B_i the average benefit per person from intervention i, and M is the total number of interventions considered. C_i are the costs associated with i, and $\overline{X}$ the total budget available (the constraint). Here actions (interventions) are arranged, and the priority setting problem can

be formulated, as a linear programming problem, with actions sorted according to their ratio of costs to effects (actions are implemented starting with the one with the lowest ratio, until budgets are used up).

Where the maximand is health, and the QALY is adopted as the relevant measure of health, as it is in NICE in the UK, the objective function becomes the maximisation of QALYs subject to constraints. This can be written alternatively as:

$$\text{Maximise: } \sum_{i=1}^{M} n_i \times QALY_i; \qquad \text{subject to budget constraint}$$

and, in a similar linear programming fashion, interventions can be sorted according to their ratio of costs to effects. This latter issue reflects the possibility of using a cost per QALY threshold to establish which interventions may be regarded as acceptable for health policy in the NHS by NICE. NICE do not have an explicit cost per QALY threshold (e.g. £25,000 per QALY), for use in their technology appraisal programme (NICE, 2005), but they do offer guidelines on what level of cost per QALY is generally regarded as being a cost-effective use of NHS resources (Rawlins & Culyer 2004, NICE, 2005).

2.3.11 Equity versus efficiency (trade-off)

The equity versus efficiency trade-off has been used as a practical and analytical concept to discuss the notion of opportunity costs, when making choices between competing demands for a limited budget.

Efficiency is generally regarded as an uncontentious objective of any health care system, but especially in a publicly funded health care system, such as the UK NHS, where public funds need to be used wisely and efficiently. Achieving an efficient allocation of resources is about comparing the costs (or resources) and the benefits (e.g. health gain) of alternative uses of resources, and ensuring that resources are allocated to maximise the gains to society. In the UK, for decisions over the merits of health technologies, the QALY is commonly referred to as the preferred measure of health output (e.g. NICE, 2004), and the efficiency objective is often framed as maximising QALYs gained, from available resources.

Achieving efficiency involves both technical (or operational) efficiency, and allocative efficiency; these being aligned to efficiency in production and efficiency in consumption (introduced in section 2.3.7). Achieving the most efficient use of resources, in the context of social objectives, is concerned with minimising the 'opportunity costs' (or 'sacrifices'), when

making choices between competing claims on limited resources. These opportunity costs, lead to the conceptual presentation of the 'equity versus efficiency' trade-off. Whereby, the opportunity costs may be efficiency gains foregone, in order to achieve a more equitable distribution of health care, or they may be a less equitable distribution of health care in order to achieve a more efficient use of resources.

Okun (1975) has been credited (e.g. Osberg, 1995) with popularising the idea of a "great trade off" between equity and efficiency. Whilst, equity and efficiency are rarely defined in a clear and concise manner, the concept of the trade off between these objectives is able to deliver an important message that there are gainers and losers when making policy decisions subject to resource constraints. Wagstaff (1991) was one of the earlier analytical contributors on the equity versus efficiency trade-off in the health economics literature, with a theoretical presentation of a SWF to depict a trade-off between efficiency and the reduction of health inequalities (applying an inequality aversion parameter).

Whilst efficiency may be an important and commonly used objective in health care systems, such as the UK NHS, it is not regarded as the only important objective (discussed further in Chapters 3 and 4). Other important objectives surround the distribution of health care resources across different individuals and groups in society. This is often referred to as a fair or equitable allocation of resources (e.g. Sassi *et al*, 2001). However, there is little guidance on what society might regard as a basis for judging what is 'fair' or 'equitable'; and exploring this issue is one of the motivations for the current thesis.

Social preferences over the distribution of health care, also referred to as distributive preferences, are frequently summarised as 'equity' considerations, but without any accompanying detail on the specific issues that are thought to define what is meant by equity. Therefore, equity is commonly presented as a very broad issue; in the health care literature there is broad discussion of 'conceptions of equity', 'equity principles', 'notions of equity' (e.g. Culyer 2001, Donaldson & Gerard 2005). There is rarely any operational detail around equity to inform specific resource allocation decisions (Sassi *et al,* 2001). Equity objectives are not stated in a clear and explicit manner in the UK NHS (Charny *et al* 1989, Newdick 2005), and importantly, given the potential to have a range of equity objectives, there is no indication in the health care literature of the relative importance of competing equity objectives i.e. the trade-offs between different equity considerations (e.g. Dolan *et al,* 2005); (again, this issue is one of the motivations for the current thesis).

The literature around the equity versus efficiency trade-offs in health care is largely analytical (e.g. Wagstaff 1991, Dolan 1998), however, there have been some empirical contributions against various equity considerations (e.g. Sassi *et al* 2001, Donaldson & Gerard 2005). Such equity considerations have not been aligned directly to empirically determined distributive preferences. In these empirical studies (and in the majority of the health care literature) more general conceptions of equity, such as equal access for equal need to consider the availability and utilisation of specific services, have been used to consider broad conceptions of equity.

Equity is frequently aligned with some notion of equality. However, equity and equality are two different, although often intertwined, concepts of distribution, and they are not always good complements ('bedfellows') or interchangeable. Equality is the state of being equal, across some determined dimension, or measurable variable. Equity is used to consider a fair, or equitable, distribution of a good (or resources/outcomes), and that may not entail equality. Although equality, across various dimensions of health and health care delivery, has been used widely to introduce, and explore, the notion of equity in health.

Equality and equity can be characterised using the definitions of horizontal and vertical equity, the former putting greater focus on equality than the latter:

> *Horizontal equity*: is concerned with the equal treatment of equals
> *Vertical equity*: is concerned with the unequal treatment of unequals

These definitions of equity have been given wide coverage in the literature around health economics and health policy, however putting either of the terms in an operational (practical) resource allocation problem requires context specific information, value judgments, and an interpretation of what it is that is 'equal' or 'unequal'. Any such operational use will require information on what social values are important, and how society might wish to trade-off competing social values, against each other (including efficiency as a social value).

In terms of horizontal equity, commonly cited objectives have been equality of access, equality of expenditure, equality of utilisation. However, all of these have to be linked to some definition of 'need'. For vertical equity it is also necessary to address the definition of need, for example in pursuing an objective of unequal treatment according to 'need'. There is a broad literature on the definition of 'need', and the use of 'need', to consider equality objectives in health care (e.g. Culyer & Wagstaff 1992, Hauck *et al* 2003, Donaldson & Gerard 2005). For example, equal access according to need, and the consideration of the

'capacity to benefit' as a measure of a need for health care. In the present outline, it is sufficient to highlight this issue, without discussing it in any detail, as the presence of some notion of need, to seek an understanding over equality objectives, in addressing equitable distributions of resources, takes us in the same direction as that which dictates the importance of exploring social values.

Within the equity and equality literature, the growth in the related research on health inequalities (e.g. Department of Health 1980, Acheson Report 1998, and the ESRC work programme on health inequalities), has made the reduction of health inequalities a prominent issue for health policy, at both a national level (e.g. UK Department of Health 2003) and an international level (e.g. WHO *World Health Report* 2000, New Zealand Ministry of Health 2000). Recent research on equity has framed it as the "absence of inequity" (Starfield 2001, p324), or the reduction of inequalities, however such broad definitions of equity still leave a vacuum at a practical decision-making level, and the articulation of social values around a varied range of health inequalities is still needed.

It is argued here (Chapter 3) that the above broad, and primarily conceptual, presentations of what might be meant by equity in the allocation of health care (resources) are not operationally useful social values, or distributive preferences; and that they are unable to inform specific decisions (such as health technology appraisal). Such broad conceptions of equity may be more useful when viewed as 'higher-order', system wide policy objectives, or principles.

In summary, the equity versus efficiency trade-off is a useful analytical tool. The economists presentation of the SWF is an example of how trade-offs can be presented in a way to raise awareness of opportunity costs, and sacrifices, when resources are constrained. Whilst resources for the delivery of health care are constrained there will always be competing demands for the same resources, and there will always be a need to weigh up gains and losses against equity and efficiency criteria. The use of distributive preferences, also referred to as social values, to inform such judgments, will provide an explicit basis for the distribution of health care across different groups, and consideration of the opportunity costs associated with different choices. Addressing equity considerations in a more operational, and explicit manner, through the elicitation (and application) of social values, will place decision-makers, and society more broadly, in a position to assess the appropriateness and acceptability of the consequences of health care resource allocation decisions.

2.3.12 Social values (terminology)

The central element of this thesis is the 'social values' over alternative uses of limited health care resources. Social values have been introduced in various forms above, when discussing the key concepts used as a framework for the thesis. It is helpful to conclude this section by highlighting the variation in terminology over such social values.

Many commentators have referred to social values and their role in health care decision-making (see Chapter 4). However, terminology varies, and social values are presented as distributive preferences (e.g. Culyer, 2001), distributive principles (e.g. Wilmot *et al,* 2004), social preferences (e.g. Bleichrodt *et al,* 2005), equity arguments (Nord *et al,* 1995), decision-making criteria, preferences over alternative distributions of health care (e.g. Sassi *et al,* 2001), public preferences (e.g. Dolan & Tsuchiya, 2005), societal values (e.g. Schwappach, 2002), impersonal preferences (Schwappach & Strasmann, 2006) as well as other variations around these terms. Therefore, in this thesis there is a variety of language used to cover the discussion of social values.

In this thesis the simple definition of social values, underlying the material presented, is as follows.

Firstly, value judgments are assumed to offer a view of what is desirable (or undesirable), often without technical or objective data, but based on considerations of overall value (for example, of value to society). As such, social values as normative propositions, are not treated here as something that can be true or false, they are considered here as being either persuasive or otherwise (Ng, 1979).

Secondly, social values are 'social' (as opposed to individual values) as they are concerned with the allocation of resources for others in society (across different groups in society), and not the allocation of resources against alternative actions for the individual offering the social values.

Finally, drawing on definitions of social value judgments in health care (Culyer 2005, NICE 2005), a social value is defined as a normative proposition, or ethical opinion, that a particular course of action, institutional arrangement, or method of analysis ought to be implemented, or is itself 'good'. As such, social values, in the context of distributional preferences over health care may say it is a good thing to do 'X', a good thing to do 'Y'; so they are statements (social values) over what is regarded as 'good' for society, in the context of the decision problem. For example, 'it is good to use resources efficiently'.

This thesis is concerned with identifying key social values that are concerned with doing good (as above) when allocating health care resources, and how they may be weighed against one another in the context of the appraisal of health technologies in a resource constrained system of funding health care. This is the basis for exploring the social value of health interventions.

The consideration of social values in this thesis draws inspiration from the work of Broome (1991), who considered the structure of 'good', and the fact that there are many different sources of doing good. Broome argued that it is necessary to 'weigh' different goods in order to consider what action may be better than others. Broome discusses this need to 'weigh' goods against one another and refers to it as a notion of 'betterness' i.e. some things being better than others.

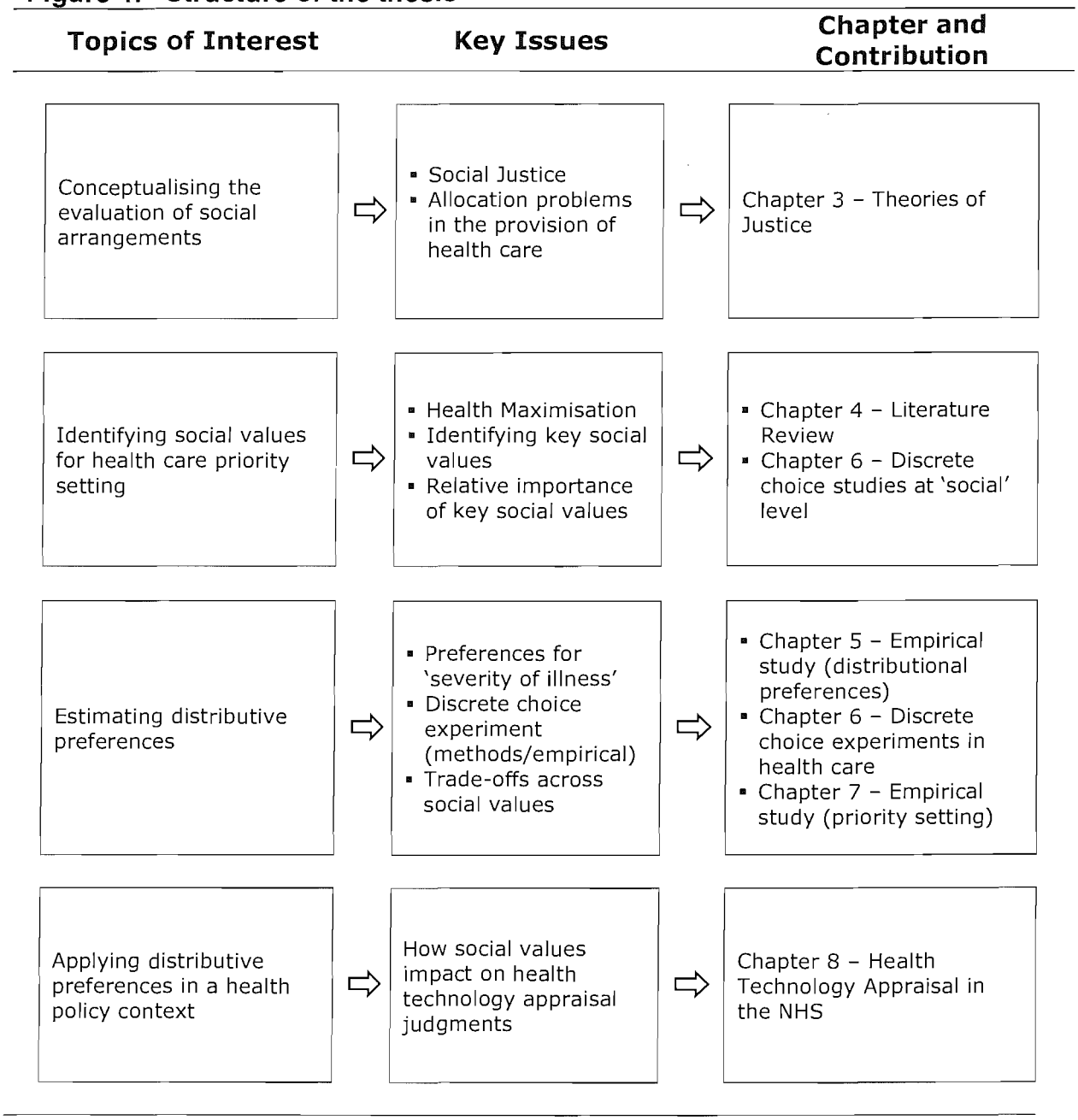
2.4 Structure of the thesis

Overall the thesis attempts to strike a balance between contributing to the debate in health policy over resource allocation problems and to methodological debates in health economics on priority setting and social values (equity values). Figure 4 shows the basic structure of the thesis, the main issues addressed and the contributions made.

The thesis offers contributions to both health economics and health policy analyses. These are set out as follows. Firstly the thesis systematically considers a range of pertinent theories of justice and fairness as general theories (i.e. not health-related) and goes on to interpret them for health technology appraisal. As a result, specific key messages are drawn out for the purposes of informing health technology appraisal. Secondly, the thesis comprehensively draws together, assesses the quality of, and identifies key gaps in, the current empirical literature on social values for priority setting choices in health care. The third and fourth contributions are empirical in nature. The first empirical study pursues, in detail, the idea of a social concern for the criteria 'severity of illness', in the distribution of health care resources. (i.e. a preference in favour of severely affected persons, *vis-a-vis* less severely affected persons). The study investigates the hypothesis that there may be a general preference to favour 'worst-off groups' in resource allocation problems, with severity indicating worst-off (or a general preference for fairness). The second empirical study applies a discrete choice experiment approach to investigate the trade-offs (relative values) between alternative social values, using a sample of the general population. This study adds further evidence, to a currently limited one, of using the discrete choice methodology

for social choices. Finally, the thesis contributes to furthering health economics knowledge by using the results from the discrete choice study to examine the consistency of the NICE health technology appraisal programme with the initial findings on public preferences across the key social values identified.

Figure 4: Structure of the thesis



The discussion, in Chapter 9, brings together the various methods and findings across the core elements of the thesis (theories of justice, empirical literature review, key social values, relative importance of the social values, and their application to health policy), to explore

what the key findings offer in terms of helping health care analysts and decision makers with a more explicit and operational approach to the consideration of social values as distributive preferences, when seeking a position of equity (i.e. a fair allocation of resources).

2.5 Summary of Introduction

This introductory chapter has set out the need for the research presented in this thesis, the scope of the research, and the structure of the thesis. The thesis explores the social value of health care interventions, using a health economic framework, and in the context of health technology appraisal. A working framework for social values, health economic theory and health technology appraisal has been provided.

The important need for research in this area, and the motivations for the research have been outlined. The introduction to the economic theory around welfare economics, and the analytical approach to the assessment of arguments around a potential trade-off between equity versus efficiency, and the related opportunity costs, sets a clear context for the research drawn together in the thesis. The following chapters, covering theories of justice, and the empirical literature on social values in health care, further set the context for the research and the additional empirical work in the later chapters of the thesis.

3 THEORIES OF JUSTICE: A SURVEY OF COMMON THEORIES OF JUSTICE, AND HOW THEY MAY APPLY TO HEALTH CARE

3.1 Introduction

This chapter presents a discussion around justice and fairness, through an examination of common general theories of justice, and their application to health technology appraisal.

Justice is a core moral and political term which is generally associated with the promotion of just human relationships as part of a well-ordered society. There are many interpretations of what it is society might mean by justice. It is possible to draw a distinction between corrective or remedial justice on the one hand and social and distributive justice on the other, the former having to do with the law and the latter with social policy, particularly taxation and welfare (Campbell, 2000). It is also possible to draw a distinction between procedural justice and distributive justice. Procedural justice is linked to the concept of justice that has traditionally been concerned with rights and duties: with giving a person their due, and not infringing their rights (other than for moral reasons), (Little, 2002). Distributive justice is more aligned to the assessment of the distributive aspects of the basic structure of society (Rawls, 1971).

This thesis is concerned with the social and distributive conception of justice, in the context of health care. Social and distributive justice, tend to bring together the notions of justice and fairness (Rawls, 1971, Broome, 1991). The consideration of justice and fairness here acts as a foundation to the examination of social value judgments and how they may characterise equity positions in health care, specifically in the area of health policy and the appraisal of health technologies.

Theories of justice and fairness are considered here in order:

- To provide a greater understanding of the evaluation of social arrangements.
- To establish the importance of identifying social value judgments as decision-making criteria.
- To explore the influence of theories of justice on the appraisal of health technologies, via a thought experiment.

To further these aims the most commonly cited theories of justice are discussed below.

3.2 Theories of justice

The theories of justice discussed here are from the literature on moral philosophy, and they are concerned with moral judgments over social alternatives. Economic evaluation is rarely value-free, and it is likely that economists will have to rely on some moral philosophy (Hausman & McPherson, 1993).

The following theories of justice are discussed:

- Utilitarianism
- Rawls' Theory of Justice (*Contractualism*)
- Nozick's Entitlement Theory (*Libertarianism*)
- Egalitarianism
- Extra-Welfarism

Firstly the theories of justice are presented in a general context, as they are not directly related to health. Thereafter, the theories are considered against their potential practical and operational use in health care policy decisions, using the process of health technology appraisal within the NHS to explore such implications. The discussion of these general theories of justice, and the common factors within them, provide a foundation for consideration of allocation problems in health care.

These theories of justice have been selected, following guidance from experts in this area (Professor AP Hamlin, e.g. Hamlin 1996), to represent a range of alternative frameworks for the evaluation of social welfare.

Through an examination of these broader general theories of justice it is possible to arrive at some key messages on the complex nature of the evaluation of social arrangements. These messages are presented to inform on health policy-making and the process of health technology appraisal in the UK NHS, and are used as a basis for presenting a proposal for a stylized framework for health technology appraisal.

This stylized framework sets out a hierarchy of requirements comprising underlying principles, procedural arrangements, and operational objectives reflecting decision-making criteria. The former of these requirements (i.e. principles and process) are the focus for much of the literature on theories of justice, however the latter requirements of stated objectives and decision-making criteria are largely neglected. This thesis examines these more neglected requirements, i.e. the social value judgments that may act as objectives

and/or criteria when decision makers are faced with allocation problems in health care (e.g. health technology appraisal). This chapter is used to establish the importance of identifying such social value judgments.

3.2.1 Utilitarianism

In its modern form utilitarianism is most commonly associated with the moral standard of maximising utility (regardless of the definition of utility), yet it is not a single doctrine (i.e. a maximising one) but a family of doctrines defined by key characteristics (Bailey, 1997). Utilitarianism, in whatever form (various presented below) encompasses four key elements:

(1) Evaluative consequentialism: utilitarianism is a consequentialist approach, where an assessment of utility or welfare is based on consequences. Consequentialism claims that we ought to do whatever maximises good consequences (the amount of good in the world), but consequentialism alone does not determine what the 'good' is.

(2) A theory of the personal good: utilitarianism involves a theory of the personal good (i.e. general happiness), where goodness (and badness) is evaluated in terms of individuals' utilities (or welfare). Consequences need to be considered in the context of personal good, i.e. how they are valuable to individuals.

(3) Interpersonal comparability: the notion of interpersonal comparability holds that we can make comparative judgments about states of the world in which some persons are made better off and others made worse off.

(4) Distributive indifference over the good: utilitarianism holds a principle of distributive neutrality, whereby increases in personal good (utility) make states superior to the same degree without regard for the person in which those increases take place (it is the size and not the location of gain that matters).

As a consequentialist theory, utilitarianism has little to say about the multiplicity of rights, or the make-up of persons outside of the definition of whatever it is that is being maximised. It is also blind to circumstances, and causes, when it comes to persons being in a particular state. With regard to the information base applied by utilitarianism, it is limited to information on personal utilities, or preferences, and does not allow the consideration of non-utility information. Under the utilitarian principle of distributive indifference it is possible to have distributions which are regarded by some non-utilitarians as substantively unjust, in that they permit the sacrificing of the interests of some individuals to promote the well-being of others,

if the latter gain more than the former lose. The principle of distributive indifference would see a relevant reason for tolerating inequalities as a gain in utility, (i.e. a gain through maximising utility). Utilitarianism would argue that society should be prepared to tolerate the fact that some individuals do less well than others as long as the aggregated personal good is greater.

These characteristics highlight the fact that utilitarianism is indeed a 'reductionist' theory, it uses the principle of utility to capture the consequences of an action and the 'good' (or bad) that results from that action. However, the classical utilitarian interpretation of the term utility, as 'happiness' or 'goodness', offers a much broader measure of outcome (well-being) than that captured using more recent references to utility according to 'preferences' (Broome, 1991).

Jeremy Bentham introduced 'utilitarianism' as a moral theory (Bentham, 1789). He presented utilitarianism as a moral theory according to which an action is right if and only if it conforms to the principle of utility. The principle of utility was set out as that which approves of an action in so far as the action has an overall tendency to promote the greatest amount of happiness; with happiness related to the presence of pleasure and the absence of pain. Bentham claimed utilitarianism should be used to consider alternative courses of action, to determine which action had the best tendency to promote happiness, and therefore should be chosen. Bentham's work was followed by John Stuart Mill, whose *Utilitariansim* (1863) moved Bentham's work forward. Mill highlighted the difference between what economics measured and what human beings really 'valued', arguing that society ought to aim at maximising the welfare of all, and that utilitarian views were consistent with fairness and a theory of justice.

Mill presents the normative principle of utilitarianism as:

> Actions are right in proportion as they tend to promote happiness; wrong as they tend to produce the reverse of happiness

Mill refers to this as the principle of utility. Mill argued that the principle of utility involves an assessment of only an action's consequences, and not the motives or character traits of the agent performing the action (see below under 'evaluative consequentialism').

Between them Bentham and JS Mill are generally judged to be the source for what is recognised as the classical moral theory of utilitarianism. Others have contributed to the

theory of utilitarianism, notably Sidgwick (*Methods of Ethics*, 1874), but classical utilitarianism remains largely as described by Bentham and JS Mill. However, the utilitarianism referred to in most instances (within the recent economics literature) has become something other than the classical utilitarianism of Bentham and Mill. Furthermore, the terminology of happiness and pleasure introduced by Bentham and JS Mill has been replaced by a different conception of utility.

Utilitarianism has been used widely 'in name' but many different variants have been conceived, some of the more formal ones listed below. Regardless of format the framework for utilitarianism embodies the four key characteristics above, all of which are required for a utilitarian approach.

The various different forms of utilitarianism, commonly referred to in the literature, are:

- Act utilitarianism, where each individual action is to be evaluated directly in terms of the utility principle.
- Rule utilitarianism, where behaviour is evaluated by rules that, if universally followed would lead to the greatest good for the greatest number.
- Ethical utilitarianism, which is concerned with social choices and not individual preferences, and where the best action is the one which gives as many people as possible what they choose.
- Democratic utilitarianism, applying a democratic interpretation of preference utilitarianism, by making social and political choices in accordance with the expressed will of the majority.
- Preference utilitarianism, which interprets utility as giving individuals what it is they choose (preference satisfaction), rather than considering the pleasures and pains they experience.

Utilitarianism is linked to a consequentialist framework, however it is much more than consequentialism when defined according to the combination of necessary characteristics stated above (including the theory of personal good). Often where utilitarianism comes under attack as a moral theory, it is not utilitarianism as defined by its key characteristics, but theories described as utilitarian but not utilitarian in accordance with its required component parts. For example, where critics argue against utilitarianism on the grounds of its use of

consequences, and the maximisation of consequences without any regard for distribution, they argue against consequentialism *per se*, and fail to see that utilitarianism is (or should be) based on consequentialism in combination with an adequate theory of the personal good, a basis for interpersonal comparability and a stance of distributive neutrality.

In such a way, discussion of utilitarianism is often accompanied by ambiguities and confusion. Much of the confusion, or debate, relates to the concept of personal good i.e. what is it that is meant by happiness, or good, or utility, or preference satisfaction or welfare? Over time, the object of utilitarianism has moved from happiness maximising, using 'pleasure and pain', to 'want-satisfaction', to the maximising of economic preferences. The current and most common form of utilitarianism is 'preference utilitarianism' (Broome 1991, Campbell 2001). But again there is a lack of clarity on what it is that is meant by preference and preference satisfaction (especially in the discussion of health and health care, where revealed preference data is rare). For example, preference can be a stated preference or a revealed preference, and the basis upon which preferences are determined may be unknown. Furthermore, preferences, and preference satisfaction, are questioned as an adequate conception of individual well-being (e.g. Hausman & McPherson 1993, Sen 1987).

There has been much debate over the use of utilitarianism as a moral theory, and a number of alternative theories of justice have arisen (at least in part) out of a general dissatisfaction with utilitarianism (e.g. Rawlsian theories, Nozick's entitlement theory, and Sen's capabilities approach). These alternative theories are discussed later in this chapter, offering a further insight to the utilitarian approach, but the main objections to utilitarianism are based on (a) its use of the principle of distributive indifference (whereby utilitarianism maximises some measure of good, regardless of how it is distributed), (b) a belief that utilitarianism disregards individual rights, and (c) a belief that the utilitarian information base for evaluation is too thin. Utilitarianism has a response to these criticisms (see Bailey, 1997). But, as discussed above, a broader issue, and one that may be linked to all of these main criticisms, is the concept of utility and the different meanings that are attached to the term 'utility' (e.g. Mirlees 1982, Sen 1991, and Broome 1991).

In summary, the use of preference and preference satisfaction is judged as a questionable means of capturing well-being, and maximising preferences is not thought to be the same as maximising utility. There are limitations with using preferences as a measure of well-being (e.g. adaptive preferences), and this impacts on the interpretation of utilitarianism as a moral theory. In effect, different people might derive different utility from the same preference (e.g.

a unit of production) and therefore, a policy that maximises preference satisfaction might not maximise utility and social welfare.

Utilitarianism is not a simple (one-dimensional) theory to oppose, due to the wide-ranging interpretations of what it is that utilitarianism and the metric of utility can and do stand for, when the approach to an allocation problem is said to be following the utilitarian theory. Utilitarianism is undoubtedly a very broad theory, some argue that it demands too much, and tries to do too much (e.g. Bailey 1997, Hahn 1982).

Many of the following theories have arisen out of dissatisfaction with utilitarianism, and they seek to provide an alternative to utilitarianism as a guiding theory of justice.

3.2.2 Rawls' Theory of Justice

John Rawls published his *Theory of Justice* in 1971. His work has been hugely influential in the area of moral philosophy and public policy. Rawls sets out his views on the basis for social justice, using the position of social institutions to provide a 'just' basic structure for the social system, and by applying a framework around social contract theory. Rawls presents what he believes to be the basis for a system of procedural justice, where 'well-being' is based on an index of primary goods (e.g. liberty, education, income). The over-riding theme from Rawls is that of 'justice as fairness' within the basic structure of society i.e. major institutions, and the basic structure of society is the primary subject of Rawlsian justice.

Rawls was dissatisfied with the doctrine of utilitarianism (maximising the sum total), he did not believe that utilitarianism in any of its various forms was a just approach for the assessment of social welfare. For Rawls, it was not just that "some should have less in order that others may prosper" (Rawls 1999, p13), and Rawls presents his theory as an alternative to utilitarianism.

Rawls sees justice as a set of principles for assigning rights and duties in the basic institutions of society and defining the appropriate distribution of benefits and burdens of social co-operation.

Rawls sets out two principles of justice:

1. Each person has an equal right to a fully adequate scheme of equal basic liberties which is compatible with a similar scheme of liberties for all.

2. Social and economic inequalities are to satisfy two conditions. First they must be attached to offices and positions open to all under conditions of fair equality of opportunity; and second, they must be to the greater benefit of the least advantaged members of society.

The two principles are presented in a lexicographical ordering[5], with basic liberties over-riding the consideration of the social and economic arrangements and inequalities (i.e. who gets what).

Rawls puts priority on justice over efficiency and welfare. He introduces what he calls 'the difference principle', which states that social and economic inequalities must be to the benefit of the least advantaged members of society.

Rawls is very clear in his theory that it sees no one single objective or criterion for the assessment of social arrangements (social welfare), unlike utilitarianism which has a clear and dominant criterion (maximising utility). Rawls, therefore, presents a theory which revolves around pure procedural justice, in order to prevail when a range of criteria may be used for the assessment of welfare. The procedural justice put forward by Rawls employs the framework of the social contract i.e. an agreement between potential citizens (or between such citizens and a potential ruler) about the terms on which they are to enter into social and/or political relationships.

'Justice as fairness' stems from the idea that the principles of justice are agreed in an initial situation that is fair (procedural justice). This captures what Rawls refers to as the 'original position', which is one of the central elements of Rawls's theory. It is a hypothetical situation where persons consider choices surrounding principles of justice under a position of uncertainty. Rawls places persons in an original position of theoretical equality, applying his novel conception of a 'veil of ignorance', which is introduced to remove all possibility of unfairness in decisions to be made. Rawls argues that rational persons in the 'original position' would adopt what decision theorists call a "least worst", or "maximin", rule, whereby, the position of the worst-off persons (least-advantaged) is of prime importance to decision makers. This maximin rule supports the difference principle, which is introduced by Rawls

[5] This is an order which requires us to satisfy the first principle in the ordering before we can move on to the second, the second before we consider the third, and so on. A principle does not come into play until those previous to it are either fully met or do not apply (Rawls 1999, p38).

as a strongly egalitarian conception, in the sense that unless there is a distribution that makes both persons better off (in a simple two-person case), an equal distribution is preferred (Rawls 1999, p65).

Rawlsian theory presents a number of simple concepts (his principles of justice, the difference principle, the original position and the veil of ignorance), which together Rawls believed would bring about his conception of justice as fairness. Rawls states that:

> ...we should attempt to find simple concepts that can be assembled to give a reasonable conception of justice. The notions of the basic structure, of the veil of ignorance, of a lexical order, of the least favoured position, as well as of pure procedural justice are all examples of this. By themselves none of these could be expected to work, but properly put together they may serve well enough. It is too much to suppose that there exists for all or even most moral problems a reasonable solution (Rawls 1999, p77)

The overall framework introduced by Rawls is that of a simple well ordered society, where everyone is presumed to act justly and to do his part in upholding just institutions.

Rawls' Theory of Justice has been subject to much debate, receiving both high praise and stern criticism, including criticism and clarification from Rawls himself (Rawls, 1993). Rawls adjusted his views on the assumption of a 'well-ordered' society, and introduced a notion of 'reasonableness' for persons in society. Both the original position and the difference principle have been subject to numerous objections (e.g. Nozick, 1974), and there have been broad discussions (e.g. Harsanyi 1976, Campbell 2001) on the theory concluding that Rawls makes some empirical leaps of faith in the application of his principles, and the framework for social co-operation.

Where Rawls is subject to criticism on the difference principle, he goes to some trouble to distinguish the difference principle from the economists 'maximin' criterion, which is generally understood as a rule of choice under uncertainty. Rawls clearly states that he placed the difference principle as a principle of justice, which he viewed as distinct from the maximin criterion (Rawls 1999, p72). Rawls presents the difference principle as a 'very special criterion', applying to the basic structure of society via representative individuals whose expectations are to be estimated by an index of primary goods. He does not wish the difference principle to be associated with the issues related to uncertainty i.e. risk aversion, and extreme attitudes to risk, as "there are many considerations in favour of the difference principle in which the aversion to risk plays no role at all" (1999, p72-3).

Rawlsian theory has been cited, in much of the economics literature around social welfare, to support the use of a maximin decision rule in resource allocation decisions. In many instances this has been in the context of micro-level allocation decisions. However, Rawls consistently rejected the use of his theory to address micro-level issues, pointing to the fact that his theory (difference principle) applied to 'macro' issues (e.g. determination of taxation policy), not the micro examples which have commonly been set out in objections to his theory. Indeed, in his theory of justice, Rawls illustrates the difference principle by considering the distribution of income among social classes (Rawls 1971, 1999).

Rawls' theory stands out as a conceptual alternative to utilitarianism. It is generally accepted that the theory is a very valuable addition to the literature on moral theories of justice, however, within the approach there are unresolved tensions between elements of the theory and broad empirical assumptions (Campbell, 2001).

3.2.3 Libertarianism: Robert Nozick's Entitlement Theory

Libertarian justice stresses individual rights as the basis for social organisation, with justice being a matter of each individual getting what they are entitled to through the exercising of their rights (Campbell, 2001). This is the approach adopted and set out by Robert Nozick in his entitlement theory (Nozick, 1974), where justice is a normative structure which dictates that all individuals may choose to do what they like as long as they do not infringe the rights of others. Libertarian justice is primarily about freedom, and does not address notions of equality in social and economic positions, or the distribution of welfare, directly. Although Rawlsian justice may be regarded as an example of welfare liberalism, as it gives prominence to individual rights, it is also distinct from the more formal presentation of minimal-state rights-based libertarianism (e.g. Nozick's entitlement theory).

Nozick's theory of justice was set out in his seminal work, '*Anarchy, State and Utopia*' (1974). Nozick took up a classical libertarian theory of a minimal State, limited to protecting its citizens against force and fraud, which he believed could arise legitimately, without violating anyone's rights. Nozick challenged the assumption that justice requires extensive State involvement for the redistribution of wealth (by coercive means like progressive taxation) in the direction of equality. His work is seen as a challenge to utilitarianism, and a challenge to the theory of justice put forward by Rawls.

The entitlement theory begins with a strong formulation of individual rights, and it thereafter contends that an extensive state is not justified.

Nozick's entitlement theory is based on three principles. Nozick uses the terminology of 'principles [theory] of justice in holdings'; the term 'holding' used to describe the goods, money and property of all kinds that people have. These principles refer to justice in acquisition, justice in transfer, and the theory also considers the rectification of injustice:

Principles:

1. A person who acquires a holding in accordance with the principle of justice in acquisition is entitled to that holding.
2. A person who acquires a holding in accordance with the principle of justice in transfer, from someone else entitled to the holding, is entitled to the holding.
3. No one is entitled to a holding except by (repeated) applications of 1 and 2.

Nozick states:

> "The general outlines of the theory of justice in holdings [the entitlement theory] are that the holdings of a person are just if he is entitled to them by the principles of justice in acquisition and transfer, or by the principle of rectification of injustice (as specified by the first two principles). If each person's holdings are just, then the total set (distribution) of holdings is just. To turn these general outlines into a specific theory we would have to specify the details of each of the three principles of justice in holdings: the principle of acquisition of holdings, the principle of transfer of holdings, and the principle of rectification of violations of the first two principles. I shall not attempt that task here" (Nozick 1974, p153).

Nozick's theory is a 'rights based' approach, and flows from the Libertarian traditional view that the state violates rights if it attempts to transfer property from some (e.g. the rich) to others (e.g. the poor). Nozick's position is that people in society have no obligation to help those worse off than they are.

The entitlement theory is put forward to reflect the opportunity for a minimal state to have a basis for justice; whilst the theory also allows for a diversity of arrangements of divided ownership (e.g. partnerships, loans, leases, gifts). Nozick does not specify the entitlement theory of justice in any detail. He uses the general outlines of the theory to "illuminate the nature and defects of other conceptions of distributive justice" (Nozick 1974, p153).

Nozick's work raises two key objections to 'other' theories of justice (e.g. Rawlsian theory and utilitarianism). He argues that they incorrectly take an 'end-state' or 'current time-slice'

perspective, when a historical view is the correct one. He also argues that they have a 'patterned' approach to distribution which is not feasible, and an unpatterned theoretical approach, as arises through the principles of the entitlement theory, is the correct one.

Nozick argues against 'current time-slice' or 'end-state' principles of justice, which are based only on how things are distributed at that time (who ends up with what), as judged by some structural principles of just distribution. Nozick is critical of current time-slice principles as they are not commonly thought of as constituting the whole story of distributive shares (i.e. he believes it is important to consider how the distribution came about). He presents his entitlement principles of justice in holdings as historical principles of justice.

On the notion of 'patterned' distributions Nozick states:

> "Let us call a principle of distribution *patterned* if it specifies that a distribution is to vary along some natural dimension, weighted sum of natural dimensions, or lexicographic ordering of natural dimensions. And let us say a distribution is patterned if it accords with some patterned principle. ... Almost every suggested principle of distributive justice is patterned: to each according to his moral merit, or needs, or marginal product, or how hard he tries, or the weighted sum of the forgoing, and so on. The principle of entitlement we [Nozick] have sketched is *not* patterned. There is no one natural dimension or weighted sum or combination of a small number of dimensions that yields the distributions generated in accordance with the principle of entitlement." (Nozick 1974, p156-157)

Nozick suggests that almost every alternative principle of distributive justice is patterned. Such as those based around distribution to each according to his moral merit, or needs, or marginal product, or how hard he tries, or the weighted sum of the forgoing, and so on (Nozick, 1974). Nozick argues that under entitlement theory no overarching aim is needed, no distributional pattern is required, what is important is that the system of entitlements is defensible when constituted by the individual aims of individual transactions.

Nozick's entitlement theory is presented to support the argument for a minimal state which enforces no patterns and does not interfere with peoples rights.

It is widely held that Nozick's arguments are incomplete and inconclusive (Campbell 2001, p68). but the force of Nozick's work and the timing of its arrival (a short time after the delivery of Rawls' *Theory of Justice* in 1971, and amidst the 1970-80's new right political

atmosphere of Regan in the USA and Thatcher in the UK), brought widespread discussion on and delivery of Libertarian views, which in themselves focused the minds of others with alternative views, and provoked an open and explicit debate on the assumptions surrounding the orthodox views on welfare assessment.

3.2.3 Egalitarianism

Egalitarianism is a moral principle based on the value of equality amongst persons. Equality is associated with a distributive ideal, a notion of distributive justice and fairness. But the question becomes 'equality of what?' Equality has conventionally been taken to be about income or wealth, but other conceptions of equality have been introduced in order to consider equality in the broader context of welfare and well-being. Some of these have been mentioned in the earlier chapter of this thesis.

Drawing on the moral philosophy literature, Ronald Dworkin has written widely on egalitarianism. His prominent work titled '*Sovereign Virtue: The Theory and Practice of Equality*' (2000) offers a deep insight to issues of egalitarianism and his views are used here to introduce some of the important approaches and issues surrounding equality. Dworkin begins by addressing the question 'what is equality?' and he argues that it is necessary to state what form of equality is important; to distinguish various conceptions of equality. Dworkin, draws a distinction between distributional equality and political equality. Distributional equality concerns some amount or share of resources, whilst political distribution is concerned with political power, and individual rights. Dworkin considers two general theories of distributional equality, equality of welfare and equality of resources. The former theory corresponding to a preference-based view of well-being, and the latter theory corresponding to a more objective measure of equality.

Equality of Welfare

Equality of welfare holds that a distributional scheme treats people as equals when it distributes or transfers resources among them until no further transfer would leave them more equal in welfare. With equality of welfare, goods are distributed equally among persons to the degree that the distribution brings it about that each person enjoys the same welfare. This approach has an immediate appeal, in so far as it is welfare that is of prime importance. However, the objection to the welfare argument is that the conception of welfare is unclear, what does it entail e.g. satisfaction, preferences? Dworkin has argued that the subjective conceptions of welfare do not offer an opportunity to consider equality of welfare in a general theory of equality (Dworkin, 2000).

Equality of Resources

Equality of resources holds that a distributional scheme treats people as equals when it distributes or transfers so that no further transfer would leave their shares of the total resources more equal. Dworkin stresses that equality of resources is based on equality in whatever resources are privately owned by individuals. Dworkin's approach counts personal characteristics and talents amongst the resources owned. The equal resources approach does not mean that each person will have an identical bundle of resources, different bundles will have the same value. To illustrate this, Dworkin introduces the 'envy test', whereby under an initial division of resources, no one would envy anyone else's bundle. Whilst the notion of equality of resources has to contend with similar issues as those discussed under equality of welfare (i.e. handicaps and expensive tastes), Dworkin introduces the economic market (e.g. insurance market, tax system) as an analytical device to illustrate how such matters may be considered under a resource based approach.

Equality of resources is presented by Dworkin (1981) as the preferred conception of equality. Dworkin finds little support for the equality of welfare, but endorses the latter resource based conception of equality (although Dworkin uses a broad definition of resources).

As well as welfare-egalitarianism and resource-egalitarianism, other theories surround the notion of equality. Richard Arneson (1989) argues for equality of opportunity, with respect to welfare.

Equality of Opportunity for Welfare

Arneson (1989), finding dissatisfaction with the use of welfare or resources, advocates the use of equal opportunity for welfare as an egalitarian ideal. He states that an opportunity is a chance of getting a good if one seeks it, and for equal opportunity for welfare to be used among a number of persons, each must face an array of options that is equivalent to every other person's in terms of the prospects for preference satisfaction it offers. Arneson, uses the analytical approach of a decision tree to describe how individuals would have a complete set of choices (life-histories). Arneson, advocates that equal opportunity for welfare prevails among persons when all of them face equivalent decision trees; equivalent expected values for each persons best (most prudent) choice of options (Arneson 1989, p86). In such an approach he states that inequalities are therefore down to voluntary choice or differentially negligent behaviour (individuals are personally responsible).

Other egalitarian views

More generally equality is an ever present notion in discussions over distributive justice. When equality is not the whole story (as it is for egalitarians) it is often taken to be part of the story. For example, the capabilities approach from Sen (1992), the theory of justice from Rawls (1971), and the general doctrine of utilitarianism all contain some elements of egalitarianism.

John Rawls has taken up the egalitarian principle, in a libertarian setting, arguing for a distribution of primary goods which favours the worst-off (least advantaged) persons, and arguing for equality of opportunity in the basic structures of society (see section on Rawls later in this chapter). Other contributors (e.g. Sen, 1992) have suggested standards by which equality can be assessed. Sen, like Rawls, is dissatisfied with the use of income and commodities for the assessment of social arrangements and responds to the question 'equality of what' with his theory based on capabilities and functioning (see section on extra-welfarism). Even utilitarianism, which advocates the maximisation of some concept of utility or welfare, has at it's core the belief that whatever metric is used to describe utility it is of equal value to all.

The principle of 'equality of access' is considered a common conception of justice and fairness. For example, in the delivery of health care, equality of access has been regarded as a central objective of many health care systems (Goddard & Smith, 2001). This objective implies that individuals should have an equal opportunity to use health services, regardless of the characteristics and circumstances of individuals or treatment groups (e.g. age, where people live, ability to pay). Equality of access, is limited in its definition and usefulness given its dependence on other notions of equality (i.e. equality of opportunity), and the commonly related notion of equal access according to need (with need itself open to ambiguity and subjective judgment). The literature around equality of access discusses the related issues of 'access as utilisation', that 'ability to pay' may impact on equality of opportunity, the definition of the services for which access is to be provided (e.g. tax-financed health care system versus social insurance system of health care), and access to what level of services (e.g. a minimum set of services, or broader level of services), (Hauck *et al,* 2003).

Egalitarianism as a moral principle

Egalitariansim as a moral principle has shortcomings, although there is a moral intuition in the equal treatment of persons, the broad egalitarian approaches described above are characterised as 'high-level theory', as 'distributive ideals', and unable to offer a practical approach to the assessment of social arrangements (social welfare), (Arneson 1989). Such

high-level issues may be present when analysing any theory of justice and fairness, and as such are part of a more specific theory of justice i.e. a special case in all theories of justice. Egalitarianism, either generally, or in its specific theoretical presentations may not be a sufficient basis for consideration as a specific conception or theory of justice (Campbell 2001, Konow 2001). However, given the prominence of equality as a moral value, some degree of egalitarianism is an essential ingredient within systems of justice.

3.2.4 Extra-welfarism: Sen's Capabilities Approach

Whereas welfarism, the standard approach of normative welfare economics, seeks to evaluate social arrangements on the basis of individuals' utilities (values, preferences) alone, an extra-welfarist approach admits non-utility information about individuals into the evaluation of alternative actions or states.

Sen (1979) introduced a theory where an individuals 'basic capabilities' (i.e. a person being able do certain things) are a particularly important class of non-utility information. Following the writings of Sen, Culyer (1989) supports the use of non-utility information in the consideration of social arrangements in the provision of health care.

Sen believed that mainstream conceptions of justice (e.g. utilitarian and Rawlsian theories) had serious limitations, and were missing some notion of 'basic capabilities'. In the assessment of social welfare, and in response to concerns over inequalities, and the question 'equality of what?', Sen presented his 'capabilities approach' as an alternative to these more mainstream theories of justice (Sen 1979, 1982, 1992).

The capabilities approach is based on Sen's view that social arrangements should be assessed through the consideration of a person's well-being; with well-being (or good) extending beyond the economists common notion of "utility" (i.e. utility as preference satisfaction). Sen presented his theory as an alternative to mainstream welfarism, where often it was an income or commodity based measure of preference (and utlity) that was used to compare alternative social states.

The capabilities approach rests on Sen's belief that:

> ... "the well-being of a person can be seen in terms of the quality (the 'well-ness', as it were) of the person's being. Living may be seen as consisting of a set of interrelated 'functionings', consisting of being and doings. A person's achievement in this respect can be seen as the vector of his or her functionings. The relevant functionings can vary from such elementary things as being adequately nourished, being in good

> health, avoiding escapable morbidity and premature mortality, etc., to more complex achievements such as being happy, having self-respect, taking part in the life of the community, and so on. The claim [Sen's claim] is that functionings are constitutive of a persons being, and an evaluation of well-being has to take the form of an assessment of these constituent elements. Closely related to the notion of functionings is that of the *capability* to function. It represents the various combinations of functionings (beings and doings) that the person can achieve. Capability is, thus, a set of vectors of functionings, reflecting the person's freedom to lead one type of life or another. Just as the so-called 'budget set' in the commodity space represents a person's freedom to buy commodity bundles, the 'capability set' in the functionings space reflects the persons freedom to choose from possible livings." (Sen 1992, p39)

The capability to function is presented as an important part of social evaluation. Capability is primarily a reflection of the freedom to achieve valuable functionings; a person's freedom to achieve well-being. The approach is presented as an alternative to relying on utility as a guide to personal well-being and as the basis of the assessment of social welfare and equality.

Sen did not believe that the standard approach of normative welfare economics, termed "welfarism" was a valid way to consider well-being and social evaluation. Sen argues that non-welfaristic concepts may have a great relevance to political discussions and practical judgments, arguing that the relevance of non-utility information to moral judgments is the central issue involved in him disputing welfarism (Sen, 1982, p363). Culyer (1989) has applied this extra-welfarist framework in the evaluation of health care.

Sen does not discuss the development of appropriate indices which might be used to describe and measure the relevant functionings and capabilties. However, he does acknowledge that the capabilities approach requires consideration of a variety of 'doings and beings', that there are difficulties in defining the functionings and capabilities that make up the relevant evaluative space (Sen, 1992). Sen recognises that there will be certain trade-offs between different functionings and capabilities. He sees such discrimination between functionings and capabilities as an integral part of the capability approach.

Sen's framework exhibits the characteristics of both the economist and the policy adviser as he explicitly accepts that any attempt to achieve equality of capabilities must take note of 'aggregative considerations', including efficiency.

Sugden has presented a useful critique of the capabilities approach (Sugden, 1993), and whilst congratulating Sen on the attractiveness of the theory he has highlighted ambiguities, confusion, and both theoretical and practical problems. The critique offers a more general observation, that the notion of equality of capability is a little abstract, especially outside of simpler applications, such as those dealing with extreme situations in developing countries (where analysis is of the basic capabilities of food, shelter and freedom from persecution). In general the capabilities approach set out by Sen is regarded as a conceptual ideal. The foundations of Sen's theory are in the definition of what constitutes well-being, or a good life, and translating these foundations into a practical means for the evaluation of social arrangements, brings some complexities. The practical and operational issues for the theory are not detailed in anyway by Sen.

Culyer (1989) has developed an extra-welfarist approach to the evaluation of health care. In this approach Culyer shifts the objective function to be maximised from utility to 'health', with the maximisation of health (non-utility information) subject to budget constraints being the basis upon which different social states may be ordered. This approach to extra-welfarism builds on the 'decision-maker' perspective to cost-benefit analysis suggested by Sugden & Williams (1978), where the decision-maker values (and respective weights) are deemed to be the objectives for any evaluative analysis. Culyer states that the characteristics (quite broadly defined) of individuals may be relevant in the evaluation of health care, although the importance of such characteristics are considered to be contingent on the context of the decision (e.g. Culyer advocates that in the context of health the characteristics that are relevant should be related to the concept of need).

The extra-welfarist debate in health care has centred around the maximand for the evaluation of health care, with Culyer's notion of 'health' being open to different interpretations (e.g. the QALY as a measure of health outcome), and the extra-welfarist approach being classed a 'non-welfarist' approach (e.g. Tsuchiya & Williams 2001, Coast 2004). In the evaluation of health care it is argued that health *per se* is only one of a range of issues that may be relevant (Birch & Donaldson, 2002).

Table 1. Summary of Theories of Justice

Theory /Approach	Principles	Decision Making Perspective
Utilitarianism	A theory of the personal good	Goodness (and badness) is evaluated in terms of individuals' utilities (or welfare). Different approaches/interpretations of utility/welfare may be allowed, but non-utility information should not be included.
	Evaluative consequentialism	An assessment of utility/welfare is based on the consequences of actions. There may be some variability in what constitutes a consequence of an action, but all considerations related to consequences.
	Interpersonal comparability	Decision maker feels able to make comparative judgements about states of the world in which some persons are made better off and others made worse off.
	Distributive neutrality	Assumes that the measure of personal good indicates who it is that benefits the most from the action, and that no further consideration is required with regard to the persons in which the benefits are seen (i.e. it is the size and not the location of gain that matters)
Rawlsian Justice	Social Justice	Societal perspective, using major institutions (institutional setting), where the institution must be just (fair).
	Justice as fairness	Political conception of justice, with a system of pure procedural justice (e.g. equality of opportunity, difference principle, original position and veil of ignorance).
	Basic liberties as a prime over-riding issue.	Each person has an equal right to a fully adequate scheme of equal basic liberties which is compatible with a similar scheme of liberties for all. Rawlsian justice does not allow that that sacrifices imposed on a few are outweighed by the larger sum of advantages enjoyed by many
	Pluralist approach	No one over-riding independent criterion for the assessment of social justice – therefore a system of pure procedural justice (employing a social contract framework)
	Primary goods as relevant	Rawls concerned with 'primary goods'. Health and vigour – regarded as natural goods (not included in Rawls' chief primary goods)
	Social Contract	Strict compliance model, where everyone (in a well-ordered society) is presumed to act justly and to do his part in upholding just institutions.
Nozian Justice (Right's Based)	Minimal state / small Govt.	Nozick (rights-based theory) would favour a system where government intervention and government provision of services was kept to a minimum.
	Historical view of justice	Nozick believes it to be important to consider how the distribution of goods came about, taking a historical view not a limited 'current time slice', or 'end-state' view, as these do not portray the whole story of distributive shares.
	The correct approach to justice is a non-patterned one.	Nozick believes that justice should result from a set of underlying principles, not from a patterned approach to justice and distribution.
Egalitarianism	Based on the value of equality amongst persons	Dependent upon the specific perspective adopted by decision maker e.g. equality of welfare, equality of resources, equality of opportunity (with subsequent definitions of objectives) Equality seen as an overriding value or aspiration.
Extra-Welfarism	Advocates the importance of non-utility information	Extra-Welfarism advocates the use of non-utility information, e.g. considering relevant characteristics over and above the utility information relevant in evaluating alternative actions. Admits the use of non-utility information about individuals into the process of comparing social states.
	Sen's Capabilities approach: stresses the importance of 'basic capability equality'	A form of extra-welfarism that assesses social arrangements through the consideration of a person's well-being, using a vector of functionings, offering a person the 'capability' to lead one type of life or another. The approach does not allow utility to be the sole guide to personal well-being, and assessment of equality (i.e. includes non-utility data).

3.2.5 Overview of theories of justice

The theories of justice outlined above are each associated with a particular perspective. Utilitarianism may be regarded as the incumbent theory with the other theories (borne from a general dissatisfaction with the utilitarian approach) developed and presented as alternatives to utilitarianism. Table 1 presents a brief summary of the main principles upon which the theories are founded and the related decision-making perspectives. Although quite distinct from one another, there is common ground between some of the theories presented. Not surprisingly there is some common ground between utilitarianism and the extra-welfarist approach. But more surprising are the areas of common ground that may be drawn from the Rawlsian and Nozian theories. Egalitarianism has some common areas with various other theories, but importantly, there is some question over the place of egalitarianism as a specific theory of justice (Konow, 2003).

The theories from Rawls and Nozick, whilst appearing to be at some distance from one another, find agreement on a number of different issues. Unlike utilitarianism (and extra-welfarism) which takes a more context specific perspective, Nozick and Rawls would advocate a societal perspective for a theory of justice where society is generally regarded as 'just'; Nozick argues that 'people want their society to be and to look just', whilst Rawls strongly argues for institutional arrangements to reflect what he calls 'justices as fairness'. Notably, Rawls and Nozick would both argue for a set of 'underlying generating principles', both agreeing that there are no dominant aims or criteria for the evaluation of actions. Nozick discusses his dislike for 'patterned' approaches in the context of theories of justice (evaluation of social states), believing that Libertarian views upset any patterned approach. Whilst Rawls pursues a procedural approach to justice believing there are no dominant criteria for social evaluation and the 'process' itself should be able to accommodate consideration of a range of decision criteria. Nozick would seem to agree with a procedural perspective, although he would argue for a minimal (small) state with limited government intervention, whilst Rawls would not follow in such a way. Rawls considers the institutional arrangements for society, and as such would appear to argue for a more formal institutional (oragnisational) role for government, a more extensive state, and this is at odds with Nozick's entitlement theory. Furthermore, Nozick and Rawls would agree that using 'end-state' outcomes (descriptions) is inappropriate, with Nozick arguing for a historical perspective when it comes to distributions of goods, and Rawls arguing through the more procedural approach for a broader view than 'end-state' in the evaluation of actions. Where Rawls and Nozick argue against a 'patterned' and 'end-state' approach to the social evaluation of actions, we can see that utilitarianism and extra-welfarism positively support a

patterned and end-state approach, using consequences and criteria (e.g. maximising utility) to determine which actions should be chosen.

Rawls and Nozick, together with egalitarians, would argue in support of individual rights (although the interpretation of rights would be somewhat different). Rawls argues that it is not acceptable that some should have less whilst others prosper (unless inequalities are in favour of the worst off members of society); arguing against the utilitarian concept of maximising utility regardless of the location of the utility gains. Nozick has as the foundation of his entitlement theory the fact that the state should not be able to violate the rights of individuals, as may be seen in a utilitarian approach that maximises utility regardless of the infringement on individuals rights. Although Nozick argues that egalitarianism is not supportive of a rights-based perspective, the many interpretations of egalitarianism confuse this issue. It is clear that egalitarians have as a basis for their arguments the fact that the state should have equal concern for all (e.g. arguments for equal opportunity), yet this says little about the protection of basic human rights as advocated by Nozick. Although Sen is not labelled an egalitarian (he would be regarded as an extra-welfarist) he argues for an egalitarian notion in his championing of equality of capabilities as an objective in the evaluation of social arrangements. Whilst it would appear that utilitarianism has very little to say about individual rights, some utilitarians have attempted to put some structure around utilitarianism as a societal approach (e.g. rule-utilitarianism where rights are seen as 'rules'), arguing that institutional arrangements and underlying principles have some part to play within a utilitarian system which can respect individual rights and be consistent with a broader theory of justice (Bailey, 1997).

Sen's capabilities approach, would seem to have many overlaps with a utilitarian approach, although it seeks to move utilitarianism away from the a historical view of utility as an 'income related measure of well-being' i.e. discussion of welfare in the context of the marginal utility of income. Sen's dissatisfaction with utilitarianism is based on its narrow informational base, i.e. he argues that welfare as a function of utility or individual preference data is a very restrictive approach, and argues in favour of the use of important non-utility (non-preference) data. However, Sen's approach would seem to agree with utilitarianism that the notion of utility, however defined, is at a personal level (i.e. what persons can do with products, benefits, income) and he strongly argues for consideration of human diversity. Both of these would be arguments strongly proferred by utilitarians. Sen would still accept that utility is a dominant criteria upon which to evaluate social actions, although the consideration of non-utility data would also have an important role. Indeed where Sen

argues in favour of consideration of the 'evaluative space' this may be aligned to the common notion of 'utility space'.

Egalitarianism shares common ground with all theories, as all other theories of justice tend to place equality in some important central role. For example, Rawls (equality of opportunity), Nozick (maximum equal liberty), and Sen (equality of capabilities) all have equality as a major factor within theories of justice, and utilitarianism too has a place for equality i.e. the unitary value of utility is of equal value to all. Equality of one form or another receives widespread acceptance as a principle of justice, and it is a required part of any plausible theory of justice (Sen, 2002). Sen argues that "an income egalitarian, a champion of democracy, a libertarian and a property-right conservative may have different priorities, but each wants equality of something that is seen as valuable – indeed central – in the respective political philosophy" (p659). Sen also argues that equality as an abstract idea "does not have much cutting power" (p660), and that there is a need to specify what it is to be equalised. The notion, or basis, of equality within competing theories may differ and as such it has been argued that equality can not form the basis of a theory of justice and/or fairness, and at best it is a special case of other frameworks (Konow, 2001). Konow (2003) presents evidence from the general literature on theories of justice to suggest almost no empirical support for egalitarianism as a specific approach to the consideration of justice and fairness.

The above discussion is outline in its assessment of the interactions and commonality between competing theories of justice. However, the discussion does encourage further consideration of the 'plurality' of arguments within the notions of justice and fairness. Konow (2000) proposes an integrated justice theory that synthesises elements from previous approaches (common theories of justice), placing emphasis on the context of the evaluation required, and this may be a worthwhile avenue to pursue given the commonality we see between competing theories of justice, and the often complex, context-specific settings that present themselves for consideration against a backdrop of justice and fairness. This notion of the 'plurality of objectives' (societal objectives) has been raised in the discussion of 'equity' and the social evaluation of alternative actions by health policy and decision makers (Culyer, 2001). Below, the theories of justice are investigated further in a more context-specific fashion, moving the discussion to the area of health care, and the social evaluation of health interventions in a resource constrained environment (i.e. the process of health technology appraisal).

The theories of justice outlined above provide a good basis for discussion of the general use of such frameworks in the area of health care, and specifically on the appraisal of health technologies. The theories outlined have been presented without an emphasis on health care. Given the current interest in the area of health care, and specifically on the appraisal of health care technologies, the next question relates to the relevance of the theories to the issues in the context of health care.

3.3 Theories of justice and their implications for the appraisal of health technologies

3.3.1 Health technology appraisal process: A thought experiment

Health technology appraisal is used in this thesis to demonstrate the decision-making process around the evaluation of alternative health care technologies. Health technology appraisal (HTAP) provides an example of how the UK NHS makes funding decisions for new and established technologies, at either a national or regional level.

HTAP is a common process. Health care commissioners and providers (local, regional or national) are continually in a position whereby limited health care resources must be assigned to competing programmes. HTAP is broader than the process of health technology assessment (HTA). HTAP makes a judgment as to the overall value of a technology to a particular population, subject to the objectives of the organisation or the health care system it is servicing, and in doing so it may comprise considerations related to institutional, political or social priorities. The consideration of health technology appraisal (HTAP) provides an opportunity to address one of the 'everyday equity [allocation] problems' (Young, 1994) in health care.

To more fully understand the place of the theories of justice discussed above in the context of health care and HTAP a "thought experiment" is undertaken asking:

(a) What would HTAP look like if it was firmly based on each of the theories?
and,
(b) Does the present process for HTAP address (to any extent) the concerns of each of the theories?

In the above thought experiment each of the theories of justice are considered in the context of question (a), and this is followed by an examination of question (b) using observations on the conduct of HTAP in the UK NHS.

In the UK the NICE health technology appraisal programme (NICE, 2004) is the most public and prominent process for health technology appraisal, and the NICE technology appraisal programme is used here to consider the theories of justice in the above thought experiment. The process of technology appraisal has been informed, for this thesis, using published policy documents (NICE, 2004), an overview of NICE guidance published up to August 2006 (discussed further in Chapter 4), and the related recent NICE guidelines on social value judgments within the technology appraisal process (NICE, December 2005, Rawlins & Culyer 2004).

Thought experiments have been used widely in many fields, but are particularly common in philosophy, physics and mathematics (Brown,1993). They are a way of presenting hypothetical, or imaginary, situations to enable a greater understanding. There are different types of thought experiment, but all types display a patterned way of thinking to allow explanation and/or prediction of potential or hypothetical events. Thought experiments can frequently be seen in economic analysis, in a limited form, as 'what if' questions or scenarios, or in a counterfactual way to speculate on other possible outcomes. Rawls (1973) presents the 'veil of ignorance' as a type of thought experiment. In philosophy, thought experiments are typically used to present an imaginary situation with the intention of examining an intuitive response about 'the way things are' (Sorensen, 1992). That is the approach taken here. The details surrounding the thought experiment undertaken are presented in Appendix 1.

The main summary findings from the consideration of the theories of justice against the NICE technology appraisal process, via the thought experiment, are:[6]

Utilitarianism

Although NICE argue that they follow a partly utilitarian approach, within the health technology appraisal process (NICE, 2005), there are good grounds for arguing their process is inconsistent with a utilitarian framework.

[6] The summary presented here is relevant to the second of the questions in the thought experiment (question b), whilst Appendix 1 presents against both questions.

Whilst NICE does take into account, in a significant way, the maximisation of health gain (health-related welfare), and this may find some association with the utilitarian ideal of maximising utility, the objective of the NHS and NICE is not solely to act as a utility maximising organisation; the NHS and NICE have other objectives to contend with. The health technology appraisal process does use consequences as a central element of the process, using the QALY as a preferred proxy measure of preference and/or utility, but it is questionable whether a measure such as the QALY is consistent with the utilitarian requirement for a theory of the personal good (the principle of utility). It is clear that a range of non-utility data is also present in the decision making process, and that outcomes may be based on a broad definition of consequences, or indeed on information that is not directly linked to the consequences of an intervention. The present process of HTAP is generally consistent with (does address) the key utilitarian characteristic of interpersonal comparisons, and there is an element of support for the principle of distributive indifference (but this is indifference across consequences rather than the value attached to them).

In conclusion, there are good grounds for arguing the HTAP process is inconsistent with a utilitarian framework.

Rawlsian Justice

From the current NICE process of health technology appraisal it is not possible to see the application of Rawlsian principles. Decision makers undertaking health technology appraisal do not appear to make any explicit policy statements encompassing the Rawlsian principles surrounding basic liberties, equality of opportunity (within process), or on the Rawlsian notion of priority of justice over efficiency, for example on priority being given to the least advantaged persons (as defined by the HTAP institutions). There is an absence of explicit statements surrounding the Rawlsian priority of justice over efficiency, and it would appear, through the emphasis placed on cost-effectiveness, and health maximisation, that institutions value efficiency arguments highly in the appraisal process. This reliance on dominant criteria in HTAP i.e. clinical and cost-effectiveness, and the prominent use of measures of health gain (e.g. QALY), lead to a conclusion that the present approach to HTAP is not consistent with the Rawlsian 'pluralist approach'. In the NICE approach to HTAP the prominence of arguments related to distributive justice (e.g. allocation of resources, distribution of health outcomes), rather than on the central issue of procedural justice (i.e. underlying principles) would indicate that HTAP does not address the major considerations of the Rawlsian approach of 'justice as fairness'.

Whilst there are indications that some of the elements of the Rawlsian framework may be present in part within the health technology appraisal process (i.e. the institutional arrangements, and a strong element of process), there appears little to support a view of the current process as Rawlsian. There is no explicit statement on higher order (Rawlsian) principles, there is an emphasis on efficiency rather than on procedural justice.

Nozian Justice

The present NICE process of HTAP does not appear to address, to any extent, the concerns of Nozicks theory of justice. The UK NHS and NICE are inconsistent with Nozick's theory of a minimal state, and they are inconsistent with Nozick's principles of justice. The NICE technology appraisal process generally takes a 'current time-slice' view of health care needs, and an 'end-state' view of resource allocation decisions, both of these being inconsistent with a Nozian theory of justice.

Egalitarianism

There is a clear presence for equality in various forms within the NHS and within the NICE appraisal process, but it is not clear at a conceptual or operational level how the notion of equality is considered when making judgments on the appropriateness or not of a particular health technology for a particular patient group.

It is difficult to consider the egalitarian approach in the context of health technology appraisal as the objectives are often ambiguous and unclear. In the context of technology appraisal, the broader egalitarian principles, such as equality of welfare and equality of resources, are abstract notions. Furthermore, less abstract notions of equality for health care, such as equality of access, have not been directly related to the overall health-related objectives of the NHS (Newdick, 2005).

The UK NHS does appear to place emphasis on equality in various dimensions. For example, it is considered to hold equality of access to NHS services as a fundamental principle, and to seek to limit inequalities in access to services across and within regions. Furthermore, it does not regard ability to pay as a basis for discriminating between persons requiring health care. Yet, within the NICE process of health technology appraisal there seems little evidence of specific equality objectives, in relation to the appraisal of technologies and the restrictions on access to (use of) technologies. There are statements surrounding methods and process, that factors such as age, gender, and ethnicity, should not be used as priority setting criteria (NICE, 2005), unless they are related to clinical

benefits or risks. There is a statement from NICE that at the present time it regards health outcomes (i.e. QALYs) as being of equal value to all persons (NICE, 2005). However, there is an absence of higher level principles on equality within the UK NHS, and at an operational level within the NICE health technology appraisal process. It is clear that egalitarianism is not a sole or dominant framework for health technology appraisal within NICE.

Extra-welfarism

Health technology appraisal would appear to follow from the extra-welfarist belief that utility information is an insufficient basis for evaluating social welfare. The NICE appraisal process places health outcome as being of prime importance. Clinical effectiveness is one of the primary inputs to the appraisal process, followed by cost-effectiveness, which seeks, as commonly applied, to maximise a measure of health gain (e.g. life years saved, QALYs gained). Even if QALYs were regarded as a measure of utility, other non-utility information is used in the appraisal process; this being consistent with the extra-welfarist perspective.

The NICE appraisal process would appear to adopt a 'decision-maker' perspective, as suggested by Sugden & Williams (1978), and later by Culyer (1989); with the values of the decision-making body leading the objectives in the appraisal process. Cost-effectiveness analysis, which is very prominent in the technology appraisal process is undertaken from a largely extra-welfarist perspective, providing information on the maximisation of health gain subject to resource constraints.

However, extra-welfarism, as suggested by Sen, also sees utility (or preference-based) outcomes as an important element in the evaluative process, and it is not apparent where such a measure of utility (other than the use of QALYs as a potential utility outcome) appears in the NICE technology appraisal process (i.e. there is no basis for assessing the value of health outcomes to persons receiving them).

General Summary

The NICE health technology appraisal process shows no one clear alignment to any one theory of justice. The broad extra-welfarist framework described by Culyer (1989) is the most prominent framework in the appraisal process described for NICE. The process does demonstrate some elements of procedural justice, which may be aligned to the 'accountability for reasonableness' approach suggested by Daniels & Sabin (1998). The elements of procedural justice in the NICE process may be suggestive of some broad alignment to a Rawlsian conception of justice, although there are many inconsistencies

between health technology appraisal in the NHS and Rawlsian justice. There is an underlying notion of equality within the NHS, and elements of equality in the foundations of the NICE technology appraisal process, and also inherent in the methods used for appraisal. But there is an absence of a clear egalitarian framework. The health technology appraisal process is a consequentialist approach, but it is not limited to consequences alone. The appraisal process fails to link consequentialism to utilitarianism due to the absence of a coherent theory of the personal good (the valuation of outcomes in terms of utility and/or well-being).

3.4 Summary: Messages from the discussion on theories of justice

It is possible to learn from the exposition of the various theories of justice. The theories of justice discussed here all have something to say about fairness and the distribution of 'goods' (resources), and can be used to translate notions of fairness into some policy related framework.

Some would argue (e.g. Young, 1994) that in addressing everyday equity problems, such as health technology appraisal, there is no need to consider the broader theories of justice which seek to address the larger challenges faced by society. Indeed the theories of justice surveyed above do seek to examine and influence the broader questions for society. For example, the theories are not directly related to the maximand of health and health related social welfare, which unlike income and physical resources cannot be easily redistributed among people or groups in society. All of the theories discussed fall short of offering a means by which a redistribution of goods can be achieved in practice, and "most are theoretical and remote from practical implementation issues" (Hauck *et al* 2003, p22). However, the consideration of these theories has provided a solid foundation for the discussion and examination of the issues facing decision-makers in the context of health technology appraisal.

Across the theories of justice discussed above, it is apparent that there are very distinct views on the manner in which social arrangements are considered, and explicitly evaluated in society. It is clear that there are different interpretations of the key terms of justice, fairness, and equity; the latter term [equity] very rarely appearing in the discussion of the more general theories of justice.

Theories of justice vary, from those with a key central criterion for the evaluation of social states e.g. utilitarianism and utility, to those with very broad pluralist approaches involving a procedural format e.g. Rawlsian justice. Yet, even those focused on key decision criteria

make it clear that there are a great number of instrumental issues which feed into the evaluation of what is 'good', or which alternative [state of the world] is the correct one (e.g. what policy makers 'should' do). The theories all offer insights, often at variance with each other, on the difficulties associated with the evaluation of social arrangements.

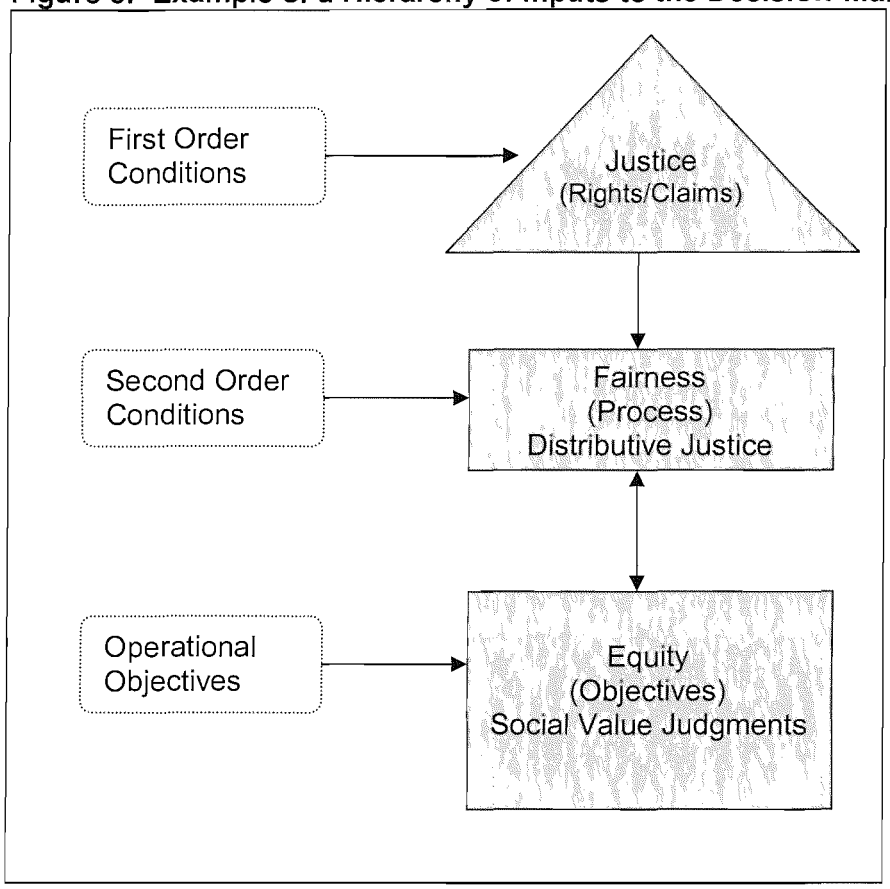
The linking of the theories to health technology appraisal has assisted with the formation of a view that there is no one single value judgment or theory that can be directly related to health care, and health technology appraisal more specifically. It is clear that in resource allocation problems, such as health technology appraisal, these are multivariate problems with a plurality of approaches. These are allocation problems, with resources in the NHS constrained. Yet, there are broader issues related to society that make health technology appraisal more than a simple division of goods among a designated number of potential recipients. There are a wide variety of important issues related to the assessment of the value of an intervention to the NHS, and to society more broadly. As Konow (2001) has suggested generally, an integrated justice system that synthesises elements from different theories of justice and places the emphasis on the 'context' of the evaluation required, may be relevant for the assessment of health care allocation problems.

It is apparent from the examination of the literature that there are different considerations at the differing levels of the decision making process. A framework is suggested below to characterise the process of health technology appraisal (Figure 5). The framework, which is stylized, sets out the research areas in the complex examination of the conceptions of justice, fairness and equity, as they impact on, and influence the process of health technology appraisal. This thesis is concerned with the social values that may inform at an operational level, to inform what might be regarded as equity positions relevant for resource allocation problems, like health technology appraisal.

There are considerations, referred to as 'first-order', that are at the very heart of the setting for health technology appraisal. These may be related to the determination of the overall budget for the NHS, or they may be related to the ideology of the NHS as a publicly funded system of health care provision. First-order considerations could capture the important aspects of the technology appraisal process related to rights, liberty, and the over-arching principles that embody societal norms. These first-order conditions are not apparent, or visible, in the present process of health technology appraisal in an explicit manner, and some recognition of these important 'background' considerations is required to set the scene for a 'just' process (i.e. an acceptance of any approach can only be possible following an explanation of it's underlying values). It may be that justice at its higher level of

interpretation is not directly manifest in much of the health technology appraisal process, but it is not absent, and it's place and influence would be a helpful guide to the more operational process and objectives of specific technology appraisal contexts.

Figure 5: Example of a Hierarchy of Inputs to the Decision-Making Process



Process is an important factor for the conduct of health technology appraisal[7]; procedural justice is a strong motivation in the theories from Rawls and Nozick, and an essential part of any allocation problem. Process is related by Rawls to 'fairness', his concept of justice as fairness, and he goes to great lengths to put in place (theoretically) a mechanism for impartiality. It may be that process can be thought of as a 'second-order' consideration, against the background of explicit first-order resource constraints or moral positions. Much of the literature on theories of justice remains in the area of overriding principles of justice (first-order considerations) and procedural issues related to fairness (second-order issues), and the operational element involved in putting these theories into practice is often missing; other than the presentation of theoretical and/or distributive ideals. Yet, the examination of

[7] This thesis does not examine process in health care in any detail. It is acknowledged that process is an important area for research. Readers are referred to Ham & Robert (2003) for an introduction to process in health care and the application of the 'accountability for reasonableness' approach from Daniels and Sabin (1998).

the theories leads to a conclusion here that the decision maker needs to define and interpret what it is they wish to pursue, in terms of operational objectives (i.e. what may be viewed as equity objectives, or social values).

The above findings that:

- there are a great number of instrumental issues which feed into the evaluation of what is 'good' (e.g. normative policy statements);
- all theories of justice and fairness stress that it is important for the decision-maker to define operational objectives;
- all theories of justice recognise that there will be trade-offs between operational objectives;
- the allocation problems in health care, such as health technology appraisal, draw on a plurality of broader approaches to distributive justice and fairness;

are used here as a point of departure, and the remainder of the thesis is focused on research around the identification of social value judgments which may inform on equity positions in the allocation of health care resources.

4 A REVIEW OF THE EMPIRICAL EVIDENCE ON DISTRIBUTIVE PREFERENCES (SOCIAL VALUES) IN THE HEALTH CARE LITERATURE

4.1 Introduction

This chapter presents a review of the literature reporting empirical findings on the social values, and distributive preferences, associated with health care interventions.

Much of the literature on social values (in priority setting scenarios) has arisen from the use of economic evaluation (cost effectiveness analysis, cost utility analysis) by health care policy makers to consider the relative value of alternative courses of action in resource allocation decisions. Whilst economic evaluation is not the only input to health policy decisions it has become a powerful tool in many health care systems (e.g. UK, Australia, Canada), in the formation of health policy (Ham & Robert, 2003).

Economic evaluation as it is commonly used (in an extra-welfarist framework) assumes that the societal objective is to maximise health (health output) from available resources (Oliver, 2004). This health maximisation objective is a strong assumption and a growing evidence base indicates that there is limited support for it as an over riding social value in resource allocation decisions. Whilst some (e.g. Williams 1996, Nord *et al* 1999, Hauck *et al* 2003) have suggested that economic evaluation can incorporate a range of other social values (other than efficiency), in the form of a weighting of health outcomes, this approach has yet to be adopted in a policy arena, and remains a theoretical suggestion (Rawlins & Culyer, 2004). In practice, the calculation of QALYS has been as a measure of health production, and the conduct of economic evaluation has been an assessment of the relative cost effectiveness of an additional unit of health outcome – i.e. assessment of efficiency and not involving the QALY as a measure of social value (outside of efficiency).

The debate concerning the weighting of health outcomes to reflect priority setting objectives, and concerns over the use of health maximisation (QALY maximisation) as the relevant maximand, have contributed to a growing body of work that examines potential social considerations that may be used to weight outcomes, or to inform health decisions outside of economic evaluation. This type of research has been referred to as 'empirical ethics' (Culyer 2001, Richardson & McKie 2005).

There have been a number of earlier informative reviews of parts of this literature (Sassi *et al* 2001, Schwappach 2002, Dolan *et al* 2005). These reviews have all adopted a specific perspective or context. Sassi *et al* base their review around economic evaluation. Reviews by Schwappach and Dolan *et al* have considered empirical evidence on the use of the QALY, and the QALY maximisation objective. The review undertaken as part of this thesis has used a broader perspective, not limited to evidence around economic evaluation or QALY maximisation, and has sought to examine the literature presenting empirical findings from both attitudinal studies and experimental studies. The earlier reviews have been used to inform the current literature search strategy, and the identification of relevant empirical literature.

4.2 Research questions

The literature review sought to address the following questions:

1. Is the maximisation of health (health gain) a valid representation of the social value of health interventions?
2. Is the social value of a health intervention dependent on factors other than health?
3. Is the social value of a health intervention related to the characteristics of eligible patient groups?
4. Is the social value of a health intervention related to the characteristics of the health intervention?

4.3 Literature search methods

This is a methodological review, and as such it is not best suited to traditional literature searching methods (e.g. such as those used for review of clinical trials). Therefore, the literature search has used a combination of methods. The review used a conventional keyword based search strategy, across electronic databases, in combination with a core set of key references (drawn from previously published review articles, and expert opinion). The use of a set of core references, and citations to these references, is analogous to the 'citation pearl growing' literature searching framework (Hartley 1990, Dolan *et al* 2005).

The search strategy (see below) was developed via an iterative process, with expert input from an information scientist. Earlier search strategies were tested to establish the sensitivity and specificity of the literature identified using a core listing of expected key references. When using very common search terms there is a trade-off between sensitivity and specificity, and the iterative process indicated the usefulness and feasibility of using specific terms in the strategy. For example, the use of the term 'ethics' was investigated for

specificity (and feasibility) prior to its omission from the search strategy used. However, given the broad nature of the topic area the search was inevitably driven by sensitivity rather than specificity, with a large number of the resulting references being excluded.

4.3.1 Search terms:

Area of literature: (applying terms 1 to 7 below)

1. Health care
2. Health policy
3. Health decision*
4. Health technolog$ and (appraisal$ or assessment$))
5. Health care rationing <mesh term>
6. Health policy <mesh term>
7. Health priorities <mesh term>

Specific search terms: (applying terms 1 to 7 below)

1. Equity
2. Justice
3. Fairness
4. Public adj3 preference$ or attitude$
5. Priority setting (or priority-setting)
6. Social value$
7. Societal value$

The above terms were applied in combination (i.e. area terms or/1-7 AND specific terms or/1-7), as required across the database options. Appendix 2 presents the exact search strategies used by database and the output from searching. A large number of titles/abstracts were expected (with a large proportion expected to be irrelevant), but a sensitive search was judged to be the best option, with an initial sifting of the abstracts undertaken to identify potentially relevant references.

Searching was limited to English language materials.

4.3.2 Databases searched

Databases used for the search comprised MEDLINE, EMBASE, ECONLit, PSychLit, HMIC, SCI & SSCI. HMIC was used to capture contributions from the grey literature. Literature searches were undertaken from inception of database (see Appendix 2) up to January 2007. Initial searching covered literature up to January 2005, however a more recent update search captured literature up to January 2007.[8]

[8] In addition, a general review of the published NICE technology appraisal guidance, reports from the NICE Citizens Council, and NICE methods and process reports, has been undertaken, to supplement the literature identified using the above methodology (see Section 4.5.4).

4.3.3 Inclusion / exclusion criteria

The initial sift of the abstracts/titles from the literature search was based on the following criteria:

1. *Excludes:*
    - References purely on patient-level clinical ethics were excluded.
    - References comprising of editorial or commentary style contributions were excluded.
2. *Health context:*
    - Papers were sought which were related to the area of health and health care (e.g. not issues of criminal justice)
    - Papers were sought which were able to inform on the areas of equity and resource allocation decisions within health, health care and general health policy.
3. *Content:*
    - Following the above general sift against criteria 1 and 2 (above), references had to present empirical findings of relevance to the decisions surrounding the matters related to the analysis of equity/fairness (i.e. empirical insights to the dimensions of equity, however defined, and/or trade-off between relevant dimensions of equity).

Given the methodological nature of the literature search it was necessary to make a judgment on inclusion of studies on a case-by-case basis

4.4 Literature search results

The literature search identified 4,878 abstracts and titles (including update searching). Many of the references identified were excluded as they were clearly related to clinical ethics, ethics committees, patient end-of-life decisions (e.g. do not resuscitate decisions, decisions over ventilation removal), abortion in a policy context, HIV/AIDS related clinical issues, mental health and criminal justice. Following a detailed application of the inclusion and exclusion criteria, often requiring case-by-case judgment, an evidence base of empirical studies was established as a basis for the empirical review. With additions from the update literature search (2005-Jan 2007) an evidence base of 101 empirical studies was established (see Table 3).

Quality Assessment

There is no basis for formally addressing the quality of the studies included, and there is no previous literature in this area covering quality assessment. In one previous literature review (Dolan *et al,* 2005) this fact is acknowledged, and the authors have simply drawn attention to the characteristics of the study design. The current review has gone further and made a

simple inference on the quality of the methods applied (as table below). This allows a simple categorical approach to be used (high, moderate, low) against the methods reported in the empirical studies. This judgment against the quality of the methods rests on the premise that typically the methods associated with a 'high' category are regarded as preferable to those for the 'medium' category, and the subsequent 'low' category (regarded as a weak assumption). These assumptions are regarded as crude, but relatively non-contentious, and have been arrived at following discussion with others engaged in the area of empirical research in health economics (health state valuation, preference elicitation). The literature in this area is reported to be simple, undeveloped, and often of a methodological nature (Sassi *et al* 2001, Schwappach *et al* 2002, Dolan *et al* 2005), and this simple assessment of methods allows further consideration of these issues.

Table 2. Categories used to consider the quality of methods used in studies:

Category/Assessment of the quality of methods applied:	Characterisitics:
High	Random sample used Face-to-face interview methods used
Moderate	Convenience sample AND face-to-face interview methods OR Random sample used AND postal/self-complete format
Low	Convenience sample AND non-interview methods *Any study where methods and data are not transparent

*Note: It is also taken that for studies to be in a high or moderate category they need to present the methods used and the data recorded in a clear and transparent manner. However, to establish such a requirement it is often necessary for a multiple reviewer methodology to form a robust judgment, therefore this issue has not been explicitly addressed as part of the current review.

4.5 Findings: Narrative review of the literature

4.5.1 Summary of the literature

Table 3 presents a summary picture of the literature identified and reviewed to inform on the empirical evidence surrounding potential social values and priority setting criteria. The table reports summary detail on 101 studies. It gives a summary of the characteristics of the empirical studies, including the simple judgment made here against the quality of the methods applied in the studies (low, moderate, or high quality methods).

The literature is of an international nature (whilst accepting that the literature search was restricted to English language materials), with over 15 different countries represented in the

literature identified. A large number of studies are from the UK (31), USA (18), and Scandinavian countries (23). Australia (7) and Canada (4) provide multiple studies, but in a number of instances the countries contributing to the literature have only 1 to 2 empirical studies published (in the literature identified). The vast majority of the studies (over 90%) have been published since 1995. In 74 of the studies reported here, the study design is judged to be of an experimental nature (at least in part) i.e. the design includes the presentation of tasks to respondents which involve some form of choice – opportunity cost – between competing alternatives. In 48 studies the study design is judged (in some substantive part) to be 'attitudinal', with data on attitudes and opinions reported i.e. the study design involves the collection of responses around the attitudes of individuals to specific questions, issues, and/or tasks presented (e.g. categorical responses against level of agreement or disagreement with presented statements). In 21 studies there is a mix of experimental and attitudinal data, or a mix of empirical and methodological (theoretical and/or analytical) considerations.

In 15 studies (15%) sample sizes used have been 30 people or less, although in some instances these studies have been of a more detailed qualitative design, or have reported findings from more than one sample.

The simple qualitative categorisation of studies, against the methods applied, provides some indication of the 'quality' of the evidence base identified. In at least 69 of the 101 studies the sample used is a convenience sample. In at least 64 of the studies the design used a self-complete format (e.g. postal questionnaire, self-complete questionnaire). Whilst the consideration of the methods applied is simple, it indicates that in over 50 (50%) of the studies the methods used would indicate a low quality threshold (methods). In only 14 (14%) studies can the methods applied be used to indicate a high quality threshold (methods). It is important to note that in a number of studies using either 'low' or 'moderate' threshold methods there may be a high 'quality' level with respect to the rigour applied to the rationale for the study and the development and exploration of findings. Conversely, in those studies presenting with high quality methods (interviews in a random sample) this may not be a true reflection of the overall quality of the study, with such studies often using simple designs and failing to address issues that may be important in the collection of data and the interpretation of findings in a health policy context.

For example, the UK study by Bowling (1996) is recorded here as having high quality methods, yet its design is simple and the findings from this study are not easily transferable to health policy decisions. The study sought the views of adults on priority setting over

health services, using a simple grouped ranking of 12 specific health care descriptions e.g. "treatments for children with life-threatening illness" or "a health screening and education programme which could prevent a large number of people needing life saving operations in the future (for example, screening for cancers)". A related question on how to allocate a given sum of money led to responses that were contradictory to the answers from attitudinal questions presented.

Where some studies are regarded as having a lower level of 'quality of methods' they may have a greater policy relevance, and a more useful set of findings. For example, studies by Nord (e.g. Nord *et al,* 1995), Ubel (e.g. Ubel *et al,* 1999), and Skitka & Tetlock (1992) offer useful policy insights, and rich discussion, although using convenience samples and self-complete questionnaire formats in some instances.

A large number of studies can be interpreted as predominantly 'methods studies', with accompanying empirical data. The presentation of these studies would appear to be to primarily inform the methods for future studies of a more policy-orientated design, rather than to make policy propositions themselves (e.g. Anderson & Lytkens 1999, Baron & Ubel 2001, Dolan & Robinson 2001). A large number of studies can be regarded as being 'context' specific, with specific descriptions of health conditions and/or health states (e.g. Zweibel *et al* 1993, Choudry *et al* 1997, Green *et al* 2001), with a lot of these studies related to organ transplant or renal dialysis health care scenarios (Ubel & Lowenstein 1995, 1996, Varekamp 1998, Abellan-Perpinan and Pinto-Prades 1999, Ratcliffe 2000, Browning & Thomas 2001).

The review of the literature presented here is of a summary nature, and is set out to highlight a number of general areas in the literature that are commonly discussed and addressed, and to allow the diverse literature to be drawn together. The review does not seek to provide summary detail on all of the studies reviewed (although a large number are used as examples), but does seek to present a summary of the key factors present in the empirical literature. In many cases studies are able to inform on more than one of the sections presented.

Greater attention is placed here on considering the evidence that suggests society may be prepared to trade-off efficiency gains (health gains) against other distributive concerns. Thereafter, such distributive concerns, and the empirical evidence around them, are introduced.

Table 3: Summary Characteristics of the Empirical Studies Included in the Review

Authors	Pub. Year	Approach*	Sample Size	Country of Study	Design*	Sample	QA Methods
Abellan-Perpinan & Prades	1999	EXP	78	Spain	SC	CONV (students)	Low
Anderson & Lytkens	1999	EXP	225	Sweden	SC	CONV (students)	Low
Annand & Wailoo	2000	ATT	118	UK	SC	RAND (public)	Moderate
Baltussen *et al*	2006	EXP	30	Ghana	SC	CONV (decision-makers)	Low
Baron & Ubel	2001	EXP	70	USA	SC (online)	CONV (public)	Low
Battista *et al*	1995	ATT	55	Canada	SC	CONV (mixed)	Low
Beach *et al*	2003	EXP	781	USA	SC	CONV (public)	Low
Bjork & Rosen	1993	EXP / ATT	60	Sweden	INT	CONV (politicians)	Moderate
Bleichrodt *et al*	2005	EXP	65 / 179	Netherlands	INT/SC	CONV (students/public)	Moderate
Block *et al*	2001	ATT	122	Mexico/South America	INT/GRP	CONV (public)	Moderate
Bowling	1996	ATT	2005	UK	INT	RAND (public)	High
Bowling *et al*	2002	ATT	337/242	UK	INT	RAND (public)	High
Browning & Thomas	2001	EXP / ATT	238	Australia	SC	RAND/CONV (mixed)	Moderate
Bryan *et al*	2002	EXP	909	UK	INT	RAND (public)	High
Charny *et al*	1989	EXP	719	UK	INT	RAND (public)	High
Choudry *et al*	1997	EXP / ATT	80	Canada	SC/PQ	CONV (Healthcare)	Low
Cookson & Dolan	1999	EXP / ATT	60	UK	SC /GRP INT	RAND (public)	High/Moderate
Cropper *et al*	1994	EXP	1000 / 564 / 1000	USA	TEL INT (SURV)	RAND (public)	High/Moderate
Dicker & Armstrong	1995	ATT	16	UK	INT	CONV (public/patients)	Moderate
Dolan & Cookson	2000	EXP / ATT	60	UK	GRP INT (focus group)	RAND (public)	High/Moderate
Dolan & Green	1998	EXP	28	UK	INT	CONV (public)	Moderate
Dolan & Robinson	2001	EXP	71	UK	SC (GRP/INT)	CONV (students)	Low
Dolan & Shaw	2003	ATT	23	UK	SC (GRP)	PURPOSIVE (public)	Low
Dolan & Shaw	2004	ATT	23	UK	INT (GRP)	PURPOSIVE (public)	Moderate
Dolan & Tsuchiya	2005	EXP	128	UK	SC (GRP)	RAND (public)	Moderate
Dolan *et al*	1999	ATT / EXP	60	UK	SC (GRP)	RAND (public)	Moderate
Edwards *et al*	2003	EXP	1101	UK	SC	RAND (public/clinical/dec-makers)	Moderate
Emmelin *et al*	1999	ATT / EXP	451	Sweden	SC/PQ	CONV (politicians)	Low
Farrar *et al*	2000	EXP	130	UK	SC	CONV (healthcare)	Low
Fowler *et al*	1994	ATT	206	USA	INT (TEL)	RAND (public)	High
Furnham *et al*	2000	EXP	107 / 24	UK	SC	CONV (mixed / students)	Low

Authors	Pub. Year	Approach*	Sample Size	Country of Study	Design*	Sample	QA Methods
Gerard	2005	EXP	798	UK	SC	CONV (public)	Low
Green *et al*	2000	ATT	100 / 101	USA	SC	CONV (public / politicians)	Low
Gyrd-Hansen	2004	EXP	3201	Denmark	INT	RAND (public)	High
Gyrd-Hansen & Slothuus	2002	EXP	1895	Denmark	INT	RAND (public)	High
Gyrd-Hansen & Sogaard	2001	EXP	750	Denmark	INT	RAND (public)	High
Holmes	1997	EXP	72	USA	SC	CONV	Low
Johannesson & Gerdtham	1996	EXP	80	Sweden	INT (GRP)	CONV (students)	Moderate
Johannesson & Johansson	1996	EXP	1000 / 2000	Sweden	SC	RAND (public)	Moderate
Johannesson & Johansson	1997	EXP	780	Sweden	INT (TEL)	RAND (public)	High
Johri *et al*	2005	EXP	160	USA	SC	CONV (public)	Low
Kinnunen *et al*	1998	ATT / EXP	1178/682/837/1133	Finland	SC	RAND (public / nursing / Drs) CONV (politicians)	Moderate/Low
Kluge & Tomasson	2002	ATT	5	Canda	INT	CONV (healthcare)	Moderate
Kuder & Roeder	1995	ATT / EXP	46	USA	INT (GRP)	CONV (public)	Moderate
Lees *et al*	2002	ATT	1004 / 357	UK	SC	CONV (public / clinicians)	Low
Lewis & Charny	1989	EXP	721	UK	INT	RAND (public)	High
Lian	2001	ATT	152	Norway	SC	CONV (clinicians)	Low
Linblad *et al*	2002	ATT	22	Sweden	INT	CONV (patients)	Moderate
Lindholm & Rosen	1998	EXP	449	Sweden	SC	CONV (politicians)	Low
Lindholm *et al*	1997	EXP	449	Sweden	SC	CONV (politicians)	Low
Mariotto *et al*	1999	ATT	443 / 189	Italy	INT	RAND (elderly public) / CONV (nurse/aide)	High / Low
Mooney *et al*	1995	EXP / ATT	283	Australia	SC	CONV (decision makers)	Low
Mullet *et al*	1996	ATT	6	Canada	SC	CONV (clinicians)	Low
Myllykangas *et al*	2003	ATT	1178/682/837/1133	Finland	SC	RAND (public / nursing / Drs) CONV (politicians)	Moderate/Low
Neuberger *et al*	1998	EXP / ATT	1000 / 200 / 100	UK	INT / INT / SC	QUOTA (pub / Drs / specialist Dr)	Moderate
Nord	1993	EXP	61 / 25	Norway	SC	CONV (public)	Low
Nord	1993	EXP	10	Norway	INT	CONV (healthcare)	Moderate
Nord	1995	EXP	14 / 53 / 52	Norway	SC	CONV	Low
Nord	1995	EXP	8 seminars (10-13)	Norway	SC	CONV	Low
Nord *et al*	1995	ATT	551	Australia	SC	CONV (partial quota sampling)	Low
Nord *et al*	1995	EXP / ATT	119	Australia	INT	CONV (public)	Moderate
Nord *et al*	1996	ATT / EXP	551/44/42	Australia	SC / INT / INT	CONV (public)	Low / Moderate

Authors	Pub. Year	Approach*	Sample Size	Country of Study	Design*	Sample	QA Methods
Oddsson	2003	EXP / ATT	913	Iceland	SC	RAND (4 of 7 groups) - Phys/Pols/public/public	Moderate
Oliver	2004	EXP	25	UK	SC	CONV (university)	Low
Olsen	1994	EXP	134	Norway	SC	CONV (students/clinicians)	Low
Pernerger *et al*	2002	EXP	1170	Switzerland	SC	Pop'n of physicians	Moderate
Ratcliffe	2000	EXP	303	UK	SC	CONV (mixed)	Low
Roberts *et al*	1999	EXP	91	UK	INT	CONV (public)	Moderate
Rodriquez-Miguez *et al*	2000	EXP	61	Spain	SC / GRP	CONV (students)	Low
Rodriquez-Miguez *et al*	2002	EXP	45	Spain	SC / GRP	CONV (students)	Low
Rosen & Karlberg	2002	ATT	1194 / 427	Sweden	SC	RAND (public / decision-makers)	Moderate
Ryynanen *et al*	1999	ATT	1156/667/803/1096	Finland	SC	RAND (pub / nursing / physicians); CONV (pol)	Moderate/Low
Schwappach	2002	EXP	127	Germany	SC	CONV (public)	Low
Schwappach	2003	EXP	154	Germany	SC (online)	CONV (students)	Low
Schwappach	2005	EXP	1253	Germany	SC (online)	CONV (public)	Low
Schwappach & Strasmann	2006	EXP	716	Germany	SC (online)	CONV (public)	Low
Shmueli	2000	ATT	2030	Israel	INT	RAND (public)	High
Shmueli	1999	EXP	2006	Israel	INT	RAND (public)	High
Skitka & Tetlock	1992	EXP / ATT	198 / 37	USA	SC	CONV (students)	Low
Stolk *et al*	2005	EXP	41	Netherlands	SC	CONV (students/healthcare)	Low
Stronks *et al*	1997	ATT / EXP	45	Netherlands	INT (GRP)	CONV (mixed)	Moderate
Tappenden *et al*	2006	EXP	37	UK	SC	CONV (decision-makers)	Low
Tsuchiya *et al*	2003	EXP / ATT	140	UK	INT	RAND (purposive)	High
Tsuchiya *et al*	2005	ATT	87	UK	SC	CONV (public)	Low
Ubel	1999	EXP	479	USA	SC	CONV (public)	Low
Ubel and Lowenstein	1995	EXP	138	USA	SC	CONV (public)	Low
Ubel and Lowenstein	1996	EXP	169	USA	SC	CONV (public)	Low
Ubel *et al*	1996	EXP	568 / 74 / 73	USA	SC	CONV (public/ethicists/dec-makers)	Low
Ubel *et al*	1998	EXP	289	USA	SC	CONV (public)	Low
Ubel *et al*	1999	EXP	241 / 66 / 74	USA	SC	CONV (public)	Low
Ubel *et al*	2001	EXP/ATT	408	USA	SC	CONV (public)	Low
Ubel *et al*	2001	EXP	615 / 68	USA	SC	CONV (public)	Low

Authors	Pub. Year	Approach*	Sample Size	Country of Study	Design*	Sample	QA Methods
van Busschbach *et al*	1993	EXP	30 / 47	Netherlands	INT	CONV (students / public)	Moderate
Varekamp *et al*	1998	ATT	31	Netherlands	INT	CONV (healthcare)	Moderate
Wailoo & Annand	2005	ATT	118	UK	SC/PQ	RAND (public)	Moderate
Wilmot *et al*	2004	EXP/ATT	22	UK	GRP/INT	CONV (public)	Moderate
Williams	1988	ATT	N/A	UK	SC	CONV (mixed)	Low
Wiseman *et al*	2003	ATT	373	Australia	SC	CONV (public/patient)	Low
Wiseman	2005	ATT/EXP	373 / 43	Australia	SC	CONV (public/patient/healthcare)	Low
Woolhead *et al*	2002	ATT	25	UK	INT	Selected patients	Moderate
Zweibel *et al*	1993	EXP / ATT	505	USA	INT (TEL)	RAND (public)	High

* **KEY:**
Approach: EXP = experimental, ATT = Attitudinal
Design: SC = self complete, PQ = postal questionnaire, INT = interview, GRP = group, TEL = telephone

4.5.2 Is the maximisation of health (health gain) a valid representation of the social value of health interventions? Empirical evidence on health maximisation.

In health economics research, and economic evaluation in particular, it has become standard practice to seek to maximise health gain inside the budget constraint (Oliver 2004). In the conduct of economic evaluation (cost utility analysis) it is now broadly accepted, especially in the UK, that this generally entails maximising QALYs with available resources. This position in some circumstances is questionable, particularly when the findings from economic evaluations are used to support recommendations in practical decision-making settings at a societal level, such as those in the UK NHS (e.g. NICE).

Considering the health maximising criterion at a group or societal level, the literature presents a general, but clear, picture that people (respondents in studies) are willing to sacrifice gains in health outcomes (e.g. life-expectancy, life-years, QALYs) in order to give priority to treatment groups that are not able to benefit greatly from health care, but whom nevertheless are able to benefit in some way.

The findings presented in the studies reviewed strongly suggest that the message from the literature is of a willingness to make some form of sacrifice with respect to efficiency in order to distribute resources according to other criteria i.e. respondents chose alternatives that are clearly not providing the greatest amount of health gain, and are not the most efficient use of resources (and in some cases it is clearly a very inefficient use of resources). These general observations are drawn from a wide range of studies, that vary in design, sample size, sample characteristics, and context (e.g. Nord 1993, 1995, Ubel 1996, Ubel *et al* 1998, Ubel *et al* 1999, Ubel & Lowenstein 1996, Ubel *et al* 1999, Cookson & Dolan 1999, Dolan & Cookson 2000, Abellan-Perpinan & Pinto-Prades 1999, Block *et al* 2001, Pernerger *et al* 2002, Lindholm 1998, Schwappach 2003, Lindholm *et al* 1997, Johannesson & Gerdtham 1996, Shmueli 1999, Charny *et al* 1989, Gyrd-Hansen 2004, Oddsson 2003, Choudry *et al* 1997, Beach *et al* 2003, Edwards *et al* 2003, Anderson & Lytkens 1999).

Nord and colleagues have published a number of studies examining preferences for health gain over other relevant decision making criteria (e.g. Nord 1993, Nord *et al* 1995). These studies have tended to explore the importance of severity of starting health state, and the lesser potential for health improvement, as decision-making criteria in priority setting decisions (see relevant discussion below). Nord and colleagues present findings from empirical studies as evidence against the use of a prominent/dominant health maximising

approach, which may put those who are in a severe health state and who are unable to show large health improvements at a disadvantage. Nord (1993) in a simple experimental study undertaken in Norway, and Nord *et al* (1995, 1996) in larger samples in Australian studies, have reported preferences which do not support a health maximising approach. In these studies many respondents have preferred to allocate resources to groups who were not able to make large health improvements, when other potential allocations of the same resources were able to provide a greater overall health gain.

Abellan-Perpinan & Pinto-Prades (1999), drawing on the work of Nord and colleagues, present findings from a survey in a sample of students in Spain. Respondents were asked to make a choice between different patient level scenarios, where funding had to be allocated across patients described in a differential manner i.e. different current and future health states. They report that very few respondents followed a health maximising objective when making choices. Respondents indicated a preference to provide an equal allocation of funding, regardless of the differential health gain potentials between patients.

Ubel and colleagues have reported, in a number of studies (USA), respondents providing preferences that are not consistent with a health maximising objective (Ubel and Lowenstein 1995, Ubel 1996, Ubel *et al* 1999, Ubel *et al* 2000). Ubel *et al* (1996) report preferences to support a less effective screening test over a more effective one, with respondents prepared to forego health gain in order to provide a wider coverage of the screening test. This study, in the USA, used three convenience samples (568 public, 73 decision-makers, and 74 ethicists), via a self-complete survey. In the survey respondents were asked to choose between providing a screening test that would be provided for all of the population, saving 1,000 lives, or a more expensive test that was more effective, saving 1,100 lives, but it would be provided to half of the population. In the samples used 53% of the public, 53% of the ethicists, and 41% of decision makers favoured making the screening test available to the whole population, the less effective test.

Ubel *et al* (2000) extend their earlier research, from the above 1996 study, by exploring the context of the stated preference to trade-off between health gain and distributional considerations, when making a choice between a less effective and more effective screening test. They present a similar choice scenario in a convenience sample of the USA public (n=495) and physicians (n=1,294), again using a simple self-complete format. However, in this follow-up study they present a scenario where the less effective test is not available to "all" of the population. Three question formats were presented (1) as per the 1996 study with 100% coverage versus 50%, (2) 50% coverage versus 25% coverage, and (3) 90%

coverage for the less effective test, versus 40% coverage for the more effective test. In all scenarios the less effective test (with wider coverage) saved 1,000 lives, and the more effective test saved 1,100 lives. The study provided similar findings in the scenario used for the 1996 survey, scenario (1), but showed a much reduced number of respondents willing to trade-off health gains against distributional considerations (coverage). Only 27% to 28% of respondents in this follow-up study chose the less effective test, when it was not available to the whole population.

Pernerger *et al* (2002) present results supporting the findings from Ubel *et al* (1996, 2000). They report preferences from a postal survey in a sample of Swiss physicians, where 75% of respondents preferred universal access to a less effective screening test over a scenario that achieved greater health gain for society by offering a more effective test to fewer people.

Ubel *et al* (1999) examine the support for health gain maximisation in a study using the person trade-off approach to elicit preferences against three choice scenarios, each presenting a trade-off between maximising health versus other considerations. A sample of the USA general public was used, analysis was restricted to data from 241 useable responses. In the analysis respondents indicated a willingness to take other considerations into account, and not to follow a health maximising approach. There are a number of contextual factors in the study, as with many of the empirical studies in this area, and the study provides insight to other issues discussed below, but also provides evidence that the health maximising objective may have limited support.

Ubel and Lowenstein (1995), in a study examining preferences in the context of liver transplantation, provide support for the argument against health maximisation. In this USA study the sample generally had a preference which was not consistent with the allocation of resources to the group of patients that had the greatest expected health improvement from treatment. Respondents expressed a preference to treat patients that were not able to provide the greatest expected health gains; 33% of respondents chose to distribute organs equally between two prognostic groups, one with a better prognostic outcome where greater health benefits were expected. Only 22% of respondents had a preference to treat the better prognostic group, with the majority of these stating the choice was based on maximising survival opportunities. In many cases where a more equal distribution of resources was preferred, respondents justified their decisions on the grounds that everyone deserved a chance at transplant. As with many studies in this area (empirical ethics) the study is very context specific. Whilst clearly demonstrating that prognosis is an important factor in how respondents chose to distribute transplant organs, findings may not be

generalisable to other priority-setting contexts. However, such studies may indicate that the public place a high value on distributional concerns, even if supporting such concerns is at the expense of health gains foregone.

Charny *et al* (1989) present evidence from a UK study, involving interviews in a large sample (n=719), that the use of health gain as a strong decision making criterion is limited. The study used a series of context-specific questions, involving specific health conditions (e.g. heart attack) and linked to factors such as age, social class, employment status. A general finding from the study was that a life-year may not be of equal worth to all, and that social values may involve distributional considerations as well as health gain.

Dolan (1998) in a UK study which considers the severity of the starting health state, reports a preference to trade-off potentially greater health gain in one group to provide a lesser health gain to a different (more severely affected) treatment group.

Cookson & Dolan (1999), and Dolan & Cookson (2000) (in a related publication) in a UK study based on the use of focus groups to examine preferences of the general public, found that people were willing to make health gain trade-offs between patient groups. The study reports a strong preference for equality of access over the maximisation of health gains/benefits. In this study the qualitative responses indicated that in this sample (n=60) none of five decision-making criteria presented (including health maximisation) had overriding importance, and that health maximisation was not seen to have a dominant role in health care decision-making. However, health maximisation was reported as an important consideration in health care decision-making.

Johannesson and Gerdtham (1996) present findings from an experimental study (Sweden) to consider the trade-off between health maximisation and equality of health (at the expense of health gains). The study is a simple one, exploring methods for investigating the equity-efficiency trade-off. It reports the views of 80 students, with results suggesting a strong preference to trade-off health gain in one group (a better-off group, in terms of remaining lifetime health) to provide a lesser health group to a less well-off group.

Anderson and Lyttkens (1999), in a Swedish study similar to that reported by Johannesson and Gerdtham (1996), report findings from an experimental empirical study that indicate factors other than the maximisation of health gains are important when choosing between different health care scenarios. The respondents in this study (convenience sample of

economics students) indicated that the size of the equity-efficiency trade-off and the level of inequality between different groups were important factors in the health care choices made.

Oddsson (2003) in a survey in Iceland reports a strong egalitarian preference in a number of random samples (general public, physicians, politicians). This postal survey examined trade-offs between treatment effect (health gain) and level of health state severity. Two questionnaire formats were used, and between 60% and 73% of respondents were not prepared to prioritise between those patient groups that could get a greater expected health benefit and those that were expected to get a lesser health benefit (in terms of health gain). Respondents preferred to offer an equal distribution of resources between these groups, and did not demonstrate a preference to maximise health gain from available resources.

Bryan *et al* (2002) report findings from an experimental study examining the preferences of a random sample of the UK general public. The study used a discrete-choice experiment, and it examined trade-offs across attributes, in a policy context. The study allowed the examination of responses against a health maximising objective. The authors of this study report that in only 8.5% of respondents could preferences be considered consistent with a QALY maximising approach.

In a number of attitudinal studies a common finding is that health gains/improvements are one of the important decision making criteria, but not the only important factor (e.g. Ryyannen *et al* 1999, Myllkangas *et al* 2003, Kinnunen *et al* 1998, Battista *et al* 1995, UK study by Bowling *et al* 2002). Many of these studies are relatively simple, and ask respondents to provide a rating of individual statements, or criteria, using Likert-type response scales. Findings from the Finnish Kuopio Study have been reported in numerous papers (Ryyannen *et al* 1999, Myllkangas *et al* 2003, Kinnunen *et al* 1998). The study was a large postal survey eliciting views of a random sample of the general public, politicians, doctors and nurses. It set out to explore the attitudes and values of respondents regarding health care priorities, and priority setting preferences. The survey used a listing of potential prioritisation criteria, and patient based case descriptions as examples of where priority setting may be necessary. Many of the issues presented were related to characteristics of patient groups (e.g. child patient, rich patient, self-induced illness). The findings from the study indicate that a range of distributional concerns are important to respondents, and that context may be an important factor. Findings across groups were very similar, although politicians are reported to be more likely to be able to set priorities. The important concerns contained factors relevant to health maximisation, but also other issues where respondents indicated that trade-offs would be made against efficiency (health gain).

For health maximisation:

From the literature identified it is possible to draw support for a health maximising decision making perspective (e.g. Choudry *et al* 1997, Beach *et al* 2003). Given the complex nature of the empirical evidence reported it is often possible to interpret findings from studies in a number of different ways. In such a way it is possible to detect both strong and weak messages in favour of health maximisation in the empirical literature. However, the results from the current review suggest that the evidence base is heavily weighted against the dominant use of the health maximisation criterion.

Choudry *et al* (1997), in a Canadian study amongst health officials (potential decision makers), report some support for a health maximising approach. This study presented what are described as 'distributional effects scenarios', eliciting preferences between providing large health gains to a few people versus small health gains to many people. A large number of respondents (55%) indicated support for providing large benefits to fewer patients (a health maximising approach). However, 23% of respondents were unable to decide or had no preference over the distribution of health care between groups, even though there were notable differences in some of the options presented in the scenarios (often the scenarios involved the same level of overall health gain, but distributed differently).

Beach *et al* (2003) in a self-complete survey in a convenience sample of the general public (USA, n=781), found that a large number of people had a preference to provide a more effective cancer screening strategy (a health maximising strategy) over a less effective screening strategy available to a larger number of people. The finding that 55% of respondents chose the health maximising option offers some support for the importance of health maximisation, but 41.5% preferred the less effective, but more available, test (3.5% chose no screening at all), showing a strong preference for equality of access. In four of the six scenarios used in the study there were approximately 70% of respondents who preferred the health maximising approach. However, the study had a very context-specific design, with the focus on cancer screening, and questions presented in the context of health insurance coverage in a USA population (where cost-effectiveness analysis and efficiency arguments are less prominent in health policy).

Ratcliffe (2000) reports findings from an experimental study (discrete-choice experiment) where UK respondents indicated a strong preference to select scenarios which offered the greatest health gain (capacity to benefit). This study was a postal survey considering the distribution of transplant organs, and was context specific, with prognosis in liver transplant being a key factor. However, in this study there was still a demonstrated preference to also

offer some level of support to those groups that were not able to benefit greatly from treatment, regardless of the opportunity cost of health gain foregone in other groups with a greater capacity to benefit.

As above, in attitudinal studies the size of possible health gains, and the capacity to benefit are regarded as important decision making criteria, although not regarded as the most important, in many cases (e.g. Myllkangas *et al* 2003, Kinnunen *et al* 1998).

A number of studies have highlighted a possible threshold effect, with preferences possibly dependent on the relative health gains possible in competing groups. For example, where benefits going to one group may be considered to be too small (to make any meaningful difference to the recipient) respondents may prefer to concentrate benefits amongst fewer people instead (Abellan-Perpinan & Pinto-Prades 1999, Olsen 2000, Ubel *et al* 2000, Choudry *et al* 1997, Dolan and Green 1998).

For example, although Abellan-Perpinan & Pinto-Prades (1999) found that only a small number of respondents in their study were acting to maximise health, they found that respondents were more likely to select a health maximising alternative where the benefits going to the group who would have relatively small gains, were very small.

Summary on health maximisation literature:

In summarising the literature in this area, it is important to be aware of the general limitations in much of the empirical literature (e.g. framing of questions, convenience samples, methodological studies) however, given the variety in the studies reported some general messages from the literature may be highlighted. It seems clear that the principle of health maximisation lacks general support as a dominant decision making criterion, at a societal level. The evidence available strongly suggests that people will, in certain contexts, be prepared to trade-off health gains against other decision-making considerations/criteria. The evidence strongly suggests that people are not indifferent concerning the distribution of benefits across health care groups. The evidence suggests that the value of health gain (e.g. QALYs) is not constant, and that simply maximising aggregate health gains across the population may not be the best use of limited resources according to people's preferences.

4.5.3 Is the social value of a health intervention dependent on factors other than health (health gain)?

If health gain is not to be used as a dominant decision making criterion, it is necessary to consider on what basis should priorities be set between different health care groups? The literature reviewed does offer some insight to this difficult question. As discussed above, the current literature strongly suggests that the social value of health interventions is dependent on factors other than health. Below, the evidence base reviewed is discussed against three sub-headings, separating potential priority setting considerations into those that cover the characteristics of the groups receiving health care, the characteristics of the health intervention itself, and other distributive preferences. All of these three areas are interlinked and overlap in terms of the empirical evidence base.

4.5.3.1 Characteristics of the patient group

Age

Age has been one of the most widely discussed issues in the context of health care priority setting. Tsuchiya (1999) has presented a detailed coverage of the treatment of age, presenting different interpretations of 'ageism', which suggest lower priority for older people. Tsuchiya *et al* (2003) have followed up on this earlier work and presented both theoretical and empirical discussion surrounding age, and the potential for age-weighting in health care priority-setting decisions.

One of the most prominent theoretical presentations for age as a priority setting criterion is the 'fair-innings approach' (FIA) presented by Williams (1997). This is a clear argument for the consideration of age in allocating health care resources, although it is concerned with an allocation according to a quality-adjusted life-expectancy, on the basis of an ethical entitlement, and not merely the use of age alone as a basis for discriminating between groups. Williams' presentation of the FIA is largely theoretical, using hypothetical data in the absence of empirical data. However, it has given rise to a number of other empirical investigations of age as a priority setting criterion, and has heightened awareness of age in the debate around priority setting.

Whilst there are a number of studies that can be cited to support a possible age-related priority setting preference (e.g. Bowling 1996, Johannesson & Johansson 1996, Lewis & Charny 1989, Busschbach *et al* 1993, Charny *et al* 1989, Varekamp *et al* 1998, Browning & Thomas 2001, Tsuchiya *et al* 2003, Schwappach 2003, Johri *et al* 2005, Cropper *et al* 1994) it is important to consider the context in which these studies are framed, and the possible

explanations for the response patterns reported. It is possible that the perspective and the framing of survey questions may play a major part in studies. For example Browning & Thomas (2001) report age as important in the context of allocating donor organs, and this preference may not be generalisable to other policy areas. It may be that the preferences reported in many of these studies are true preferences to favour certain age groups over others, but it may also be the case that the preferences are in some way related to the experimental designs used. In many studies it is often difficult to interpret the findings, in either an experimental or policy context. For example, where Cropper *et al* (1994) report a preference for younger versus older age groups, concluding that age does in fact matter in a decision-making context, it is important to note that the study is presented as a trade-off between present versus future lives. In a series of questions the study uses age bands of 20-year olds versus 60-year olds, examining preferences using an equivalence of numbers question. In these questions the young are preferred to the old, but this is the case even where the choice was between saving 6,000 lives of 60-year olds versus 200 lives of 20-year olds. A study by Johannesson and Johannsson (1996), also estimates such trade-offs between stated age groups (equivalence methods), and it suggests that treating one person at age 30-years is equivalent to treatment of 35 people aged 70-years. Such studies highlight the need for the reader to be familiar with the methods used in these studies, as well as their headline messages.

There is empirical support for a preference against the use of age as a priority setting criterion (e.g. Zweibel *et al* 1993, Kuder and Roeder 1995, Nord *et al* 1995). Kuder and Roeder (1995) have reported findings from a mixed-methods study to suggest that respondents did not support age as a priority setting criterion, but when asked to make a choice they did so, indicating a preference for young versus old. It would seem that when faced with a scenario where treatment is available to either a child or an elderly person, and it is stressed that only one of these persons can be treated, respondents indicate a preference to treat young people over old people. Yet, this may not be a true reflection of how respondents would wish to allocate scarce health care resources.

Attitudinal studies, where there is no 'opportunity cost' involved in the questions asked, indicate a preference for a positive discrimination in favour of the young, but also indicate that respondents are not willing to discriminate against the older age groups (e.g. Ryynnanen *et al* 1999, 2000).

The majority of empirical studies examining age appear to be supportive of giving greater weight to younger people. However, the many studies examining age are often very simple

designs. They are often primarily methodological studies, and frequently present quite specific scenarios to respondents (e.g. Bowling 1996). Furthermore, almost all of the studies examining age as a priority setting consideration do not control for other potential confounding factors that may influence preferences (e.g. Lewis & Charny, 1989). One exception is the study by Tsuchiya *et al* (2003), where, using a UK general population sample, the study attempted to control for the size of the benefit. In some studies there are inconsistencies in the observed pattern of preference for young versus old. For example, where respondents are asked to make a choice between those aged 2-years versus 8-years of age they indicate a preference for the older age group (Lewis & Charny, 1989).

There appears to be no clear message from the empirical literature with respect to age as a basis for setting health care priorities. Earlier reviews of the literature in this area (Sassi *et al* 2001, Scwappach *et al* 2002, Dolan *et al* 2005), conclude that evidence may be interpreted as being supportive of giving lesser weight to older people, versus younger people, but that this empirical evidence is simple, often using crude methods, and such findings may be related to issues such as greater capacity to benefit (larger health gains) and may not be a direct reflection of a preference against a specific age group. Indications from health policy forums report that the general public are not prepared to use age directly as a decision making criterion, although if it is related to other factors (such as the clinical benefits available to age groups) it may be a relevant consideration (NICE, Citizens Council 2003, *Report on Age*). The NICE Citizens Council (see discussion in 4.5.4) does not recommend that NICE should be more generous in its judgments of cost effectiveness merely because of age (Rawlins & Culyer, 2004).

At the present time it is difficult to interpret the literature on age. It is clear that further empirical research is needed to disentangle age from other confounding factors. At the present time it is not possible to say how a given benefit should or could be weighted [allocated] across different age groups. It is also clear that in the UK NHS (e.g. NICE) there is no explicit basis upon which to use age as a priority setting criterion (unless it is an indicator of either risk of benefit).

Social role

The conceptual argument surrounding 'social role' is that some individuals (or groups) may be valued more to society than others on the basis of their role in society (at that time). Social role is one of the patient characteristics that could be linked to the 'age' issue, as social role and age are linked via the life-cycle stages at which social role may be defined e.g. caring for young children, caring for elderly dependents, employment and productivity

within society. It may be that the empirical studies reporting age and/or social role get these two issues 'tangled' up in some way, and are unable to report preferences against either in a specific way (Sassi *et al,* 2001).

There is limited empirical evidence directly available on this issue. The evidence available suggests that public preferences are opposed to discriminating on the basis of productivity losses, and between individuals or groups on the basis of social role (Schwappach 2002). However, Schwappach reports that "while there is commonly a vast majority of survey participants that refuses to discriminate according to working status, retirement, wealth or poverty of participants, or their position in society, more people are prepared to prioritise in favour of patients with dependants or other social responsibilities" (p213).

The evidence is suggestive of giving some priority to those groups with dependants (Williams 1988, Charny *et al* 1996, Browning & Thomas 2001, Neuberger *et al* 1998, Dolan *et al* 1999, Holmes 1997, Olsen *et al* 1998). For example, Olsen *et al* (1998) asked respondents to indicate preferences for treating groups of employed people, parents, people taking care of others, compared to reference groups (same illness and will gain as much from treatment). The study found a preference to give priority to parents (47% or respondents), or people taking care of others (45% of respondents). For employment status, only 27% of respondents gave priority to groups who were in the paid workforce. Williams (1988), in one of the earliest empirical reports, from a simple self-complete survey, argues that the public preference may be to discriminate in favour of those with children over people without children. However, the literature on this issue is not very clear, and other studies have reported a preference not to take responsibility for dependents into account (e.g. Edwards *et al,* 2003).

There is a theoretical literature presenting debate surrounding the inclusion of productivity losses in cost-effectiveness analysis (e.g. Brouwer *et al* 1997, Weinstein *et al* 1997), but again there is no clear concensus, and in practice it is left to the decision-making perspective employed and the interpretation of the analyst.

NICE have stated that differential productivity "ought not to be used to disadvantage people who are not in regular paid employment, including children and those who are retired" (Rawlins & Culyer 2004, p226). The UK NICE Citizens Council (2003) does not recommend that NICE should use social role in its judgements on cost effectiveness.

Severity

The conceptual argument here is that the societal value of a health improvement of a given size is greater when given to a patient group that has a severe (more severe) health condition (Nord, 2005). There is some empirical support for favouring those people that are most severely affected by their health condition, or who have the most severe illnesses, even if they are unable to gain as much as others in terms of health improvement (e.g. Nord 1993, Nord *et al* 1995). Such empirical evidence has been highlighted (above) to draw attention to the lack of support for the efficiency arguments surrounding the maximisation of health gains. Several studies have suggested that focusing on health gains, and therefore the capacity to benefit, does not reflect the social preference to treat severely ill patients (e.g. Gyrd-Hansen 2004, Oliver 2004).

Empirical studies have asked respondents to choose between patient groups who will gain equally from treatment, but differ in the severity of their current (starting) health condition. They have also asked respondents to choose between groups where health gains are different across groups that are at differing levels of pre-treatment and post-treatment disease severity. In such studies many respondents want to give priority to the severely ill patient groups, even when they have much lower health gains than other patients in the choices offered (Nord 1993, Nord *et al* 1995, Ubel *et al* 1998, Ubel 1999, Oddsson 2003).

Other studies, often exploring a number of factors other than severity of health, have found support for severity of health as an influential factor in setting health care priorities (e.g. Ubel & Lowenstein 1995, 1996a, 1996b, Cookson & Dolan 1999, Shmueli 1999, Gyrd-Hansen 2004, Edwards *et al* 2003, Oliver 2004, Wiseman 2005).

Cookson and Dolan (1999), and Dolan and Cookson (2000), report that findings from deliberative group interviews, and self-complete questionnaires, suggest that severity of health was one of a number of important considerations. Oliver (2004) considered prioritisation of life-saving health care between groups who differed according to the presence of disability, finding some support that those in severe health states should receive at least equal priority regardless of capacity to benefit. Edwards *et al* (2003) report that the general public include the level of disability (in patient groups), in the listing of the most influential factors that should be considered in setting priorities choices (in determining waiting times). Gyrd-Hansen (2004), in a sample of the Danish general public, reports support for favouring those groups with a more severe health condition.

Shmueli (1999), in an interview study in the Israeli general public (n=2,006), found that respondents were prepared to give more weight to those patients with severe pre-treatment health conditions, regardless of differential outcomes. The study reports that people felt it was not enough to compare the outcomes of treatment. This study used very specific patient descriptions to elicit preferences. For example, one question was framed using two patients in the emergency room. Both patients are 50-year old male road accident patients, both are married with children. Respondents are told that one has a high chance of dying, if not treated immediately, and may die even if treated. Whilst the other will survive, but if not treated immediately may be paralised in both legs. The question posed in this instance was that if only one patient could be treated at a time, which of the two should be treated first? The study used variants of the above format, with different life-expectancy for the first patient (likely to die). In all variants there was a large proportion of respondents who favored treating the patient who was likely to die if not treated immediately, even where the patient was only expected to live for one month if treated. The number of respondents favouring treatment of the first patient, in the more severe pre-treatment health state, increased as the life-expectancy after treatment was increased (to 1-year, and 5-years).

In some cases (e.g. Richardson & McKie 2005, Cookson and Dolan 1999) the severity of the illness or condition is often referred to as a variant of the 'rule of rescue' (Haddorn 1991). This is a questionable association, dependant on the interpretation given to the rule of rescue. The current author regards the rule of rescue as being about identifiable individuals and the avoidable death (or suffering) of these identifiable individuals. At a policy level, there may be an element of identification in groups of patients, especially where specific cases are presented in the media. However, policy decisions (e.g. UK NICE) are at a wider policy level and not directly related to the classic presentation of the 'rule of rescue'.

The empirical literature indicates a level of support for giving at least equal priority to those groups in severe health states, regardless of a lower capacity to benefit, and in some cases giving greater priority to such groups. The studies in this area may be capturing a number of different concerns (e.g. pre-treatment health, post treatment health, health gain, patient characteristics), and are open to some ambiguity in interpretation of preferences in a decision-making context. But overall there seems support for severity of health condition to be given some place in priority setting decisions. Dolan *et al* (2005) conclude that "in general terms, and across a range of decision contexts, the empirical evidence currently available suggests that people are willing to sacrifice quality of life gains in order to give priority to the most severely ill" (p 205).

Health related behaviour / Lifestyle

Issues of health and lifestyle, health-related behaviour, and desert, have been addressed in the empirical studies reviewed. From the evidence available, health-related lifestyle (e.g. smoking behaviour, alcohol consumption), and the related 'cause' of ill-health arguments (e.g. drug addiction, obesity), have some impact on preferences of those questioned on priority-setting, but there is not a great deal of evidence available and that available is unclear.

A number of studies report a level of support amongst respondents to give a lower priority to those who are considered to be in some way responsible for their ill health (Bowling 1996, Williams 1988, Lewis & Charny 1989, Nord *et al* 1995, Stronks *et al* 1997, Ryynanen *et al* 1999, Ratcliffe 2000, Neuberger *et al* 1998, Skitka & Tetlock 1992). There are also studies where respondents are strongly opposed to discriminating on the basis of health related behaviour such as smoking, alcoholism (e.g. Ubel *et al* 2001, Edwards *et al* 2003).

Dolan *et al* (1999), following a series of group interviews, regard the issue of lifestyle to be a contentious area, with competing and conflicting messages from respondents. As with other areas of the empirical literature, as well as the potential for framing and design issues in the studies, the preference studies in this area are also open to possible confounding with respect to health gain, or efficiency, arguments. It may be that those who are regarded as having a less healthy lifestyle are less likely to have a health gain as great as others with a healthier lifestyle (related to risk of complications, or general prognosis) (Schwappach, 2003).

Social (socio-economic) status

Distinct from social role, the issue of socio-economic status, or social class, has been considered in a number of studies. This has been in the form of preferences for the allocation of health care (resources or outcomes) for the richer versus the poorer groups in society (Dolan *et al* 1999), or against health profiles against different categories of socio-economic status (Mooney *et al* 1995, Wiseman 2005), or levels of education (Dolan *et al,* 1999). Mooney *et al* (1995), and Wiseman (2005) report a strong preference to favour the lower social class groupings, whilst Anand and Wailoo (2000), and Block *et al* (2001) do not find a preference according to social class groupings. Charny *et al* (1989) do report preferences by occupation type, but there was no clear message, with some respondents preferring high social status occupations, and others preferring lower social status occupations (e.g. lorry driver, versus a teacher). However, Charny *et al* report that the

majority of respondents did not want to discriminate on the basis of occupation. Dolan *et al* (1999) report 23% of respondents suggested a lower priority for the richer groups, and that 8% of respondents were in favour of giving higher priority for the less educated groups in society. Emmelin *et al* (1999) report a small number of respondents (12%) indicate a preference to give priority to those groups regarded as socially disadvantaged.

The literature in this area is undeveloped and unclear. One area of empirical research has been in the area of inequalities, and specifically around the reduction of inequalities, and in these studies (outlined below) socio-economic status has been one of the considerations. It is difficult in such studies to separate out considerations around social status specifically and the issue of inequalities.

Gender & Race (Ethnicity)

The empirical literature around issues of gender and ethnicity is sparse and unclear. A number of studies have considered, or commented on gender, with no indicated support for setting priorities on the basis of gender (e.g. Holmes 1997, Browning & Thomas 2001, Mooney *et al* 1995, Dolan *et al* 1999). Few studies state race/ethnicity as a consideration in priority setting. Ubel *et al* (1996) suggest that race/ethnicity is not supported as a basis for setting priorities.

There is no indication from the current literature that gender or ethnicity have any support as considerations in priority setting debates.

Prior health care consumption

The conceptual argument here is that there may be a different social value attached to health care provided to those who have not previously received significant health care provision compared to those who have already received a significant amount of health care provision (e.g. in the case of organ transplant the category of 're-transplant'). The underlying hypothesis is that society might feel that people are entitled to a certain amount of health care, i.e. to have their lives saved once, and that everybody should get a first chance before others get a second (Schwappach 2002). There is little direct empirical evidence in this area, although a number of studies offer an insight whilst addressing a number of other decision making considerations (e.g. health related behaviour and lifestyle), indicating that prior consumption of health care alone is not a strong, or relevant, criterion when making choices over allocation of health care, and that prognosis and other factors are important contextual inputs (e.g. Ubel & Lowenstein 1995, Ratcliffe 2000).

4.5.3.2 Characteristics of the health intervention / health effect

The 'start-point', 'end-point', 'health improvement' available, 'nature of the health improvement', and 'direction' of health improvement are all factors that can be discussed in terms of the health intervention, as well as being important for discussion of patient characteristics. The 'number' of persons to be treated is also a factor that may be relevant in decision making, and may be related to the 'patient group' and/or the 'health intervention'.

Start-point/pre-treatment health

Considerations at the level of the health intervention are often distinct from the characteristics of patient groups, but they may also overlap in places. The 'starting point' for health care (i.e. pre-treatment health state) is one such overlapping factor i.e. at what point is the health intervention used (mild, moderate or severe illness). This issue is very much related to the arguments presented above on the severity of health, and the potential for giving preference to those in health states regarded as more severe than others. Whilst it may be argued the there may be indifference over equal health improvements given to patient groups at different starting health states (e.g. some mild/moderate versus some severe), the evidence presented above argues that this may not be the case, with preference suggested for the most severely affected groups.

Health 'end-point' after treatment

Arguments over the 'end-point' (the post-treatment health state) after treatment are linked to the discussion and debate surrounding the maximisation of health, subject to available resources. It may be that two patient groups in a similar starting health state may be treated with an equivalent intervention, but the two groups have a different capacity to benefit (e.g. one can be completely cured and returned to full health, while the other can be returned to an improved state which is less than full health). Applying a health maximising objective would favour one group and discriminate against the other, on the basis of capacity to benefit. The empirical evidence around this issue has been discussed in the above section on maximising health gain (e.g. Abellan-Perpinan & Pinto-Prades 1999, Shmueli 1999). However, it is important to be aware that the issue can appear in a number of different conceptual presentations.

Number of people treated

The argument here is that the societal value of a health improvement will increase as the number of people treated increases. Whilst a number of studies offer insight here, they are

initially seeking to explore the acceptability of the maximand of health gain. Studies from Ubel *et al* (1996, 2000) have indicated that preferences are for a wider coverage (in the context of screening), even if it is at the expense of overall health gain. However, the same studies have indicated that the extent of the coverage available (all versus less than all patients) has some impact on the strength of preference. A number of the studies identified have used equivalence of numbers techniques to elicit preferences, however, these studies have not in the first instance set out to explore the 'number of people treated' as a decision making criterion. They have used equivalence numbers to guage the strength of preference against competing groups of patients, on the basis of patient or health intervention characteristics. A useful review of the person trade-off technique, which uses the number of people treated in competing treatment groups to elicit preferences, has been reported by Green (2001), but it is not relevant for the consideration of the specific preferences attached to the coverage of a health intervention.

Often it is the case that the numbers to be treated with a health intervention is of greatest relevance to 'budget impact' considerations, whereby health care funding bodies are exploring the flows of resource use and costs over time. It is not an issue that is relevant in the consideration of the cost effectiveness of a health intervention. On the issue of 'affordability', and budget impact, NICE have stated that they do not take these factors into account when making judgments about the cost effectiveness of health interventions (Rawlins & Culyer, 2004).

Health improvement/gain

Evidence reported above, surrounding health maximisation, provides support for using the level of health improvement, or the magnitude of health gain, as a basis for setting priorities. There is empirical support for the fact that health gain does have a large influence on the priority setting preferences of respondents (e.g. Cookson & Dolan 1999, Roberts *et al* 1997, Bowling 1996, Dolan & Green 1998). However, it has been noted that this is not at the expense of all other distributive considerations. Efficiency, and health improvement, are both related and well supported considerations in the empirical literature reviewed. Efficiency in the context of health policy is of prime importance (e.g. Rawlins & Culyer, 2004), but not as the sole consideration in the decision making process.

Nature of health intervention/improvement

The nature of the health intervention, or health improvement, (i.e. life-saving versus life-enhancing), and the direction of the health improvement (i.e. preventative health care versus

treatment of current health conditions) have received some attention in the empirical literature. The former of these issues (i.e. life saving) has received greater attention in the theoretical and conceptual literature surrounding the provision of health care (e.g. Ubel *et al* 1999). Nord and colleagues have contributed in both theoretical debates (e.g. Nord 1993, Ubet *et al* 2000) and through empirical enquiry (e.g. Nord *et al,* 1995), reporting a theoretical rationale for saving lives, around a broad interpretation of the rule of rescue, and empirical support for the saving of lives over life-enhancing interventions.

Whilst life-saving interventions are considered in a number of the empirical studies identified, it is rare that the effect of the intervention (i.e. saving lives) is isolated within the general scenarios presented, and rarely is any inference drawn directly on the nature of the intervention. Many of the studies examining severity of health, or level of health gain, have some form of 'noise' from the nature of interventions (e.g. Johri *et al,* 2005), and the literature is currently too crude and simple for conclusions to be drawn surrounding the many factors that may be at play when respondents provide preference, or attitudinal, data.

In terms of 'direction of change' (e.g. prevent vs. cure) studies by Ubel *et al* (1998), Shmueli (1999), and Schwappach (2002) have considered this issue. The study by Ubel and colleagues is the most frequently cited in this area. Ubel *et al* (1998) examined preferences over preventative versus curative health care, finding no strong preference either way. This study does report a greater number of people preferring to fund preventive versus curative care, but the differences were not large, and the strongest preference was to give equal importance to the different forms of health care. A number of different choice scenarios were used in the survey and most had other issues present in the scenarios given (e.g. different levels of health improvement between options, different levels of severity and disability). In one of the choices given between prevent and cure the levels of health improvement were the same, and 37% of respondents preferred the prevent option, 21% preferred the cure option, whilst the remainder preferred equal importance. It is important to note that, as with many studies, a convenience sample was used, a self-complete format was used, and the questions were framed in a very specific way using two patient groups who were described as residents in a nursing home. It is therefore difficult to establish the generalisability of the findings, and their policy relevance.

The study by Shumueli (1999) included many contextual factors, including a notion of preference over prolonging life compared to the prevention of severe and permanent disability. The preferences reported were dependent on the relative outcomes (life-expectancy) for the patient groups described, and it is not clear what specific preference

there was over prevention versus treatment. Schwappach (2002a) considers the social value of avoiding a decline in health, comparing preferences for 'upward movements' on the health scale versus avoidance of 'downward movements' in health. Results indicate a preference towards curative (health improving) health care rather than preventative (avoiding decline), with the majority of respondents (69%) preferring the allocation of resources to already diseased patients. But responents also indicated that the number of people involved was an important factor in the decision making process, as they were prepared to trade-off a preference for curative versus protective health care when the number of patients was different i.e. when a larger number of patients could be saved from decline. This study was predominantly a methods study, in a convenience sample (Germany) which was not representative of the wider population. It used the person trade-off approach for elicitation of preferences, and the design of the study raises some concerns in terms of policy relevance.

4.5.3.3 Other general distributive preferences

Health inequalities

Health inequalities across different groups in society are generally regarded as being 'inequitable', and there has been a great deal of research to identify and quantify health inequalities. However, how to address such inequalities remains a challenge. The empirical evidence is indicative of a preference for reducing health inequalities across different groups in society (e.g. Lindholm *et al* 1996, Johannesson & Gerdtham 1996, Lindhom *et al* 1998, Lindholm & Rosen 1998, Emmelin *et al* 1999, Anderson & Lyttkens 1999, Dolan & Robinson 2001, Cuadras-Morato *et al* 2001). Studies have suggested a preference to act on inequalities across groups described according to social class, those being socially disadvantaged, and groups described as poor versus richer. There is no evidence of a preference to discriminate against groups according to age, gender, or ethnicity, to address health inequalities (as discussed above). The evidence is sparse, and experimental, and the inference in favour of reducing inequalities is often from studies exploring the assumption that health (QALY) maximisation is an appropriate decision making objective.

Whilst the evidence indicates a willingness to trade-off efficiency gains (health improvements) to reduce inequalities, the evidence suggests that respondents would not do so at all costs, and that there are thresholds at which the relative gains and losses in respective groups (in choice sets presented) are important (e.g. Lindholm *et al* 1998, Anderson & Lyttkens 1999). There is also evidence that respondents are not prepared to trade-off health gains to address inequalities. Dolan & Robinson (2001) report findings from

two empirical studies, with one suggesting inequality aversion, and another that respondents were inequality neutral.

It would seem that society regards health inequalities as a bad thing, and there is evidence for a reasonable level of support to address them in some instances. But action to reduce inequalities, and the associated opportunity costs (some losers and some gainers), needs careful consideration before respondents are prepared to indicate a strong preference to allocate health resources to address health inequalities directly.

4.5.4 Policy guidance from the National Institute for Health and Clinical Excellence (NICE)

The review of the published literature has also been supplemented by an outline review of the guidance from the NICE technology appraisal programme, together with consideration of the reports from the NICE Citizens Council, and policy documents informing on the NICE appraisal process (e.g. NICE 2005, Culyer & Rawlins 2004, Bryan *et al* 2006, Williams & Bryan 2007). This outline review consisted of a general appraisal of the literature in this area, including published NICE guidance, but it did not consist of a formal systematic review of NICE guidance.

NICE Technology Appraisal Guidance

Guidance from the technology appraisal programme up until August 2006 comprised 86 current published guidance reports. These reports were consulted, in a general way, to identify any key and recurrent equity considerations, outside of the clinical and cost-effectiveness of health technologies under review. These guidance reports are relatively brief, providing a commentary on the evidence submitted to NICE and the consideration of the evidence submitted to the NICE Appraisal Committee upon which guidance was determined. The main focus in all of these reports is on summarising the health technology, its licensed indication and treatment group, with broad coverage of the evidence reported on the clinical and cost-effectiveness of the technology. Whilst clinical evidence and a judgment on cost effectiveness are present in all guidance from NICE, other factors are not covered in a systematic manner.

The presentation of the guidance has changed over time, as the NICE appraisal process has developed (NICE, 2004). More recent guidance includes discussion against a greater number of the considerations by/from the Appraisal Committee. However, these considerations are almost entirely those related, in various ways, to the clinical effectiveness evidence, the estimates of cost-effectiveness, the methods used to assess clinical and cost-

effectiveness, and the issues related to the delivery of the intervention. In only a small number of cases is there coverage of considerations which 'may' be interpreted as specific 'equity' issues. Devlin and Parkin (2004) present a detailed review of NICE guidance published up to May 2002, comprising the first 39 published guidance reports from NICE, reporting that in only three instances (of 39 guidance reports), did NICE make specific mention to other variables influencing its judgments. These three cases referred to treatment for motor neurone disease, pancreatic cancer and non-small cell lung cancer, with reference being made to health status (poor pre-treatment health status), poor prognosis, and/or low survival rates in these treatment groups. The general review of NICE guidance undertaken here suggests that the guidance published since May 2002, up to August 2006, is consistent with the findings of the detailed review by Devlin & Parkin. In only a small number of cases is it possible to identify a 'specific' mention of other equity considerations that influence NICE judgments. There are cases where the reader 'may' interpret some influence from equity considerations, but there is rarely a specific statement that equity considerations have influenced the judgment made. For example, with reference to the appraisal of imatinab for the treatment of chronic myeloid leukaemia (NICE, 2003), Rawlins and Culyer (2004)[9] state that although the cost-effectiveness of treatment in this case was outside of the range usually acceptable to NICE, "in the absence of any effective alternative treatment ... imatinab was considered to be cost effective in the chronic phase after interferon alfa" (p225). They go on to state that imatinab was made available to patients in the blast cell phase of chronic myeloid leukaemia "on the grounds of equity" (p225). However, in the published NICE guidance there is no specific mention for either of these equity considerations.

Whilst specific mention of equity considerations is rare in NICE guidance, there is specific mention in many instances where no other effective treatment is available. This information is usually presented in the contextual description of the technology and current practice.

This issue of 'other treatments' may be of greater interest given the attention given to 'orphan treatments' by NICE (i.e. a treatment for a disease for which no alternative curative treatment for patients exists). In a consultation document on social value judgments (NICE, April 2005) NICE indicated, in a draft recommendation, that NICE should give special consideration to innovations that provide significant improvements in health for previously untreatable conditions (that is beyond 'best supportive care'), but taking account of the prognosis, the magnitude of the gain in health, and the cost. Although this recommendation

[9] Professor MD Rawlins has been Chair of NICE since its inception in 1999. Professor Culyer was Vice-Chair of NICE from 1999-2003.

did not appear in the final NICE guidelines on social value judgments, it may be the case that 'orphan treatments', or the more general availability or not of other treatment options, might be a societal consideration in priority setting decisions such as those undertaken by NICE. Indeed, Williams *et al* (2005), in a study observing the NICE appraisal process, note that the availability of alternative treatments, or not, was a potential modifying factor in the NICE decision-making process. Tappenden *et al* (2006) have also identified the issue of 'other treatments' as an important attribute in the NICE appraisal process.

NICE Citizens Council

NICE uses a Citizens Council to help it find out what members of the public think about key issues which may inform the development of guidance NICE issues.

The Citizens Council consists of 30 members of the general public. It meets twice a year in public and each meeting lasts up to three days. NICE decides on the topics that the Citizens Council meets to discuss. Meetings are deliberative and draw on a range of presented evidence, including expert witnesses and case studies.

NICE has used the information from the Citizens Council to inform its methodology, and to develop a NICE guidelines document on scientific and social value judgments that inform the work to NICE and its advisory bodies. The guideline on social value judgments is discussed below.

At the request of the NICE Board, the Citizens Council have reported on (i) issues affecting clinical need, (ii) on the circumstances where age should be taken into account when making decisions, (iii) issues related to the National Confidential Enquiries funded by NICE, (iv) drugs to treat very rare diseases ('ultra orphan drugs'), (v) mandatory public health issues, and (vi) the 'rule of rescue' (e.g. saving lives versus improving the lives of people whose lives are not in immediate danger), (see www.nice.org for all reports). The Citizens Council has also been asked to consider the issue of health inequalities from NICE's perspective, and were due to report in 2007.

The reports from the Citizens Council vary in the level of prescriptive recommendations they make. With some specific statements on the views of the citizens council that may be seen as normative, and influential in assisting NICE with social value judgments (e.g. report on age), but others essentially 'scoping' the potentially relevant issues (e.g. clinical need), or not drawing any strong views which may directly help NICE (e.g. report on 'rule of rescue').

The reports from the Citizens Council have assisted NICE in its development of guidelines on social value judgments for NICE and its advisory bodies (NICE, December 2005).

NICE guideline on social value judgments

NICE have issued guidelines on social value judgments that may help NICE and its advisory bodies in developing NICE guidance (NICE, December 2005). The guidelines were issued by NICE in December 2005. There was an earlier consultation document on social value judgments circulated in April 2005. Rawlins & Culyer presented a paper in the *British Medical Journal* in 2004, on social value judgments in the context of NICE. These references have been referred to earlier (in Chapter 3, and in discussion above in this Chapter).

The NICE guidelines on social value judgments describe how such judgments should, generally, be incorporated into the NICE health technology appraisal process and the subsequent guidance. They acknowledge that there will be circumstances where the general principles presented will not be appropriate, stating that in such cases NICE should explain any departure from the social value judgments reported in the guidelines.

The guidelines discuss the problems associated with addressing 'distributive justice'. They discuss the directions to NICE by the Secretary of State for Health (Statutory Instrument 1999, No. 220), and the main competing theories of distributive justice. It is clear from the policy documents covering the NICE social value judgments that NICE do primarily consider the clinical and cost-effectiveness of technologies when making guidance decisions. They do not take affordability into account, and do not specifically consider the number of people to be treated, or budget impact considerations, when making judgments (Rawlins & Culyer, 2004); although these items may be discussed in guidance documents. NICE clearly take efficiency as a major influence when determining guidance, but do reject efficiency as the sole criterion for deciding cost effectiveness.

The guidelines present 13 principles. These principles cover process, the fact that NICE should only support health technologies that are shown to be clinically effective, that NICE must take account of economic considerations (and the range of cost per QALY estimates that are generally regarded as cost-effective), that cost-utility analysis (and therefore the QALY) is the preferred analysis but not the sole basis for decisions on cost-effectiveness, and thereafter specific principles that may be related to distributive value judgments across patient groups. On these latter principles, a summary is that:

- NICE do not support the use of age, gender, racial (ethnic) group, or sexual orientation as priority setting criteria, unless these are clearly linked to clinical effectiveness (benefits, risks).
- NICE do not support the use of social class, social roles, or individuals' income when considering cost-effectiveness. But in its public health guidance there is a wish to promote preventative measures likely to reduce health inequalities associated with socio economic status.
- NICE should avoid denying care to patients with conditions that may be self-inflicted, unless this issue impacts on clinical and cost-effectiveness.
- NICE should not generally give priority to conditions associated with social stigma, unless the additional psychological burdens related to these conditions have not been adequately covered in the economic analysis used.

The NICE guidelines on social value judgments are presented as a 'living document', and NICE expect future updates to cover additional issues, for example, the 'rule of rescue' and 'ultra orphan drugs'. Whilst no principles are presented on the issue of rule of rescue or orphan drugs, the latter issue was covered in the earlier consultation document. No statement is made on the specific issue of 'severity' of health condition, although this may be covered in future additions to the guidelines around a broad definition of rule of rescue.

4.5.5 Social values: Summary of observations from the literature review

- On the basis of the framework used in this review many of the published studies would be regarded as being of low methodological quality (i.e. convenience samples, postal responses).

- Many of the empirical studies are 'methods studies', primarily looking at the way survey questions could or should be framed, rather than seeking to inform health policy decisions.

- The evidence base does not support the use of health maximisation as a dominant and over-riding social value (the use of cost-effectiveness analysis alone does not reflect the distributive preferences of the general public).

- Health gain is a key consideration in a number of empirical studies, but the evidence base unequivocally shows that respondents indicated a willingness to trade-off health gains for a non-health maximising distribution/choice.

- A number of studies have demonstrated that the social value of a health intervention may be dependent on factors other than health e.g. a wish to treat those people with a more severe health condition, the context of the choice presented (life-saving versus life-enhancing), the impact of health care on inequalities between groups.

- The social value of a health intervention is not generally related to the characteristics of eligible patient groups, i.e. limited support for setting priorities on the basis of age, gender, social class, ethnicity, social roles.

- A large number of studies have considered age as a priority setting criterion. Whilst there are a number of studies which show a preference to treat the young over the old (i.e. age as important), the broader evidence base does not support the general use of age as a priority setting criterion. When faced with pairwise choices between old versus the young respondents to studies do tend to select the younger groups ahead of the older groups, but this is not reflective of more general findings (from attitudinal studies, and more qualitative studies) indicating age is not a valid basis on which to discriminate. This finding is supported by the work of the NICE Citizens council.

- A number of studies have shown that respondents to surveys have indicated a preference to treat the more severely affected persons (groups) – those whose pre-treatment health state is severe - ahead of the less severely affected (regardless of the relative capacity to benefit from the alternative groups).

- Studies have shown that the size of the health gain offered by interventions may be an important priority setting criterion. However, findings are difficult to disentangle as there may be a range of issues that are important when respondents have provided preferences – with studies only tending to consider one issue at a time (e.g. age, health gain, or severity). Studies indicate that there may be some form of 'health gain threshold', whereby respondents are not willing to prioritise on the basis of capacity to benefit (i.e. show a preference towards those who can receive the most health gain), but are prepared to give groups less weight (lower priority) if they are only able to receive very small health benefits (regarded by the respondents as 'not worth it').

- Policy statements from NICE in the UK report that NICE take the QALY to be the principal measure of health outcome, and that a QALY (health gain) is regarded as being of equal value to all persons. NICE state that the assumption underlying most of its technology appraisals is that "the weight given to the gain of a QALY is the same regardless of how many QALYs have already been received, how many are in prospect, the age or the sex of the beneficiaries, their deservedness, and the extent to which the recipients are deprived in other aspects of health. The decision to give no differential weight is the result of a social value judgment that an additional adjusted life year is of equal importance each person" (Rawlins & Culyer, 2004, p226).

- NICE have developed guidelines on social value judgments. These guidelines state that age, gender, sexual orientation, income, social class, racial (ethnic) grounds, should not generally be taken into account into account when setting priorities. Social roles, at different ages, should not influence considerations of cost-effectiveness. Gender and sexual orientation could be taken into account where they are risk factors, and age may be taken into account where it is an indicator of benefit or risk.

- Policy guidance from NICE in the UK states that in the appraisal of technologies the Appraisal Committee do not take affordability into account when making judgments. This may be interpreted as not taking into account the numbers involved in the treatment group, and/or the budget impact considerations for the UK NHS.

- Policy documents from NICE suggest that one issue that may be important in the appraisal process is whether other treatments area available to the patient group under consideration. For example, where a drug is subject to appraisal by NICE the Appraisal Committee are aware if this is the only potential treatment for the patient group, or whether other treatments area available.

- NICE have stated that social value judgments have a critical role if resources are to be distributed with efficiency and equity. They have stated that NICE and its advisory bodies have no particular legitimacy to determine the social values of those served by the NHS

- Importantly, a major weakness in the evidence base is that almost all studies have looked at specific social values in isolation e.g. severity versus efficiency, or age versus efficiency, and there is a need to undertake empirical work looking at the relationship

between key social values. This is a common recommendation in three previous literature reviews in this area (Sassi *et al* 2001, Schwappach 2002, Dolan *et al* 2005).

- Social preferences are regarded as important in the formation of health policy. For example, NICE state that one underlying principle is that "advice from NICE to the NHS should embody values that are generally held by the population that the NHS serves" (Rawlins & Culyer, 2004 p226).

4.6 Discussion

The review of the literature, and related NICE policy documents, builds on and extends the earlier reviews of the literature reported in this area. The review has identified over 100 empirical studies that may inform on distributive preferences. The review has presented a summary of the evidence base, against study design, sample size, methods employed and country of origin. In addition, it considers the general quality of the methods used in the empirical studies reviewed.

The evidence based reviewed is an international one, although studies from Europe and North America form the majority of the evidence base. The evidence base includes many studies that use simple methods that may not be regarded as the best quality methods available. The general findings on the methodological nature of the studies reviewed are similar to those presented in earlier reviews, but the review draws on a much broader evidence base. It is still the case that much of the evidence base available is not robust enough to make policy suggestions. However, in some cases the evidence base is wide and varied and is able to put across a clear, albeit general, policy message. This is the case for arguments against the maximisation of health as a dominant decision-making objective. The literature also clearly indicates that the level of benefit (health improvement) is an important social consideration, that efficiency is an important social consideration, and that other factors may impact on the preferred distribution of health care.

The specific nature of these 'other factors' is still uncertain, and much of the empirical evidence can be interpreted in different ways, or may be subject to criticism over the methods used. Potential prime candidates for 'other factors', may be the 'severity of the current (pre-treatment) health state or condition' and the reduction of health inequalities (around socio economic status).

The policy guidance published by NICE in the UK is helpful in considering the factors around distributive justice that may influence judgments on the cost-effectiveness, and ultimately the availability, of health technologies included in the NICE health technologies appraisal programme.

The evidence base reviewed has included mostly studies that have asked respondents to trade-off one issue (at a time) against health gain. Studies that ask respondents to weigh up a number of different issues at the same time are rare. As suggested by previous reviews (Schwappach 2002, Dolan *et al* 2005) one finding from the current review is the need for larger population based studies that seek to estimate the relative importance of key priority setting criteria, to determine some general conclusions in this area. Hauck *et al* (2003), following a review of the literature around priority setting, conclude that "there is a need to make progress in eliciting the publics views about what constitutes a 'fair' distribution of health and health care" (p35).

One general observation, from the very diverse literature reviewed, is that the literature points to an observed general preference not to abandon, and/or to give equal or greater priority too, those groups that may be described as the 'worst off' groups in the choice scenarios presented. These may be severely affected patient groups *vis-a-vis* less severely affected groups. These severely affected patient groups may be those that are in a severe health state (i.e. severe health impairment) or in groups that are presented as at a disadvantage in some way (e.g. poorer prognosis, shorter life-expectancy). For example, studies from Nord (1995), Dolan and Cookson (2000), Ubel (1999), have shown a preference from respondents to place importance on the severity of the health condition, in most cases regardless of the capacity of the more severely affected patients to benefit. Ratcliffe (2000) in a study eliciting preferences for the allocation of donor organs between two patient groups, found that respondents allocated some of the available organs to a patient group that had a poorer prognosis (i.e. shorter life-expectancy). Whilst most respondents elicited a preference that indicated the capacity to survive and benefit from transplant was the most important criterion in the selection of transplant recipients, the majority of respondents also allocated organs in a way that did not abandon the 'less well-off' patient groups. Ratcliffe suggests that respondents felt a proportion of individuals in both groups should be given some chance or hope of receiving a transplant, regardless of prognosis or patient characteristics. Other studies have demonstrated this type of preference for fairness (e.g. Neuberger *et al* 1998, Schwappach 2003, Wiseman 2005). The current review indicates that one way to draw much of the literature together may be in characterising a preference to treat, in some way, the worst off groups in society. This

implied preference is often at the expense of efficiency gains, or in treating competing patient groups. This general observation requires further investigation, and the examination of a preference for the worst off group is explored further in Chapter 5, using severity of health as a starting point.

There are limitations with the review of the literature presented here. Firstly, methodological reviews of this type are not subject to rigourous guidance on the methods to be used for searching, data extraction and evidence synthesis, as is the case in reviews of clinical trials on the clinical effectiveness of health technologies. Therefore reviews of such a methodological nature are open to some elements of bias. The review presented here is thought to be comprehensive, explicit in the methods applied, and balanced in the presentation of findings. However, it may be that there are studies that have not been included that should have been, and it may be that there are areas of the literature that are deserving of greater attention. The review has used English language materials only, and this reflects the constraints of the research in this thesis, i.e. funding, time, opportunity for double data review. However, as indicated, the contributions to this area are predominantly from English source materials. Although it is acknowledged that many of the Asian and South American (Spanish speaking) countries are now beginning to publish in this area.

The review presented is of an outline and summary nature, with material presented in a simple tabular format. The material reviewed has been the basis for background theoretical and analytical thinking to inform this thesis. But it is acknowledged here that material is deserving of greater attention when it comes to the specific presentation of methods and results from the empirical studies included.

Policy guidance from NICE in the UK has been included here for insight. It is accepted that NICE is just one of many policy forums in the UK where health care resource allocation decisions are common, and it may not capture the national or international debates in the area of distributive justice in heath care. However, it is presented here for context, and is regarded as a 'first mover' in international policy forums with its Citizens Council framework, and its explicit statements on social value judgments.

4.7 Research in this thesis

The findings from the above literature review have indicated that further work is required to investigate the preferences of the public for the allocation of scarce resources to competing patient groups. The research in this thesis contributes to research in the area of empirical ethics, through additional empirical contributions, and by providing further empirical support

for the importance of distributive preferences, rather than the use of a dominant health maximising decision maker objective.

In the next chapter of the thesis the preference literature surrounding the issue of severity of health is considered, and it explores and develops the general notion that there may be a social preference for fairness towards the 'worst off' groups in a priority setting task.

The literature review has also highlighted, and reinforced the conclusions from earlier reviews, that the current literature is almost entirely looking at social values in isolation, and that there is a need to investigate the relative values of competing social values in a multi-criterion approach. In chapter 7 of this thesis this issue is considered further, in a multi-attribute study looking at the relative importance across four social values.

5 Empirical Study 1
EXAMINING A SOCIAL CONCERN FOR SEVERITY OF ILLNESS: A PREFERENCE FOR SEVERITY OF ILLNESS, OR A PREFERENCE FOR THE 'WORST OFF' GROUP?

5.1 Introduction

In the previous chapter, the literature review of the empirical studies on social preferences identified several studies that indicated the severity of a persons (group of persons) health state (pre-treatment) may be a valid basis upon which to prioritise the use of scarce resources, when faced with allocation problems (Nord 1993, 1995, Dolan 1998, Ubel 1999, Cookson & Dolan 1999, Oddsson 2003, Gyrd-Hansen 2004). A number of these studies have been widely cited in the health care literature to support a broad hypothesis that health state values (e.g. QALYs) often fail to capture and represent public preferences. Of potentially greater importance is the citation of these empirical studies to support the more specific view that the severity of a persons health state is a basis for giving priority to particular patient groups (the more severely affected), regardless of their capacity to benefit.

The basic hypothesis of the 'severity of health' approach is that the social value of a health improvement (of a given size) is greater the greater the severity of the initial health condition (e.g. Nord, 2005). This hypothesis may indeed be the case, and as Nord (2005) has suggested it is a matter of empirical testing. However, current empirical research comprises studies that are largely simple experimental studies, and the findings are deserving of greater attention in the context of broader distributive preferences. Frequently cited studies are discussed further below (Nord 1993, Ubel 1999). In this thesis these studies are used as a point of departure for further empirical research into the relevance of severity of health as a basis for setting priorities.

The motivation here is not to disprove the specific hypothesis of a preference for severity in such instances, but to further investigate the meaning of the preferences reported against a number of these studies. The hypothesis explored here is that a preference registered for the more severe of two competing groups in an allocation problem reflects a preference to treat 'the worst off group', a general preference for 'fairness'.

5.2 Background

Nord has published a number of studies that suggest respondents have a preference to prioritise in favour to the most severely ill patient groups, regardless of the efficiency gains (health output) forgone. In one early study (published in Norwegian, but reported by Ubel, 1999), Nord (1993) reports a study to elicit the preferences of Norwegian health policy planners. He presented subjects with the following scenario: .....

> *Imagine an illness A that gives severe health problems and an illness B that gives moderate problems. Treatment will help patients with illness A a little, while it will help patients with illness B considerably. The cost of treatment is the same in both cases. There is insufficient treatment capacity for both illnesses and an increase in funding is suggested. Three different views are then conceivable. (1) Most of the increase should be allocated to treatment for illness B, since the effects of these are greater, (2) Most of the increase should be allocated to treatments for illness A, since these patients are more severely ill, (3) The increase should be divided evenly between the two groups. Which of these views comes closest to your own?*

The results reported show that very few respondents favoured moderately ill patients, most preferred to divide resources equally, the remainder favoured treating the severely ill.

Ubel (1999) discusses the empirical work presented by Nord (1993), and its indication of a preference to treat the most severely ill patient groups. The format used by Nord was a simple one. The study was one of the early studies to show such a preference, and Ubel sought to examine how stable the preference for treating the severely ill would be when the wording used by Nord to elicit preferences was modified. The study by Ubel appears to have been motivated by methodological interests and concerns, although the study undertaken and reported by Ubel is widely cited (e.g. Sassi *et al* 2001, Dolan *et al* 2005, Nord 2005) to support a view that severity is an important social value. Ubel replicated Nords survey (applying the same text as far as possible, due to translation needed from Norwegian) in a USA sample and also asked additional similar questions with modified text to get a comparable response. Ubels study was more complex, as he used 6 different questionnaire formats (see Table 4 for details of the questionnaire formats used by Ubel). Three of these questionnaire formats were similar to those of Nord, with three response options (i.e. A, B, or equal). A further three questionnaire formats used a 'forced choice' approach, asking respondents to make a choice between option A (severely ill) or option B (moderately ill). Two of the 6 questionnaire formats applied by Ubel were used to consider the relevance of perspective in the response. In one of the questionnaires with three response categories

and in one of the forced-choice questions, respondents were asked to think about their own self interest when making the policy recommendation. The issues of perspective and forced-choice scenarios, addressed by Ubel, are not considered further here. The focus in the current study is on the preferences contrasted by Ubel with those of Nords earlier study (Nord, 1993), Ubels scenarios Q1 and Q2.

Table 4. Questionnaire format used in the study by Ubel (1999)

Questionnaire format	Description/Format	Response categories
Q1	Original Nord wording	A, B, or divide equally
Q2	Nord format with text amendments	A, B, or divide equally
Q3	As Q2, but with self-interest statement	A, B, or divide equally
Q4	As Q1	Forced-choice: A or B
Q5	As Q2	Forced-choice: A or B
Q6	As Q3	Forced-choice: A or B

The questionnaire administered by Ubel (his Q2) was similar to that of the Nord (1993) study but differed in the text used to describe the response options available to respondents. These alterations were subtle, but important differences. In Ubels amended format the introductory text was the same as above, however, respondents were asked which of the following three views came closest to their own:

(1) Most of the increase should be allocated to treatment for illness B, involving moderate health problems which improve considerably with treatment
(2) Most of the increase should be allocated to treatment for illness A, involving severe health problems which improve a little with treatment
(3) The increase should be divided evenly between the two groups

Ubel reports that for the replication of the Nord study (Q1) results were similar. Very few respondents favoured moderately ill patients (9%), most preferred to divide resources equally (64%), the remainder favoured treating the severely ill (26%).

Ubel reports results similar to those of Nord for the adjusted questionnaire format, in so far as the majority of respondents were willing to give at least equal priority to the severely ill patient group, regardless of the foregone health benefits. But only 6% of respondents expressed a preference in favour of the more severely ill patient group (Group A). Ubel reports that 21% of respondents favoured moderately ill patients, most (73%) preferred to divide resources equally, and only 6% favoured treating the severely ill.

Where Ubel stressed a self interested perspective (Q3) the proportion of respondents favouring the more severely ill patient group increased to 12%, with 75% choosing to divide resources equally. Where Ubel had a forced choice (Q4-Q6) i.e. severe or moderate (no even distribution possible) 40%-57% supported the most severely affected group.
When considering the justification given by respondents for their choices, "fairness" was said to play a role when deciding on the distribution of scarce resources (especially in those that opted to evenly distribute the available funding).

Ubel offers some considerable discussion on methods used, and on the results presented in the context of the general literature. Ubel concludes that the study "showed that public preferences for treating severely ill patients are not as strong as suggested by Nord's study. However, the data still suggest that many people prefer helping severely ill patients, even when they benefit significantly less than moderately ill patients; almost half of the subjects favoured treatment for the severely ill patients despite this large difference in treatment benefit." ... "The present study adds to this accumulating evidence [preference to severely ill] while, at the same time, illustrating the effect that small changes in the wording of scenarios can have on peoples stated allocation preferences" (Ubel 1999, p902). Ubel states his study has no immediate policy implications, however these studies (Nord 1993, Ubel 1999) are frequently cited (e.g. Schwappach 2002, Dolan *et al* 2005) to support 'severity of health' as a priority setting (preference) criterion.

Whilst these studies are often cited in support of the 'severity' criterion, there are some uncertainties in the confidence we can place in the results presented. The methods used by Ubel and Nord are simple self-complete questionnaires, in convenience samples. The studies are experimental in nature, with methodological motivations rather than being aimed at influencing health policy. The nature of the preferences presented, with a large proportion of respondents choosing to divide resources equally between the two groups, also urges caution in drawing conclusions that the preference is indeed an accurate one, and not just an artefact of the choices available i.e. respondents indicating an equal division of resources as they were unwilling to make a difficult treatment choice.

Oddsson (2003) has added to the empirical evidence reported by Nord (1993) and Ubel (1999) by using a similar questionnaire format, but with the addition of a fourth response category allowing respondents to indicate that they "...are not able to make a decision and would prefer that the choice be made by others". The survey by Oddsson was administered via a postal format in random samples of physicians, politicians and the public in Iceland. Oddson uses a number of different questionnaire formats, to address issues other than

severity, but for the severity related choice results indicate that the majority of respondents prefer to distribute resources equally between the severe and less severe groups, regardless of the different capacity to benefit. However, the methods used by Oddsson differ to those of the two earlier studies as they place greater emphasis on the specific capacity to benefit of each of the treatment groups. The scenarios present specific data on the initial severity and the specific treatment effects, therefore the interpretation of the findings is less clear (in comparison to the earlier studies by Nord and Ubel). Interestingly, when given the option to avoid the difficult choice (i.e. reponse option 4 as above), less than 10% of respondents selected this option in the survey.

5.3 Empirical study

The empirical study presented in this Chapter builds on the work of Nord and Ubel *et al.* It examines the preferences of a sample of the UK general public for supporting a concern for fairness in the allocation of resources according to severity of illness and/or level of disadvantage. The study considers the severity hypothesis (e.g. Nord, 2005) and also considers a 'worse-off' group at a level broader than health. This is thought to be the first such study, including a description of disadvantage, to be undertaken in the UK.

The aims of the empirical study were:

a) to add to the empirical evidence on the role of severity of health in the priority setting debate;
b) to consider the meaning of a choice to give equal preference to competing groups;
c) to explore a general hypothesis that there exists a preference to treat the "worst off" groups in society, and that the 'severity' findings may be a reflection of such a preference.

5.3.1 Survey instrument

A survey instrument was developed in which respondents were presented with one question, containing two scenarios, and were asked to indicate a preference from the response categories available. The design of the survey instrument was driven by the previous studies by Nord (1993) and Ubel (1999). The survey design was developed from earlier studies, but underwent a series of pre-pilot and formal pilot stages, to assess the level of understanding, and acceptability, in a general population sample.

The survey instrument used four differing questionnaire formats (Q1-Q4) to address the research questions posed. These formats are summarised below, and presented in more

detail in Table 5. Appendix 3 provides the exact format used, together with the 'Showcards' used in the interview schedule.

- Q1: replicated the question used by Ubel 1999.
- Q2: was identical to Q1 but with the addition of a fourth response category giving the respondent the opportunity to state that they were not able to make a decision and would prefer that the choice be made by others
- Q3: used the same approach as Q1, but replaced the terminology for severely and moderately ill with 'disadvantaged patient group' and 'more advantaged patient group'.
- Q4: was identical to Q3 but with the addition of a fourth response category giving the respondent the opportunity to state that they were not able to make a decision and would prefer that the choice be made by others

The use of "disadvantaged" to reflect the "worst off" group

The use of the terminology of 'disadvantaged' and 'more advantaged' to change the labeling of the worst off group in the scenarios for Q3 and Q4, reflects the common use of the term "disadvantaged" in the social science literature, to reflect a range of possible disadvantages. The language of 'disadvantage' is used by the Office of National Statistics (ONS) in analysis of general household surveys (e.g. *Disadvantaged Households. Results from the 2000 General Household Survey*, 2002), and has been used for data collection and reporting purposes in national reports on social inequalities (e.g. *Focus on Social Inequalities*, ONS, 2004).

The Office for National Statistics in the UK (2002) defines disadvantaged households as follows:

- Workless households (households containing at least one person of working age, but no-one in paid employment);
- Households in which state benefits provide the only source of income;
- Lower income households (gross weekly household income of £250 or less);
- Households in which all adults are aged 60 or more (referred to as 60 plus households);
- Households in which all adults have a long-term health problem;
- Households with no adults born in Britain; and
- Households comprising a lone parent with dependent children (children aged less than 16 or aged 16 to 18 in full time education).

From this listing it is clear that some categories used to identify disadvantage are more explicitly associated with disadvantage than others. For example, where disadvantaged households are categorised by income/worklessness they are by definition economically disadvantaged, while those defined by size or composition are identified as households more likely to suffer social and/or economic disadvantage. Furthermore, there can be considerable overlap between the categories used to define disadvantage. For example, the ONS (2002) state that 71% of households dependent on state benefits were workless households, and that 52% of households comprising lone parents and dependent children were workless households (from UK General Household Survey, 2000).

Table 5. Format used for the interview questions (Q1-Q4)

Q1 –Q2: Severe versus Moderate

Imagine an illness – illness A - that gives severe health problems, and an illness - illness B – that gives moderate health problems. Treatment will help patients with illness A a little, while it will help patients with illness B considerably. The cost of treatment is the same in both cases. An increase in funding is available but we are unable to treat both patient groups. Which of these three (four) views come closest to your own?

Q1	Q2
(1) Most of the increase should be allocated to treatment for illness B, involving moderate health problems which improve considerably with treatment (2) Most of the increase should be allocated to treatment for illness A, involving severe health problems which improve a little with treatment (3) The increase should be divided evenly between the two groups	(1) Most of the increase should be allocated to treatment for illness B, involving moderate health problems which improve considerably with treatment (2) Most of the increase should be allocated to treatment for illness A, involving severe health problems which improve a little with treatment (3) The increase should be divided evenly between the two groups (4) I am not able to make a decision and would prefer that the choice be made by others

Q3-Q4: Disadvantaged versus More Advantaged

Imagine an illness – illness A – where the patient group is disadvantaged, for example, from a low income family, and an illness - illness B – where the patient is from a more advantaged group. Treatment will help patients with illness A a little, while it will help patients with illness B considerably. The cost of treatment is the same in both cases. An increase in funding is available but we are unable to treat both patient groups. Which of these three (four) views come closest to your own?

Q3	Q4
(1) Most of the increase should be allocated to treatment for illness B, in a more advantaged patient group, which improves patients considerably with treatment (2) Most of the increase should be allocated to treatment for illness A, in a disadvantaged patient group, which improves patients a little with treatment (3) The increase should be divided evenly between the two groups	(1) Most of the increase should be allocated to treatment for illness B, in a more advantaged patient group, which improves patients considerably with treatment (2) Most of the increase should be allocated to treatment for illness A, in a disadvantaged patient group, which improves patients a little with treatment (3) The increase should be divided evenly between the two groups (4) I am not able to make a decision and would prefer that the choice be made by others

In the survey instrument used, lower income was used as an example of a disadvantaged group i.e. disadvantaged group (e.g. lower income family). This followed a series of informal pre-pilot interviews (which tested alternative examples of disadvantage, and asked respondents which in their opinion – in the context of the survey - best reflected disadvantage). A formal pilot study of 25 interviewees was undertaken prior to the main survey, using the terminology of disadvantaged (e.g. lower income family) and more advantaged, and the pilot findings supported the acceptability of the terminology (including income as a measure of disadvantage). Feedback from the data collection agency (MORI) also endorsed the use of disadvantage in the context of the current survey.

It is important to note that the terminology of 'disadvantaged' was used to reflect a 'worst off' group, when there was a choice between two groups. It was not the intention of the study to specifically elicit preferences on setting priorities according to this form of group description – i.e. disadvantaged and more advantaged. Indeed pre-pilot work investigated other ways to describe two groups where one group was clearly worse off than the other. For example, clinical prognosis (good or poor), length of life-expectancy, and different categories of social class, were all considered as potential means for describing the two groups in the survey. The initial focus was on an alternative presentation of a 'worse-off' group, but disadvantaged was used to present a broader notion of "worse-offness" (Nord, 2005), not limited to health.

5.3.2 Sample

The main survey instrument was administered on a sample of the adult general population (aged 18 and over) in the Southampton (UK) City Council area. Interviews were carried out in a face-to-face format in-home by the MORI Social Research Institute between 19th September 2005 and 9th October 2005.

The adopted approach to sampling was a random location quota sample. This sets fixed quotas of people to be interviewed in a number of randomly selected sampling points. Sampling points were based on 'Output Areas' (OAs) in the Southampton City Council area, the smallest building block of the Census. For the sample used in this survey 32 OAs were randomly selected (by MORI) proportionate to population size, controlling for socio-demographic composition. Quotas were set by MORI – individually at each sample point – to reflect the socio-demographic profile of residents, on gender, age and work status, using profile data from the 2001 Census (Southampton City Council, 2004).

MORI staff were briefed in detail on the nature of the survey instrument and the candidate worked with them through the formal pilot study. Thereafter, interviewers were briefed by

MORI on the survey instrument and the interview schedule – covering survey aims and background, contact procedures, detailed explanation of the questionnaire. Interviewing was conducted by MORI's own fieldforce of face-to-face interviewers. Interviewing was subject to MORI's in house quality assurance procedures, including a back check of at least 10% of all interviews by telephone during fieldwork. All MORI interviewers had extensive experience, and were members of the Interviewer Quality Control Scheme (IQCS) and were recognised by the Market Research Society (MRS).

5.3.3 Assignment of questionnaire

There were two versions of the main survey instrument, each contained two versions of the above question format i.e. Q1-Q4. Version 1 of the survey instrument contained Q1 & Q3, version 2 contained Q2 & Q4. Each respondent was asked one of the available two questions. Respondents were randomly assigned one of the two main survey instruments and interviewers rotated the questions Q1 or Q3 / Q2 or Q4 in subsequent interviewees. Respondents answered only one of the questions Q1 to Q4.

5.3.4 Sample size

The survey aimed to interview 250 people. There are no formal sample size calculation methods for this type of public preference study. As seen in Chapter 4 preference studies have used samples sizes ranging from very small numbers to samples in excess of 2000 persons. The use of face-to-face interview methods was a limiting factor in the sample size used here. The target sample size was determined on the basis of available time and resources. Overall 250 respondents was judged to be a reasonable target for a study of this nature, when considering the literature generally. However, what made the sample size more important was the use of four different question formats, which leads to analysis on four groups of 60-65 respondents. The study by Ubel *et al* (1999) reports findings from sub-groups of respondents circa. 70 per question frame, therefore the use of a sample size of 250 across Q1-Q4 was thought to be acceptable, given that there was some scope for combining some of the subgroups (Q1-Q4) for some of the analyses.

5.3.5 Data analysis

Differences between the proportion of respondents choosing between each category (group A, group B, divide equally, others to choose) have been evaluated using *Chi-squared* (X^2) tests. The stated preferences for treating group A (severely ill / disadvantaged) or group B (moderately ill / more advantaged), from those selecting A or B, are tested using X^2 tests. The preference for treating group A or B at least equally are tested using X^2 tests.

Respondents socio-demographic characteristics are compared across questionnaire versions (Q1-Q4) with analysis of variance or X^2 tests.

5.4 Results

Interviews were completed by 261 respondents. The socio-demographic characteristics of the sample are presented in Table 6. The sample corresponds roughly to that of the Southampton area (Census 2001), being similar in gender, age profile, work status, ethnicity, home ownership, experience of illness/disability, and health status in general (Southampton City Council, 2004). However, the sample does consist of a higher proportion of retired households, part-time workers, and home workers than the more general population, as might be expected of a face-to-face 'in-home' survey of this kind. For example, in the Southampton area Census (2001) 45% of 16-74 year olds were in full-time employment, whilst 36% of the study participants were in full-time employment.[10] In the sample here 16.1% were working part-time whilst in the 2001 Census 11.1% (16-74 year olds) were working part-time.

Table 7 presents sample characteristics by questionnaire version/group (Q1-Q4), and shows that characteristics did not differ by questionnaire version.

Table 8 shows response data against choices for Q1 to Q4. There was a high rate of engagement in the survey, with 251 of 261 respondents providing a preference when presented with the survey question (1 question from Q1-Q4). Indications from interviewers were that respondents were interested in this type of health related survey question and were enthusiastic about participating. There were only 10 (3.8%) respondents who did not provide a preference across the response categories presented, indicating 'don't know' (or other) when asked to make a choice.

As with previously reported studies (Nord 1993, Ubel, 1999) the most frequent response was to divide resources equally between groups A and B. Where Q1 and Q2 used the format of Ubel (1999) the results do differ. A larger proportion of patients, around 30%, preferred to direct resources to treat group B (the patients with moderately health problems who could benefit considerably), compared to the 21% reported by Ubel, and between 19.7% and 25% of respondents preferred to direct resources to group A (the patients with severe health problems who could benefit a little) compared to 6% reported by Ubel. Overall Ubel reports (in his version 2) 79% preferring either the severely affected group or to divide resources

[10] The sample in this study were 18 years and over, no age limits were specified.

equally between the two groups, and the current study (Q1-Q2) finds between 59-63% support this preference profile.

Table 6: Socio-demographic characteristics of the sample used

	Number of People	%
Total	263	100
Gender		
Male	116	44
Female	146	56
Age		
18-34	91	35
35-54	92	35
55+	80	30
Social grade		
AB	36	14
C1	71	27
C2	65	25
DE	85	32
Work Status		
Working full time	94	36
Not full time	169	64
Working part-time	43	16
House person	23	9
Retired	60	23
Registered unemployed	5	2
Unemployed (not registered)	4	2
Permanently sick/disabled	14	5
Student	20	8
Household income		
Below £17,500	92	35
£17,500 - £29,999	47	18
£30,000 - £49,999	39	15
Above £50,000	16	6
Ethnicity		
White British	221	84
BME[1]/other	40	15
Home ownership		
Owner occupier	150	57
Social renter	60	23
Private renter	47	18
Illness/disability		
Yes[2]	103	39
No	152	58
Household composition		
With children	87	33
Without children	171	65

[1] Black and minority ethnic
[2] Respondent and/or someone else in household

Table 7. Sample characteristics (percentage) by question group (Q1-Q4)

	Question Q1	Question Q2	Question Q3	Question Q4	
Total (number)	68	66	63	64	*Chi-sq*
Gender					*0.872*
Male	44.1	43.9	49.2	42.2	
Female	55.9	56.1	50.8	57.8	
Age					*0.419*
18-34	29.4	28.8	41.3	40.6	
35-54	39.7	34.8	25.4	37.5	
55+	30.9	36.4	33.3	21.9	
Social grade					*0.724*
AB	14.7	13.6	15.9	10.9	
C1	35.3	25.8	22.2	23.4	
C2	19.1	27.3	25.4	26.6	
DE	27.9	31.8	31.7	39.1	
Work Status					*0.173*
Working full time	39.7	33.3	36.5	34.4	
Not full time	60.3	66.7	63.5	65.6	
Working part-time	11.8	13.6	15.9	23.4	
House person	10.3	6.1	9.5	7.8	
Retired	23.5	34.8	20.6	12.5	
Unemployed	1.5	6.1	4.8	1.6	
Permanently sick/disabled	5.9	0.0	6.3	9.4	
Student	7.4	6.1	6.3	10.9	
Household income					*0.528**
Below £17,500	23.5	37.9	41.3	37.5	
£17,500 - £29,999	20.6	21.2	15.9	14.1	
£30,000 - £49,999	17.6	13.6	14.3	14.1	
Above £50,000	10.3	1.5	4.8	25.0	
Refused/not stated	27.9	25.8	23.8	26.6	
Ethnicity					*0.557*
White British	80.9	89.4	82.5	82.8	
BME[1]/other	19.1	10.6	17.5	17.2	
Home ownership					*0.227*
Owner occupier	63.2	53.0	54.0	56.3	
Social renter	20.6	25.8	23.8	21.9	
Private renter	14.7	18.2	20.6	18.8	
Illness/disability					*0.294*
Yes[2]	39.7	48.5	38.1	31.3	
No	54.4	50.0	60.3	65.6	
Household composition					*0.216*
With children	30.9	28.8	27.0	43.8	
Without children	67.6	68.2	69.8	56.3	
Health in general					*0.919*
Good or very good	73.5	74.2	66.7	70.3	
Fair	14.7	15.2	20.6	21.9	
Bad or very bad	10.3	9.1	9.5	7.8	
Health Insurance					*0.63*
Yes	14.7	10.6	14.3	18.8	

[1] Black and minority ethnic
[2] Respondent and/or someone else in household
*combining £30,000-£49,000 and Above £50,000

However, results are generally consistent with those presented by Ubel, in so far as the clarification (amendment) to the text used by Nord (1993) resulted in a significant difference in the percentage choosing group B, i.e. Nord reports that 'very few' chose group B in his original study, and Ubel reports only 9% preferring group B where his study used the original text from the Nord study. The amendment to the text led to between 21% (Ubel) and 30% (in this study) choosing to treat group B.

Where the experiment has been amended to present the choice as one between patients who are disadvantaged with treatment helping a little versus patients who are more advantaged with treatment helping considerably (Q3 & Q4), the response data shows a stronger preference to treat the 'worse-off' [disadvantaged] group (stronger than in the severity of health questions). In these questions, between 28% and 46% of respondents indicated that they preferred to treat the disadvantaged group (37% when pooled), and between 33% and 55% preferring to divide the resources equally between the two groups (44% when pooled). Only 12.5% to 17.5% (15% pooled) preferred to direct resources to the more advantaged patients, regardless of their greater benefit from treatment.

Where the study offered respondents an opportunity to avoid such a difficult decision and to let others choose (Q2 & Q4), very few respondents selected the "others to choose" option, only between 3.1% and 7.4% (7/132 people in total, 2.68% when pooled). This indicates that where respondents are stating a preference for equality it may be interpreted as a true preference to divide resource equally between the two groups. Whilst there are only a small number of cases (10 in total) where respondents have not provided a preference (i.e. don't know or other), the numbers are smaller in the questions where respondents were able to opt to 'let others choose' (Q2 & Q4).

Table 8 reports *p*-values comparing (X^2 tests) differences between those respondents who expressed a preference for group A or B (i.e. the statistical testing ignores those respondents giving equal funding to both groups). Only in Q3 and Q4 (with scenarios for disadvantaged versus more advantaged) are there statistically significant differences between those choosing one group over another. In the combined results, over the full sample, there is no statistically significant difference (p=0.123) between those respondents choosing either A or B (77 vs. 59 respondents).

Considering the combined results (Table 8), combined by group description, there is a stronger preference to support the worse off group when it is described using terminology of 'disadvantaged' versus 'more advantaged' (Q3 and Q4). Combining responses from Q3 and

Q4, 37% expressed a preference to treat the disadvantaged patient group, with only 15% preferring the more advantaged patient group, who could get greater help (benefit). Over 80% of respondents (Q3 and Q4) preferred to give at least equal preference to the two groups, even though patient group A were only able to benefit 'a little' from treatment. The preference to treat the worse off group, who could only benefit 'a little' (compared to considerable benefit) was weaker when the group were described according to severity of health condition (Q1 and Q2). In the case of severe versus moderate health problems, combining findings from Q1 and Q2, 30% of respondents expressed a preference to treat patient group B, who had moderate health problems and were able to get considerable help from treatment. Only 22% preferred to treat the patient group with severe health problems, getting a little help from treatment. Slightly over 60% of respondents (Q1 and Q2) preferred to give at least equal preference to the group A, with severe health problems, regardless of the limited help available from treatment.

When combining response data from all four of the subgroups, regardless of descriptions used and response categories available, there is a strong preference to give group A (the worse-off group) at least equal priority, with over 70% of respondents selecting either A or the equal priority option. Although, a substantial proportion (over 20%) of respondents show a preference to treat those who are better off and able to get the greater help from treatment.

Table 8. Responses to preference questionnaire (Q1-Q4)

			Number (percentage) of respondents expressing preference by group					
Version	Description	N	Group A (severe/disadv)	Group B (moderate/adv)	Equal for A & B	Others to choose	Other / Don't know	*p*-value[a]
Q1	Ubel text – 3 response categories; *Severe health problems where treatment helps a little vs. moderate health problems where treatment helps considerably*	66	13 (19.7)	19 (28.8)	29 (43.9)	NA	5 (7.6)	*0.289*
Q2	Ubel text – 4 response categories; *Severe health problems, treatment helps a little vs. moderate health problems where treatment helps considerably*	68	17 (25)	21 (30.9)	23 (33.8)	5 (7.4)	2 (3)	*0.516*
Q3	Adapted text – 3 response categories; *Disadvantaged patient group where treatment helps a little vs. more advantaged patient group where treatment helps considerably*	63	29 (46)	11 (17.5)	21 (33.3)	NA	2 (3.2)	*0.004*
Q4	Adapted text – 4 response categories; *Disadvantaged patient group where treatment helps a little vs. more advantaged patient group where treatment helps considerably*	64	18 (28.1)	8 (12.5)	35 (54.7)	2 (3.1)	1 (1.6)	*0.05*
	Total	261	77 (29.5)	59 (22.61)	108 (41.38)	7 (2.68)	10 (3.83)	*0.123*
Q1+Q2	Results combined for questions on patients with severe vs. moderate health problems	134	30 (22.4)	40 (29.9)	52 (38.8)	5 (3.7)	7 (5.2)	
Q3+Q4	Results combined for questions on disadvantaged vs. more advantaged patients	127	47 (37)	19 (15)	56 (44.1)	2 (1.6)	3 (2.4)	

[a] *For columns titled 'Group A' and 'Group B' only*: Chi-squared test of proportions preferring A or B; to test if patients had an overall preference for treating Group A (severe or disadvantaged) or Group B (moderate or more advantaged), from those selecting A or B.
NA = not applicable/not contained in that version of the question

5.5 Discussion

This study has further explored the use of a severity hypothesis for setting priorities in health care. Using a sample of the general public, and face-to-face interview techniques, the study supports the severity hypothesis (Nord, 2005). The results are similar to those presented by Ubel (1999), showing a less strong preference for the more severely affected patient group than the study of Nord (1993), when text is changed in a subtle but important way. But, like findings from Ubel's study, there is support for the severity of health hypothesis. These previous studies, and the current study, show that where a patient group is not able to benefit greatly from treatment, compared to a group competing for the same resource that are able to show greater benefits from treatment, respondents still wish to give at least equal priority to that more severely affected patient group. Also, in a large number of cases there is a positive preference to give priority to the more severely affected patient group.

Furthermore, the study has explored a more general hypothesis that respondents will prefer to give priority to those groups that may be regarded as the worse-off groups, regardless of the basis for being worse-off. Where the worse-off groups in the current study have been described using severity of health condition and a broader level of disadvantage, respondents have given a strong indication that worse-off groups should; (i) at least be treated equally, even though they are able to get only a little help (compared to considerable) from treatment, (ii) could be considered for greater priority, with 22% to 37% of respondents preferring to give priority to the worse-off groups, compared to between 15% and 30% who preferred to give priority to the better-off groups. This latter finding should be interpreted against the differential help available, to alternative groups, from available treatment.

The study findings have also offered a more general insight to the response categories available and used in such empirical studies. In earlier studies by Nord (1993) and Ubel (1999), particularly this latter study, the majority of respondents indicated an equal preference against the two scenarios presented in the questions answered. Given the absence of alternative response categories, the general meaning of this common 'equal' response was uncertain. One suggestion was that it may have been used to avoid a difficult decision between two competing options. However, given the findings in the study presented here, it seems more likely that a preference to give equal priority is a true preference for equality. Very few respondents in the current study (frames Q2 and Q4) opted to avoid the difficult decision and to indicate that the 'choice should be made by others'. This finding may also be generalisable to other studies of a similar format, although

it remains experimental and indicative, rather than robust, and should be explored more widely in future studies.

Whilst at one level the findings presented here support a general 'severity' hypothesis, they also cast some doubt over the specific preferences elicited against 'severity' as a priority setting criterion. Findings offer some initial support for a more general hypothesis that the general public have a broader preference to support those that may be regarded as the worst-off groups, and that severity of health condition is just one presentation of such a preference, or potentially a 'proxy' for this broader preference. Nord has championed the use of severity of health as a basis for setting priorities when resources are constrained, but he has also supported the notion that there is a broader concern for the worse-off (Nord, 2005). Rawls (1971) is often cited as suggesting that a theory of justice should place special importance on the position of the worse off in society (discussed in Chapter 3 of this thesis). Parfit (1991) has suggested a general concern, across social choices, for the worse-off, with the presentation of 'The Priority View'. Brock (2001) has offered some rationale for giving greater concern to the worse-off groups, who he suggests may have more urgent needs.

Chapter four of this thesis, reviewed the empirical literature around social values, and priority setting criteria, suggesting some support for a more general preference to treat the worse off. The findings in the literature review presented indicate a number of studies where the worse off from the available choice scenarios were favoured, or at least given equal preference by respondents (when efficiency gains were much smaller). The current study, offers some support for a more general preference to treat the worse-off groups, regardless of the way these groups are described. The findings here indicate that when patient groups are described in a more general way (not limited to health) there is a stronger preference towards the worse-off groups.

Whilst severity of health as a basis for setting priorities may be an accurate reflection of social values, it may also be the case that it is not severity of health *per se* that is driving the preferences given. It may be a broader preference to support those regarded as worst off, and such a social value would draw together much of the empirical literature around health care priority setting, to reflect a general preference for 'fairness'. Such a preference for fairness is evident in the broad literature that argues against the use of efficiency arguments, in the form of health gain maximisation, in heath care priority setting decisions. Although, the literature does also indicate potential thresholds around where and when it may be appropriate to give priority to the better off groups (e.g. where benefits from treatment are not regarded as meaningful to patient groups). Therefore, the current study does support a

potential preference for 'fairness', but it does not explain what is meant by fairness (or equity), and it is still down to empirical testing to determine what balance of gains and losses is an acceptable presentation of fairness.

The current study has limitations. The study uses hypothetical scenarios, although this is difficult to avoid given the nature of priority setting choices in a sample of the general public. The current study uses previous studies, and study frames, presented by Nord (1993) and Ubel (1999), and is therefore limited in the text presented and subject to the framing and context effects present in the original text. As with earlier studies from Nord and Ubel, it is difficult to quantify the differences between alternative treatment scenarios, with differences represented by qualitative labels ('a little' versus 'considerable' help). Earlier studies (Nord and Ubel) have regarded the differences presented as 'extreme' scenarios, although this is down to subjective judgment and interpretation (and deserving of further qualitative research and investigation).

As with earlier studies, the current study uses simple scenarios, and asks respondents to indicate a preference using the general allocation of a hypothetical budget (constrained). Given the simple experimental methods used, Nord (1991) and Ubel (1999) have stated that the findings from their studies have no immediate policy implications, as the text used is judged to be "too vague" to guide policy decisions (e.g. Ubel 1999, p901). However, these studies have been influential in the health care literature, and potentially in health policy forums acting on the available literature. The current study has used a more formal and rigorous sampling frame and interview techniques, and the questions were asked at the end of a 20-25 minute health related priority setting interview. During the interview respondents had had the opportunity to consider priority setting scenarios within the UK NHS, and to deliberate on the relative merits of 'severity of health condition' as a priority setting criterion, as well as other criteria (including efficiency and non-health attributes).

5.6 Conclusions

This study makes a contribution to the empirical evidence base on health care priority setting. It supports findings from earlier studies (e.g. Nord, Ubel) that respondents to such studies prefer to give at least equal priority to persons who are severely affected by their health condition, regardless of a lower expected health gain in these patients groups. This study, as in the previously mentioned, also adds to the evidence base that indicates that society (general public) does not support a strict health maximisation objective when it comes to difficult priority setting choices, and the allocation of health care resources.

The study replicates the study by Ubel (1999), but importantly it makes a contribution by testing the methodology in a quasi-random representative sample of the general public, using face-to-face interviews. It builds on the work of Ubel by testing the meaning of a preference for dividing resources equally between groups, and it also makes a novel contribution by framing the question differently, using the different descriptors of disadvantaged and more advantaged for the alternative patient groups in the priority setting choice.

There may be some relationship between severity of health condition and disadvantaged groups more generally, and it may be that the evidence frequently cited in support of a public preference to treat the most severely affected groups in terms of health condition, are reflecting a more general preference to treat the worst off groups, a general preference for fairness. This more general preference for 'fairness' fits with a hypothesis that society values the way health care benefits are distributed, as well as the total health care benefits available from limited resources. The presentation of a general preference for fairness reinforces this hypothesis, but the challenge of weighing up the efficiency gains and losses against distributive preferences is a difficult one. The following chapters (Chapters 7 and 8) consider this issue further.

6 DISCRETE CHOICE EXPERIMENTS: INTRODUCTION TO THE GENERAL METHODOLOGY AND A REVIEW OF THE STUDIES ELICITING SOCIAL VALUES IN HEALTH CARE

6.1 Introduction

The purpose of this chapter is to introduce the methodology of discrete choice experiments (DCEs). The chapter presents a brief introduction to the theory underpinning discrete choice modeling, the process of carrying out a DCE (steps involved), the current practice of DCEs in the health economics literature, and a short review of DCE studies undertaken in the context of social preferences. The chapter provides introductory information and background on DCE methods as a lead-in to the next chapter which presents a DCE experiment undertaken as part of this thesis. By highlighting the methodological status of discrete choice modeling in health care, especially in a social context, the chapter emphasises that the thesis is making a contribution to the advancement of the DCE methodology, and that it contributes to a sparse evidence base.

Alternative approaches to the discrete choice experiment, for the elicitation of public preferences, have been described in detail by Ryan *et al* (2001). Other techniques available for the elicitation of preferences cover both qualitative and quantitative methods. Qualitative interview-based methods are one approach to the investigation of social values in health care. But following the literature review reported in Chapter 4, it was judged that whilst they may provide some useful general information (such as findings reported by Cookson & Dolan, 1999), they were unlikely to provide a specific quantitative insight to the relative weights for competing social values. Quantitative methods available include attitudinal surveys and experimental choice sets, as examples discussed in Chapter 4, and other techniques such as contingent valuation.

Following the detailed review of the literature around social values, and general consideration of other techniques in the context of health technology appraisal, the DCE approach was selected to explore the social value of health care interventions. This decision was based on the techniques suitability for the elicitation of public preferences (Louviere *et al* 2000, Ryan *et al* 2001), the opportunity to present findings in terms of quantitative estimates indicating the relative weights for attributes and levels used, and its intuitive appeal in the description of social values, in a choice scenario, across different levels of impact or effect. The evidence base currently available, mainly in the elicitation of individual (non social) preferences, reports that DCE is a promising technique for the investigation of

weights attached to different attributes that are influential in a choice scenario (Ryan & Gerard, 2003). Indications from the small number of DCE studies eliciting social preferences[11] (5 references available up to end of 2002, Table 12), supported the use of the DCE framework in this context. More recent contributions to the literature (2003-2004 specifically for the decision taken here, Table 12), supported the use of DCE for the elicitation of social preferences.

6.2 The DCE Approach

Discrete choice experiments use survey (e.g. questionnaire) data to explore preferences over alternative options (e.g. services, products) in a choice set. These surveys present hypothetical choices using attributes to describe the options in the choice set. These common attributes are varied across a specified and plausible range of levels. An example of a discrete choice question is presented in Table 9, where the attributes are related to characteristics of the options in the choice set, and the characteristics vary across levels (e.g. quantitative time periods, or qualitative categories).

Table 9. Simple example of a discrete choice question on preferences for type of surgery

Attributes	OPTION 'A'	OPTION 'B'
Type of surgery	Open	Laparoscopic (key hole)
Side effects	Frequent	Rare
Waiting time	1-week	1-month
Recovery time	4-weeks	2-weeks
Chance of readmission	0.05%	5%
Which option do you prefer? (tick one box)	Prefer A ☐	Prefer B ☐

The choice experiment is a stated preference technique. Stated preference analysis is based on the idea that whilst people may have difficulty providing a specific value for a good or service, they can typically indicate which of several goods or services they would be most likely to choose from the alternatives presented to them. Stated preference analysis then makes a link between the utility function and the stated choice. Whilst the DCE approach asks respondents to choose between alternatives presented in the survey questions, other stated preference techniques are used whereby respondents are asked to rank or rate

[11] Social studies, as described in the introductory Chapter 2, and later in this chapter, refer to the elicitation of preferences (social preferences) from individuals in the context of resource allocation for 'others' in society.

scenarios/alternatives. Economists have tended to favour the discrete choice approach due to its theoretical foundations. In general, the DCE approach involves a blend of economic theory, econometric methods and survey design theory (psychology and statistical design). The approach is broadly consistent with microeconomic theory of consumer behaviour (rationality, preference completeness, and continuity of preferences, see below), and it draws upon Lancasters theory of consumer choice (Lancaster 1966, 1971), and random utility theory (McFadden 1973, Hanemann 1984). These theoretical underpinnings are discussed below.

The historical development of the discrete choice modeling approach is described by Ryan (1996, 1999). It has its origin in market research, where it has been used to identify factors that influence choice and subsequent demand for commodities. Whilst economists have traditionally relied on actual market behaviour, revealed preference data, to understand the factors affecting decisions by consumers (e.g. sales figures for goods/services), there are some areas of the economy where there are limitations on the availability of market data. In such cases stated preference data, from hypothetical choices, are often used to explore preferences and choice. This has been the case in the analysis of choice by analysts working in the areas of transport economics and environmental economics, where discrete choice methods have been widely used (Mark & Swait, 2004). From the mid 1990's there has been a growing interest in the application of DCEs in health care, where it has been used as an alternative approach for health benefit valuation and to value the broader outcomes associated with health care provision (e.g. non-health attributes, satisfaction with services), (Ryan & Gerard 2003, Fiebig *et al* 2005). Some of these applications are discussed below. The technique has been used more recently (in a small number of studies) to explore preferences in a 'social' context, related to resource allocation and priority setting choices involving other people in society (e.g. Farrar *et al* 2000, Ratcliffe 2000, Schwappach 2003, Gyrd-Hansen 2004). This latter application is the primary focus of the current thesis.

6.3 Theoretical underpinnings of discrete choice methods

6.3.1 Microeconomic theory of consumer choice

Discrete choice methods contain elements of the traditional microeconomic theory of consumer choice. This theory underpins the economists' notion of preference, and preference elicitation. A brief overview of consumer theory has been presented in Chapter 2, covering the axioms of (1) completeness, (2) transitivity, (3) continuity and (4) non-satiation, together with a general insight to other aspects of consumer theory. Essentially

consumer choice revolves around the concept of "preference": when an individual reports that "A is preferred to B" it is taken to mean that all things considered they feel better off under situation A than under situation B. This preference is usually assumed to be consistent with the basic properties (axioms) of consumer choice theory.

Given the axiomatic approach to consumer choice theory, it is possible to show formally that people are able to rank in order all possible situations (alternatives) from the least desirable to the most. Using the notion of utility, it also follows that the more desirable situation offers more utility than the less desirable ones. Therefore, if a person prefers A to B, the utility assigned to A exceeds the utility assigned to B. Individual preferences are assumed to be represented by a utility function of the form $U(X_1, X_2, \ldots Xn)$, where $X_1, X_2, \ldots Xn$ are the quantities of each of n goods that might be consumed.

6.3.2 Characteristics approach to theory of consumer choice

Lancaster (1966) presented an alternative approach to the conventional theory of consumer behaviour. Lancasters theory of consumer choice states that commodities are consumed for their attributes (or characteristics) rather than for their own sake. The theory focuses on the intrinsic qualities of goods, and regards the goods bought by consumers as inputs into a process of consumption which transforms the bundle of goods bought into a bundle of *characteristics*. For example, different mixtures of foods will produce different mixtures of characteristics. The second stage in the theory is to assume that consumers have preferences for characteristics of goods rather than for the goods themselves. This second point is very important for the discrete choice approach, where alternatives in a choice set are described according to characteristics (attributes). In discrete choice experiments alternatives are usually part of a choice set of a common commodity (i.e. different ways to describe the same type of good or service), and the set of attributes of interest are common to all alternatives in the choice set. Discrete choice modeling, in line with Lancastrian utility theory, assumes that there is a latent utility for the commodity which can be separated into 'part-worth' values. Preferences for different commodities are examined according to preferences for defined characteristics. Within discrete choice experiments alternatives (commodities) can be described in terms of attributes and an individual's valuation of the commodity depends on the levels of these attributes (Ryan, 2004).

The standard Lancaster approach (presented by Louviere *et al*, 2000) suggests that goods (X) are transformed into characteristics, t, through the relation:

$$t = BX,$$

where B defines the consumption technology (turns goods into a set of characteristics) i.e. uses a common set of characteristics (attributes) to describe alternatives in the choice set. The theory imples that utility is a function of the characteristics of alternative options in the choice set:

$$u = U(t_1, t_2, \ldots, t_R)$$

where t is the amount of the characteristic that a consumer obtains from consumption of commodities (1,....,R). However, the theory presented by Lancaster has limitations. For example, it assumes goods are infinitely divisible, frequently purchased and of low unit value. And these limitations are frequently relevant to discrete choice applications. Rosen (1974) presented a characteristics approach for indivisible (or discrete) goods that is more appropriate to the discrete choice framework. This approach is not presented in any detail here, as the important feature of the theory is that it continues to link utility directly to the characteristics of goods, but in a way that is consistent with discrete choice modeling. The theories from both Lancaster (1966) and Rosen (1974) involve a complex set of functions, and to operationalise the theories would require the analyst to have detailed information on the preferences of individuals, characteristics of goods, and thereafter to model such influences on utility. This is not feasible in practice, and the link from theory through to discrete choice modeling is essentially conceptual. However, a link between traditional consumer theory, the theoretical appeal of the characteristics approach, and empirical preference analysis is provided by random utility theory (RUT).

6.3.3 Random Utility Theory

Random utility theory (RUT) is based around the notion that consumers choose alternatives that provide them with the greatest utility. In the choice experiment context, the probability that an alternative is chosen, from a choice set, increases as the utility associated with the alternative increases.

For RUT the utility that an individual derives from an alternative is considered to be associated with the attributes of the alternative, with the utility function composed of a deterministic component (V) and an unobservable or stochastic (random) component (ϵ):

$$U_{in} = V_{in} + \epsilon_{in}$$

Where V_{in} is the deterministic component of the utility for individual n and option i; ϵ_{in} is the random or unobservable component for individual n and option i. V_{in} is the indirect utility function in which the attributes are arguments. V is therefore characterised as:

$$V_{in} = \sum_i \beta_i X_{in} + \sum_p \psi_p W_{pn}$$

Where V is the utility, X is a matrix of attribute levels, W is a vector of p individual characteristics, β_i are the coefficient estimates for each attribute in the matrix X (marginal utilities), and the ψ_p coefficients represent the extent to which personal characteristics influence choice (patient characteristics are considered as interaction effects in the econometric model, see below).

Applying random utility theory it is possible to consider the probability of choosing option i over any other option j belonging to the same choice set (C):

$$\begin{aligned} Pr(i) &= Pr(V_{in} + \epsilon_{in} \geq V_{jn} + \epsilon_{jn}) \quad \text{[choice set C of J]} \\ &= Pr\{ (V_{in} - V_{jn}) > (\epsilon_{jn} - \epsilon_{in}) \} \end{aligned}$$

For example, assuming the ϵ term is independently and identically distributed with a Type 1 extreme value, e.g. a Gumbel distribution (Louviere *et al*, 2000), then the logistic distribution can be used to derive the probability outcomes and the conditional logit model used.

6.3.4 Discrete choice models (conditional logit model)

DCE data are modeled using the random utility maximization framework (Louviere *et al* 2000, Hensher *et al* 2005). There are a number of models available that can be used for analysis of DCE data, and the class of model used is known as categorical dependent variable models (CDVMs). In CDVMs the dependent variable (left hand side) is categorical (neither interval nor ratio in nature). There are a range of CDVMs available, and the choice of model will be dependent on the design of the DCE, amongst other things. DCEs may be binary choice designs (two alternatives), or multiple choice designs. For binary choice models the most frequently used models are the logit and the probit model. For multiple choice models multinomial, nested and mixed logit or probit models are commonly used. In this thesis the DCE study reported (Chapter 7) is a binary choice design, therefore that is the focus in this outline on modeling methods, and readers are referred to other references for further information on modeling in multiple discrete choice designs (Hensher *et al,* 2005).

The binary choice models, logit and probit, are the easiest to apply and interpret. These models (as with all CDVMs) adopt the maximum likelihood estimation method, which requires assumptions about probability distribution functions. The logit and probit models are regarded as very similar, and although they produce estimators on a different scale, they

are expected to provide similar conclusions. The core difference between the logit and probit models lies in the distribution of errors (ϵ), with the logit models using the standard logistic probability distribution, and the probit model assuming a standard normal distribution. Both of these distributions have an 'S shaped' curve (cumulative distribution function) that ranges between 0 and 1, which makes them very suitable for dealing with probabilities. The difference between the logit and probit is regarded as trivial and is based on the weighting in the tails of the distributions (Kjaer, 2005). Greene (2003) argues that there may be practical reasons for favouring one model over the other, but that it is difficult to justify the choice of one distribution over another on theoretical grounds.

The logit and probit model are used to model binary (or dichotomous) outcome variables, from a set of explanatory variables. For example, in cases where the outcome (dependent variable) is success or failure, high or low, yes or no i.e. the interval of the response is an interval such as 0 or 1 (e.g. probability scale between 0 and 1). The logistic regression (or logit) model uses a transformation of the dependent variable (e.g. probability scale transformed to an odds scale) in order to model binary or categorical outcome variable, and is discussed here to demonstrate the theoretical presentation of the binary choice model (the probit model works on a similar basis, but uses the standard normal distribution and does not operate on a log-odds scale).

The logisitic regression model for the odds (e.g. odds of success) is of the form:

$$\pi / (1-\pi) = e^{(\beta 0 + \beta 1 X1 + \beta 2 X2 + \ldots + \beta k Xk)}$$

An alternative and equivalent for the log of the odds is:

$$\text{Log}[\pi / (1-\pi)] = \text{logit}(\pi) = (\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k)$$

The logit is just another transformation of the underlying probability π. And the model can also be written in terms of the underlying probability of a success outcome:

$$\pi = e^{(\beta 0 + \beta 1 X1 + \beta 2 X2 + \ldots + \beta k Xk)} / 1 + e^{(\beta 0 + \beta 1 X1 + \beta 2 X2 + \ldots + \beta k Xk)}$$

All three forms of the logistic regression model above are equivalent.

When using discrete choice data, the data is of a matched format (such as in a case controlled study) with two or more options and one response category. In such data the conditional logistic regression model is used.

The conditional logit model takes the form:

$$Pr(i) = e^{Vi} / \sum e^{Vj} \qquad [C \text{ of } j=1..,n]$$

This derivation of the conditional logit model is attributable to Luce and Suppes (1965) and is used to compare, in a choice between two alternatives, the probability that $U_{1n} > U_{2n}$. With this form also generalisable to choice sets greater than two alternatives. When specified as a linear-in-parameters functional form, an increase in the value of one of the attributes in an alternative in the choice set dictates a proportional rise or fall (depending on the specific estimator, β) in total utility (Skoldberg & Gryd-Hansen, 2003).

The conditional logit model is more precisely written as:

$$Pr(i) = e^{\lambda Vi} / \sum e^{\lambda Vj} \qquad [C \text{ of } j=1..,n]$$

Where λ is a scale parameter which is inversely related to the standard deviation of the error terms. This scale parameter affects the estimates of the β coefficients, and it varies with each model. Louviere *et al* (2000) and Louviere (2001) have discussed this scale parameter in some detail, highlighting that it is not possible to identify this scale parameter (due to its multiplicative form), and identifying it as a limitation of the assumptions of the econometric techniques used. However, Louviere (2000) has also stated that it is legitimate to set the λ scale parameter equal to one in those cases when single choice experiment models are estimated and reported.

6.3.5 Assumptions of the conditional logit model

The conditional logit model is a very popular approach to data analysis in choice experiments, this is mainly due to its simplicity and the availability of easy to use software packages, and to the use of the log-odds scale allowing interpretation of the regression estimators as odds ratios, indicating the effect of a unit change in the explanatory variables on the independent variable. However, there are four implicit assumptions when using the conditional logit model with the random utility model in the choice experiment context:

1) all individuals have the same (average) preferences;
2) homoscedasticity;
3) independence of irrelevant alternatives (IIA); and
4) orthogonality of choices.

These assumptions apply to both the logit and probit models, and they have both strengths and weaknesses. Where they are restrictive they may be relaxed to varying degrees either within the framework of the conditional logit model, or through the use of an alternative econometric model.

The first of the above assumptions can be relaxed by investigating the influence that individual characteristics have on choice. This is undertaken by the introduction of covariate interaction terms in the discrete choice modeling analysis. Alternatively, more sophisticated models, such as random parameter logit or latent class models, may also be used (Adamowicz, 2003).

Homoscedasticity is an assumption that the variability of unobserved influences on choice (i.e. the ϵ term) is roughly the same regardless of the non-chosen items in the choice set i.e. the variance is uncorrelated to the other options. The assumption of homoscedasticity simplifies mathematical and computation treatment of data in the estimation process and may lead to good estimation results even if the assumption is not true. The assumption may not be true in choice experiments because each individual is required to answer a series of choices and the unobserved component for each individual may thus be correlated. The model can be adjusted to take this into account by using a random-effects model.

The third assumption on the independence of irrelevant alternatives (IIA) states that:

> *the ratio of the probabilities of choosing one alternative over another (given that both alternatives have non-zero probability of choice) is unaffected by the presence or absence of any additional alternatives in the choice set.* (Louviere *et al* 2000, p44)

The IIA property implies that the random elements in utility (i.e. the ϵ term) are independent across alternatives and are identically distributed. This requires that the probability of choosing one alternative over another depends only on the utility of those two alternatives, i.e. within a set of alternatives there is a simple substitution structure for all alternatives. Louviere *et al* (2000) discuss the IIA assumption as being both a strength (computational convenience) and a weakness (potential for biased estimates) of the choice model. They claim that the satisfaction of the IIA condition should not be of general concern due to the fact that it is neither desirable nor undesirable *a priori*, with empirical factors being important in the consideration of the IIA assumption. For example whether the attributes and levels used offer a simple substitution structure, and whether the choice presented in the survey includes close substitutes. To relax the IIA assumption a random parameter model can be used (Adamowicz, 2003). In this respect the IIA assumption is more restrictive in the logit

model than with the probit model, as the software more commonly available (and used in this thesis in Chapter 7) is not able to transform the standard conditional logit model into a random parameters model. The probit model is more readily available in a random parameters presentation (e.g. in STATA software).

Orthogonality is a mathematical constraint requiring that all attributes be statistically independent of one another. The assumption on orthogonality of choices dictates that the choices presented in a survey are balanced and that there is no multicollinearity in the design. An orthogonal design is one in which the columns of the design display zero correlation, and where there are no interactions between the attributes. Experimental designs which are orthogonal can be identified from design catalogues, software or experts (Burgess & Street 2003, 2005, Louviere *et al*, 2000). If the experimental design used is orthogonal, the assumption of orthogonality of choices within subjects is, by definition, reasonable. However, there is a potential trade off between orthogonality (statistical efficiency) and the presentation of realistic choices, and such considerations should be taken into account by analysts (Louviere 2001, Swait and Adamowicz 2001, Louviere *et al* 2002, Bateman *et al* 2002, Hensher *et al* 2005).

6.4 Stages in a discrete choice study

Applications of discrete choice experiments generally follow the six steps outlined below (Table 10).

Once the decision problem has been characterised, and the appropriateness of the discrete choice methodology has been determined, the challenges faced in the design of the survey surround the identification of the attributes and levels, and the development of the survey instrument in conjunction with the experimental design theory for such studies.

The number of attributes and the values and labels for the levels for each attribute needs to be appropriate for the decision problem of interest, and the context of that problem. A number of approaches can be taken to determine the attributes of interest. These include a review of the relevant literature, discussions with the sample of interest and with experts (either groups or individual interviews), and preliminary pilot work. It may be that the study is targeting a pre-defined policy question, or building on earlier research findings, and that these factors present the attributes that need to be considered (or at least some of them). Often, at the outset, large numbers of attributes are identified as potential attributes, but

some are then excluded based on the examination of the question more closely, the review work (e.g. literature, policy documents) and pilot survey work undertaken.

Table 10. Stages in conducting a discrete choice experiment

Stage	Activity (key tasks)
1. Characterisation of the decision problem	Identification of the questions to be considered. Consider whether a discrete choice experiment can/should be used to address the issue. Consider context and how the question / decision problem could be framed, and to whom the survey will be presented.
2a. Attributes and levels	Define the attributes of interest, and determine how they will be presented to the respondents. Assign levels to these attributes (with levels realistic and capable of being traded off against one another).
2b. Identify covariates relevant to attributes	Decide if contextual variables (such as age, income, education) may be important covariates for the attributes identified in the sample to be used for the study.
3. Development of experimental design	Once attributes and levels have been determined, experimental design procedures are used to construct the choice tasks that will be presented to the respondents. There is a large literature on experimental design that details many options for design of choice tasks. For example, consideration of full or fractional factorial design, the method to be used to present choices to respondents, the manner in which choices will be presented e.g. yes/no, A or B, or broader multiple-choice questions.
4. Questionnaire development	Develop a survey instrument, with pre-testing and pilot work a necessary component of the development process.
5. Data collection	Obtain preferences from respondents.
6. Analysis and interpretation	Use appropriate regression techniques to analyse preference data. Interpretation of results in the context of the decision problem – this may involve an estimate of (i) the marginal rates of substitution, (ii) utility scores for given scenarios, (iii) estimate of the probability of choosing specific scenarios.

The attributes must be presented in a manner that is understandable to the respondent and meaningful in terms of the policy problem. Some studies have used a large number of attributes (6 or more), whilst many studies have limited the problem to no more than 5 attributes with between 2 and 4 levels. Each attribute must be assigned at least two levels to represent the range over which the analyst expects subjects to have preferences. This range will be dictated by the research problem. These levels should be plausible, feasible and, crucially, capable of being traded (Ryan, 1996).

The appropriate number of levels for a given attribute is generally dictated by the research question, and the context chosen for the presentation of survey choices. It should be noted though that if only two levels are chosen, linearity between levels must be assumed whereas if there are three or more levels other assumptions are possible. It is important that appropriate measurement units for each attribute are used, whether quantitative or qualitative. Quantitative units (e.g. cost, time and distance) may be more easily interpreted in the analysis of results, but often they do not capture the policy context of important attributes. Qualitative scales may be used to characterise the attributes (often they are the only option for certain attributes e.g. yes or no, or presence of characteristic or not), but particular care is needed when using these due to the issue of interpretation of results; a) depending on how they are expressed, subjects may interpret the levels differently; and b) when comparing attributes, the analyst must take care to point out the marginal unit of change.

When presenting qualitative attributes to respondents there ought to be minimal ambiguity in meaning and connotation thereby seeking to avoid confounding real attribute effects with a variety of other interpretations. For this reason precise communication of the intended meaning of attributes and levels to the respondent is important. This is clearly a factor that can influence the presentation format, and complexity, of the choice experiment.

There is a large literature on experimental design that provides many options and alternatives for designing discrete choice experiments and their specific tasks (e.g. Louviere *et al* 2000, Hensher *et al* 2005). From the attributes and levels identified all possible combinations of attributes and levels can be determined (the full factorial), however in most cases the full factorial is usually very large. For example, using 5 attributes each with 4 levels gives a full factorial of 1,024 combinations. If a complete factorial design is used, all possible combinations of the attribute levels are included in the survey and the probabilistic choice model. The advantage of this is that all effects of attributes on choice can be investigated, and by definition this is statistically efficient. The disadvantage, of course, is that the design results in the number of options becoming impractically large, and in practice it is not possible to ask respondents an unlimited number of questions. In such instances some statistical efficiency must be forfeited and a fraction of the full factorial must be selected for use in the survey. These are purposeful samples drawn from the complete factorial such that certain effects of interest can be estimated independently of one another (but not necessarily independently of unobserved effects). Reduction of the experimental design then comes at the price of having to assume that all unobserved effects are non-significant (i.e. equal zero).

In understanding this issue it is helpful to recognise that estimable effects are typically grouped into 'main effects', which are the individual effects of attributes on choice and 'interaction effects', which account for combinations of attributes varying simultaneously with choice[12]. The number of main effects and interaction terms depends on the number of attributes in the model. If there were 4 attributes A, B, C, D each with 2 levels, this would give a total of fourteen estimable effects: four main effects (A, B, C, D); six 2-way interactions (AB, AC, AD, BC, BD, CD); three 3-way interactions (ABC, ABD, BCD); and one 4-way interaction (ABCD). Under random utility theory (RUT) the total economic value of a commodity is given by this complete set of part-worth utilities.

However, if an assumption was made that a main effects design would be appropriate/adequate, in the above example where only the four main effects were estimated, this could be achieved from a much smaller experimental design. Should important interaction terms be missed out however then the results will be biased. When using a 4 attribute model, where each attribute had four levels (full factorial of 256), applying a design from a published catalogue of experimental designs (Kocur *et al*, 1981) shows that only 9 scenarios are needed to estimate a main effects model but twenty seven to accommodate a further three two-way interactions. Encouragingly, evidence suggests that main effects designs explain over 80% of variance in respondent behaviour (Louviere, 1988).

There have been recent developments in the research surrounding experimental design, with catalogues of experimental designs becoming more helpful and accessible to the analyst (e.g. Burgess and Street, 2003, 2004).

Street and Burgess (2004) have presented work on the 'D-optimality' criterion to produce statistically optimal experimental designs (where the difference between a D-optimal and orthogonal design is that the former formally rules out extreme choice sets, i.e. those containing a dominant or inferior scenario). It is not necessary for the non-specialist user of experimental designs (such as the candidate) to become expert in this aspect of experimental design, but it is important to acknowledge that the designs that are produced by this means have known levels of statistical efficiency, and can provide more precise estimates of model parameters. Prior to this work by Street and colleagues, all choice

[12] Not to be confused with other interaction effects interacting main effects with individual characteristics.

experiments in health economics were designed in the absence of knowledge on level of statistical efficiency (Gerard, 2005).

Issues related to the analysis of discrete choice data have been introduced earlier in this chapter, and will be considered further in the analysis of empirical data collected as part of this thesis. The random utility model is typically used with conditional logistic regression. Results from the estimation of the conditional regression model are accompanied by a standard set of summary statistics which provide important information about the quality of the estimated model. The summary statistics include the maximum likelihood ratio (overall significance), the pseudo R^2 statistic[13] (for goodness of fit), the *Rho* test statistic (significance of correlation, in random effects models), and it is common practice to present the percentage of correct predictions made by the model (as an indicator of model quality).

The size and statistical significance of coefficient estimates (β_i) determine the relative importance of individual attributes. It is important to consider the scale and context of the attributes when assessing relative importance. The sign on the coefficient estimates provides the direction of the effect (increase/decrease in utility with an increase/decrease in the attribute level). In most cases there will be an *a priori* prediction on the expected sign for the coefficients, the direction of effect, and a 'correct' prediction of the sign (+/-) can provide evidence of theoretical validity.

The results from the discrete choice analysis can be used to generate utility scores for scenarios described (using attribute and level combination), to estimate the probability that a particular scenario may be chosen relative to all others in the choice set, and to assess the impact of differences in attribute levels.

Marginal rates of substitution (MRS) between any two attributes can be calculated by taking the ratio of any two coefficient estimates (e.g. β_1 / β_2). This allows important information to be conveyed about the average rate at which respondents are willing to give up a unit of one attribute for an improvement in another. It is important to consider the context of the attributes and the scale used when making such comparisons. If a monetary attribute is included the interpretation of the trade-off can be more easily presented, with attributes

[13] The standard R^2 statistic cannot be used to test overall goodness of fit for logit models, but several alternative 'pseudo' R^2 statistics are used. The most usual is McFaddens R^2 (output from the statistical software STATA labels this the pseudo R^2). This measure compares the proportional difference in log likelihood ratios of a model without parameters and one with parameters. It is scaled between 0 and 1, with explanatory power increasing as the R^2 value increases. Values that lie in the range 0.2 to 0.4 (analogous to 60%-80% in the standard R^2 statistic, Hensher *et al* 2005) are considered a 'good' fit, but there is no commonly accepted threshold value. Bateman *et al* (2002) suggest that analysts should be concerned if the statistic is less than 0.1.

calibrated in monetary units and an indirect willingness to pay (WTP) estimated for given attributes (Ryan and Gerard, 2003).

6.5 Discrete choice framework: Current practice in the application of health economics

The discrete choice framework has developed within health economics greatly over the past 10 years. Commonly referred to as conjoint analysis, it was introduced as an alternative technique for benefit valuation but has developed to become popular in a number of different applications. There have been two specific published reviews of DCEs in health economic applications (Viney *et al* 2002, Ryan and Gerard 2003).[14] Viney *et al* (2002) did not present a detailed review of specific studies. They found the literature relatively undeveloped, with many studies of a methodological nature, used to demonstrate the potential value of the discrete choice framework, and how it could be applied across different settings. Ryan & Gerard (2003) present a more comprehensive review of 34 studies published between 1990 and 2000. More recent personal communications with the authors of this review have indicated that whilst the literature has moved on, their review is still seen as a reasonable reflection of current practice for choice experiments in the health economics literature. Therefore their review is outlined here to summarise the applications of the DCE framework to date. A more recent unpublished review by Fiebig *et al* (2005) updates the review by Ryan & Gerard, providing similar review information on a further 25 studies published between 2001 and 2004. This more recent review reports a similar picture on 'current practice' to that presented by Ryan & Gerard. A short review of social DCEs in the following section helps the reader to supplement the findings from the 2003 review by Ryan & Gerard outlined here.

The 34 studies reviewed by Ryan & Gerard were published in the UK (20), USA (7), Australia (6), and Canada (1). The majority of studies were published during 1999-2000 (20 studies), with 17 presented as applications for economic evaluation, 5 in studies discussing insurance plans, and 12 other contexts (e.g. examining labour supply issues). Preferences were from patient groups in 10 of the studies, from community based samples in 11 studies, from health insurance plan consumers in 5 studies, and other professional and/or clinical groups in the remaining 8 studies. Almost all studies elicited individual preferences, as they applied to the respondents. Two studies elicited social preferences (Farrar *et al* 2000, Roberts *et al* 1999), whilst a number of other studies discussed preferences in a social

[14] Other related reviews have been presented by Hanley *et al* (2003) and Ryan *et al* (2003); these reviews found the literature relatively undeveloped and called for further research into design, data analysis, validity, and interpretation of findings.

context (at least in part), (Scott *et al* 2003, Jan *et al* 2000). In all studies the source of attributes and the levels assigned were judged to be clearly described. Table 11 presents a summary of the characteristics of the studies presented by Ryan & Gerard (2003). The studies reviewed included between 2 to 24 attributes, and the scope of these attributes (e.g. time, risk, health status), were linked to the topic of application. Many studies (19, 58%) included a monetary attribute in order to consider willingness-to-pay estimates for included attributes. In 25 studies (74%) some element of time was used in the attributes (e.g. 16 included waiting time), as this can be used to estimate a monetary value indirectly using a money value for unit of time.

Administration of the survey was via self-complete questionnaire in 27 of the studies; 3 used interview surveys and 3 studies used computerised interviews. Thirteen studies used 8 choices or less per respondent, and 18 studies used between 9 and 16 choices per respondent. Two studies used more than 16 choices per respondent.

The reporting of experimental design was variable across the studies. Twenty-five studies used a fractional factorial design (with design often restricted to main effects), 4 studies used small full factorial designs, but 5 studies did not specify design type. Whilst in the majority of studies (19) software packages were used as the design source, in 9 studies design source was not reported. Four studies used experts to define the experimental design and in 2 studies design catalogues were used. A multiple choice format was used in at lease 29 of the 34 studies (only 5 used > 2 choices), with 2 binary choice studies (yes or no) and 3 studies not reporting detail on the choice set used. Ryan & Gerard graded study design as either 'strong' or 'weak' based on the reporting against the principles of a good design, however, 21 studies were judged to have 'weak' design, and 13 were not graded due to a lack of information – therefore none of the studies were judged to be strong on design.

When estimating econometric models the majority of studies (26 studies) used random effects probit or logit models, allowing for multiple observations from respondents. The consideration of validity was commonly reported across studies. This is not surprising given the methodological focus in the literature on testing the DCE framework and demonstrating its value in a health care setting. Twenty two studies report against theoretical validity, although this may simply be a presentation of the coefficients from models and discussion of the sign (+/-) against *a priori* expectations. Eighteen studies report against rationality of response, which is to be expected given the 'normative' basis of the underlying theory (i.e. how individuals should behave). Fifteen of these studies tested for rationality (or internal consistency) through the inclusion of dominance tests (choice sets where one alternative is

clearly superior) and three studies included transitivity tests. Ryan & Gerard note that there was a lack of research investigating reasons for 'irrational responses'. No studies were identified addressing external validity, i.e. comparing hypothetical choices with actual choices.

Table 11. Background characteristics for DCEs, from review by Ryan & Gerard (2003)

Item	*Category*	*Number of Studies*
Design type	Full factorial	4
	Fractional factorial	25
	Not reported	5
Design source	Software packages	19
	Design catalogue	2
	Expert	4
	Not reported	9
Number of attributes	2 – 3	5
	4 – 5	10
	6	9
	7 – 9	4
	10 or more	6
Number of choices per respondent	8 or less choices	13
	9 – 16 choices	18
	More than 16 choices	2
	Not clearly reported	1
Administration of Survey	Self-complete questionnaire	27
	Interviewer administered	3
	Computerised interview	3
	Not reported	1

(Summary of Tables presented by Ryan & Gerard, 2003)

The response rates to studies varied. Seven studies achieve a 90-100% response (via self-complete format in a 'clinic-like' *(semi-supervised)* setting, or via interviews); 7 achieved 70-89%; 7 achieved 50-69%; and 6 achieved a 30-49%. Lower response rates were reported in studies using postal questionnaires. Five studies did not report response rates. Ten studies asked respondents to indicate level of difficulty in completing the questionnaire; 6 of these studies report between 10-20% had difficulty, and 4 studies report between 20-35% had difficulty, although the studies were not altogether clear whether it was difficulty in understanding or in the choices themselves.

Following their detailed review Ryan and Gerard conclude that discrete choice experiments have been increasingly used in health economics in a number of different contexts (the authors have indicated in personal communications that this pattern has continued from 2000-present).

An update of the review by Ryan & Gerard (2003) has been presented by Fiebig *et al* (2005), extending the studies identified from 34 to 59, with a further 25 studies being published during 2001 to 2004. The update provides evidence of a rapid increase in the use of DCE methods in health. Fiebig *et al* use the same review framework as Ryan and Gerard, providing evidence of a similar pattern of practice, in general, across DCE studies in health. Interestingly, in the more recent 25 studies only 4 of these used general public (community) samples for preferences (compared to 11 of the earlier 34 studies). There was a growth in interview based studies, with 8 of the 25 more recent studies being interview administered, compared to 3 from the earlier 34 studies. There was a move towards fewer choices per respondent in the more recent 25 studies, with 15 of the 25 studies presenting 8 choices or less, compared to 13 of the earlier 34 studies (presenting 8 choices or less).

6.6 Summary review of social discrete choice experiments

From the literature review reported in Chapter 4, and from discussions with experts, 12 social context DCE studies have been identified (7 of these references date from 2003 onwards, 3 of which are unpublished). The characteristics of these studies area summarised in Table 12, and an outline discussion of this literature is presented below. As stated in Chapter 2, the 'social' context is due to the nature of the preferences elicited, with choices over the allocation of resources being for 'others' in society (across different groups in society), and not the allocation of resources against alternative actions for the individual themselves, i.e. offering the social values.

Farrar et al (2000): This study elicited preferences over different dimensions of service developments (health services) from 216 consultants (clinicians) working in one hospital trust. It is amongst the first studies to apply DCE in a social context and this early study suggested that discrete choice modeling had potential as a useful tool in the area of priority setting more generally. In this application the results from the discrete choice study were used in a policy context to score a range of specific clinical service developments, in order to estimate their relative value. Service developments were scored (rated) against decision criteria (e.g. size of health gain, strategy area) and thereafter the developments were assigned a 'benefit score' using these scores multiplied by the coefficients from the discrete choice model presented. The output from the discrete choice model showed all attributes (dimensions) were important in decisions, with all having the expected sign (+/-) and significant at the 1% level. Health gain and evidence on clinical effectiveness were the two

Table 12. Characteristics of 'Social' DCE studies

Study/Context	Design (Type/Source)	Administration method & Sample	Attributes (levels)	No. choices per respondent
Farrar *et al* (2000) Investigates preferences for clinical service developments	Fractional factorial used. 16 scenarios from possible 216 scenarios ($3^3 \times 2^1 \times 4^1 = 216$). Source of fractional factorial was software package SPEED.	Self-complete questionnaire. Sample: UK Hospital Trust consultants (n=216); 130 respondents (60%)	5 attributes: Level of clinical evidence *(A/B/C)* Size of health gain *(large/medium/small)* Contribution to professional development *(improve/no change)* Contribution to education, training and research *Contributes to (0/1/2/all 3)* Strategy area *(none/local or national/local and national)*	8 pairwise choices
Ratcliffe (2000)* Social preferences over distribution of transplant organs.	Fractional factorial used. Fractional design not stated. Full factorial of 108 scenarios ($3^3 \times 2^2$). Source of fractional factorial was software package SPEED.	Postal, self-complete questionnaire. Sample: 800 randomly chosen university employees (UK). (303 completed responses, 38%)	5 attributes: Age *(40/56/60)* Alcoholic liver disease *(yes/no)* Exp length of survival *(5/10/15 yrs)* Time spent on waiting list *(3/6/12 mths)* Re-transplanted *(yes/no)*	8 pairwise choices
Roberts *et al* (1999) Pilot study for study reported by (Bryan *et al* 2002)	as Bryan *et al* (2002)	Structured face-to-face interviews. Sample: General public (UK), convenience sample (n=91), of adults in Health Authority region.	as Bryan *et al* (2002)	as Bryan *et al* (2002)
Bryan *et al* (2002) Social preferences for different health care programmes.	Fractional factorial used, 16 scenarios: from a full factorial of 96 ($4^2 \times 3^1 \times 2^1$). 16 scenarios used in 24 choices. Source of fractional factorial was software package SPEED.	Structured face-to-face interviews. Sample: General public (UK). Random sample of adults in Health Authority region; n=1762 invited, n=909 (51.6%) completed	4 attributes: Number of people *(1/10/100)* Chance of success *(0.1%/1%/10%/50%)* Survival (1 year / 5 years) Quality of life *(For survey: EQ-5D descriptions, For analysis: 1.00/0.893/0.566/0.401)*	8 pairwise choices per respondent. [3 questionnaire formats used (randomly assigned)]
Gyrd Hansen & Slothus (2002) Preferences over willingness to forgo private consumption in order to obtain improved health care services. Alternative health systems presented.	Fractional factorial of 26 scenarios from full factorial of 800 ($2^5 \times 5^2$). Source of fractional factorial was software package SPEED.	Interviews. Random sample, Danish population aged 20-74 yrs. N= 1,895, response rate of 69%.	In summary, 7 attributes (5 with 2 levels [yes/no], 2 [money attributes] with 5 levels): Health system tries to offer all possible treatments irrespective of cost. More screening programmes introduced. Free choice of public hospital. Treatment in private hospitals is subsidised. Focus on preventative measures to reduce life-style related diseases. Extra tax payment per year. Max out of pocket payment for health services.	3 pairwise choices.

Study/Context	Design (Type/Source)	Administration method & Sample	Attributes (levels)	No. choices per respondent
Schwappach (2003) Social preferences in allocating budgets for life-saving treatments.	Fractional factorial used 18 scenarios, from full factorial of 243: [3^3X2^3] Source: SPSS ORTHOPLAN	Computer survey - web based survey (interactive options) (Germany). Sample: Undergraduate students invited to complete the questionnaire (convenience n=154)	6 attributes: Healthy lifestyle *(yes/no)* Socioeconomic status *(lower/higher)* Age *(20/40/60 yeas)* Life-expectancy *(1/10/30 years)* Quality of life *(Very good/Limited/Bad)* Prior receiver (life-saving treatment) *(Yes/No)*	11 pairwise choices (including methods/consistency questions) Graded choice used (% of budget), rather than discrete choices.
Gyrd-Hansen (2004) Social preferences over health states (EQ-5D states).	Fractional factorial, using 42 scenarios (for 23 choices), from full factorial of 243 (3^3X2^3). Purposive selection of scenarios, using previous sample from a research study (MVH Study, York).	Interviews Sample: Random sample of the Danish population, n=3,201 (response rate of 49%).	5 attributes: Used EQ-5D health descriptive system, 5 health dimensions with 3 levels each: Mobility, Self-care, Usual activities, Pain/discomfort, Anxiety/depression.	One pairwise choice.
Gerard (2005) *(unpublished PhD thesis)* Preferences on the principles that should be used to underpin service delivery, for out-of-hours (OOH) services.	Fractional factorial, 18 scenarios. Full factorial of 81 scenarios (3^3). Source: Statistical design catalogue Used 3 different versions to explore impact of perspectives, and compared results (personal vs. social vs. personal/social).	Self-complete postal survey. Sample: General public (UK), n=798 (response rate 36%) [non-random]	4 attributes: Waiting time (15/60/300 mins.) Fairness towards lower social class *(none/limited priority/top priority)* Choice of doctor *(none/2 doctors/3 doctors)* Extra taxation for OOH consultation *(£0/£15/£75)*	9 pairwise choices
Baltussen *et al* (2006) Social preferences of interventions, in priority setting context	Fractional factorial, 24 scenarios. Full factorial of 64 scenarios (6^2). Source: Statistical design catalogue.	Not clear: self-complete (at workshop), or possible interview administration. Sample: Policy makers (Ghana), convenience sample, n=30.	6 attributes: Cost effectiveness *(yes/no)* Poverty reduction *(neutral/positive)* Age of target group *(young/adults)* Severity of disease *(severe/not severe)* Health effects *(high gain for few/low gain for many)* Total budget impact *(high/low)*	12 pairwise choices
Tappenden *et al* (2006) Explores the social values used by the NICE technology appraisal process in the UK	Fractional factorial, 18 scenarios. Full factorial of 108 scenarios (3^3x2^2). Source: SPSS ORTHOPLAN	Self-complete questionnaire. Sample: Members of a health policy/decision making body (UK), n=37 (response rate 45%).	5 attributes: Cost per QALY *(£15k/£25k/£35k)* Uncertainty *(low/high)* Age under *(18/18-64/over 64)* Pre-treatment health status (0.25/0.50/0.75) Availability of other therapies *(yes/no)*	18 binary choices (yes/no)

Study/Context	Design (Type/Source)	Administration method & Sample	Attributes (levels)	No. choices per respondent
Schwappach & Strasmann (2006) Social preferences in allocating budgets Primarily a 'methods' study.	Fractional factorial used – 16 scenarios. Full factorial of 256 scenarios (4^3x2^2). Source of fractional factorial was software package SPEED.	Computer survey - web based survey (interactive options), using a panel of respondents (Germany). Sample: convenience sample, public, n=716 (response rate 72%)	5 attributes: Age (child, teen, employable age, senior) QOL (low to low, low to high, mod to high, high to high) Effect on Life-expectancy (loss of 5-yrs, no effect, gain of 5-yrs, gain of 10-yrs) Frequency of disease (rare, common) Cost of treatment (above ave., under ave.)	4 pairwise choices (plus 3 additional methods choices)

* related publication by Ratcliffe *et al* (2005); note that the abstract by Bayoumi A, Hoch J (2005) offers insufficient detail for inclusion in the above table.

most influential attributes in the study. In this study most respondents rated the tasks between fairly easy and fairly difficult, only 9% found the exercise 'very difficult'.

Ratcliffe (2000): This study asked respondents for a clear 'social' preference, through the use of questions where respondents were asked to allocate 100 donor organs across 2 specified patient groups (specified using differing attribute levels). This question format differed from the typical 'forced-choice' or 'binary choice (yes/no)' format used in discrete choice experiments. Patient groups were described as being equally as ill as each other, and respondents were told that both patient groups would die if no transplant was given. Results showed that, on average, members of the general public would give greater priority to younger patients, those without alcoholic liver disease, those with a greater expected length of post-transplant survival, those who had been on the waiting list for the longest period and patients who are being re-transplanted. Discrete choice data were presented as a model that could be used to predict how respondents would make allocation decisions between two groups on the basis of their characteristics (the study attributes). Results from the early survey questions (non-DCE questions, ranking criteria) indicated the 'capacity to survive and benefit' (expected length of survival) should be the most important criterion. However, when making choices respondents did value all of the other attributes, and of note they did not abandon the less attractive group, i.e. they allocated a proportion of the donor organs to the less attractive patient group, sacrificing some gain in efficiency for what they may have regarded as equity or fairness. Respondents indicated that both groups should be given some chance or hope of receiving a transplant, regardless of situation.

Roberts et al (1999) & Bryan et al (2002): The study by Roberts *et al* is a publication of a pilot study undertaken to inform the study presented by Bryan *et al*. As stated in Table 12 the methods for both studies are very similar, with the pilot study used to test different presentation formats for the DCE questions. The pilot study used the same attributes and levels as the main study, but tested three formats for the presentation of questions (i) presenting full clinical information on the scenarios/choices, (ii) providing clinical information on scenarios/choices when requested, (iii) providing no clinical information on scenarios/choices. Importantly, Roberts *et al* found in the pilot study that the use of clinical scenarios around discrete choices was not necessary, as there was no significant difference in the level of engagement with the study between these three groups/formats.
The study set out to investigate assumptions underlying the QALY maximising approach often considered in priority setting decisions. In pursing this objective it used DCE questions to elicit the preferences of the general public for different descriptions (scenarios) for health care programmes. The study analysis (Bryan *et al*) had three principal components;

investigating response patterns and acceptability of the DCE task, investigating the level of support for QALY maximization in general terms, and discrete choice modeling to provide evidence on the level of public support for assumptions in a QALY-maximisation model. Inconsistent responses were seen in 9.7% of the sample (these tended to be in older responders, and responders from lower social class groups, and those with lower self reported health status). In almost 30% of respondents a pattern of 'non-trading' was noted, with lexicographic preferences against one of the attributes (e.g. 43 respondents for 'number of patients treated', 96 'for chance of success'). Overall only 8.5% of respondents provided preferences consistent with a QALY maximizing objective. In the discrete choice modeling all 4 attributes had the expected sign and were highly significant, indicating the choices made were sensitive to variation in levels across attributes.

Gyrd-Hansen & Slothus (2002): In this study respondents were asked to make a choice between two descriptions of health care systems (A & B), and in each pairwise choice there was a constant reference case (option A, was unchanged). Five of the attributes described specific principles of the relative health care systems, with two attributes presenting monetary levels for tax payments, and additional out-of-pocket payments expected per year.

The authors report that all attributes were significant, and influence choices made between health care systems. The coefficients reported against the attributes were all in accordance with the *a priori* expected sign (i.e. positive or negative). The presence of positive distributive (quality) principles e.g. 'system tries to offer all possible treatments irrespective of cost', increased the desirability of the health care systems presented. Additional financial impact of the systems made the systems less desirable. Self-reported health and age-related characteristics were each reported as interactions in two of the attributes on distributive principles of the health care systems.

The authors conclude that respondents were able to complete the discrete choice survey, and that results generally demonstrated a willingness to increase the quality of the public health care system at the expense of private consumption. However, they do stress a number of caveats over the use of willingness to pay estimates.

Schwappach (2003): This study asked respondents to allocate a fixed budget to hypothetical patient groups competing for life-saving treatments. As in the study by Ratcliffe (2000) this study did not follow the typical discrete choice format. It asked respondents to make a 'graded choice' by allocating a fraction of a total budget, expressed in percentage points of a 100% budget available. Whereas the Ratcliffe study used an allocation of donor

organs, this study used a financial distribution, without being specific about the level of funding for the budget to be allocated.

Discrete choice modeling showed that all attributes were statistically significant (at 1% level). In 6 of the 7 attributes the coefficients had the expected sign (+/-), with the healthy lifestyle attribute having a negative coefficient when a positive one was hypothesised. Respondents favoured larger health outcomes (when considering the impact of quality of life and life-expectancy). They favoured younger over older people for prioritization (regardless of life-expectancy), and favoured patient groups that were not prior receivers of extensive medical care. The respondents favoured patients with a non-healthy lifestyle in allocating the budget, and this preference (direction) was unexpected by the study author, and the reasons for such a preference were subject to some speculation in the discussion of the study findings (e.g. protest vote against current health policy messages).

Schwappach notes that the vast majority of respondents were willing to trade efficiency for a more equal distribution of resources. Respondents avoided 'extreme distributions', with only 3% of decisions made by respondents allocating the whole budget on one patient group. The study suggests that respondents wished to give some resources to both groups, regardless of attribute descriptions.

Schwappach comments on an observed discrepancy between the stated theoretical importance of attributes, obtained via a rating of attributes according to level of importance at the beginning of the survey, and the distributional choices of respondents. For example, the fact that socioeconomic status of patients was stated to be of no or little importance by the majority of respondents, but contributed significantly to allocation decisions. Schwappach also comments that the format of the survey may have been too burdensome on respondents, and suggests that the same framework (although deemed to be reasonably successful in this survey) may not be successful in across a sample of the general population.

Gyrd-Hansen et al (2004): This study investigates the validity of a health maximizing objective in the preferences of a sample of the general public in Denmark. It presents choices between alternative health states, using the EQ-5D to describe these states, and elicits an individual preference and a social preference. The study considers the relationship between individual and social values. Each respondent was presented with one choice between two health states. Respondents were asked to imagine themselves in each of the health states and to indicate which of the health states they found to be the worse of the two.

This rating of the worse of the two states was used in the interview to get respondents to consider two patient groups in each of these health states, and they were asked to prioritise between treatment of these two groups; with the treatment of the worse state moving patients to the other of the two health states, and treatment for the better of the two states moving patients from that state to full health. Respondents were informed that an equal number of patients would be treated in each group (although half of the respondents were then informed that the numbers in the two groups were different), and they were not allowed to be indifferent between treatments (forced choice). Using this approach the study elicited an individual preference and a social preference.

Study results showed a difference in preference patterns and structures between individual and social preferences, although the choice task was different between the two stated preferences (choice between states vs. choice between differences/treatments). The results showed a social preference to treat the worst off group. The study found some variation between preferences across the dimensions of the EQ-5D and notes that although there was an obvious wish to have a more equitable distribution, this argument may not apply with equal force on all dimensions of the health states. The social preferences indicated a wish to relieve patients of extreme problems (level 3 of EQ-5D, for pain/discomfort and anxiety/depression). The study indicated that significant changes were needed (moves from level 3 to 1 on at least 3 to 4 of the EQ-5D dimensions) in the treatment of the better off group to override a general preference to treat the worse off group. The study found no differences in preferences when respondents were informed that the numbers treated in each group would be different. However, it was the numbers in the better off group that would be lower so this may have been overshadowed by a preference to treat the worse off group. Analysis between individual and social preferences suggested that the level of health improvement was relatively unimportant in the decision making process. The study suggests that equity is a priority for respondents who are willing to make substantial efficiency losses as a consequence.

Gerard (2005): This study is one component of an unpublished PhD thesis, it investigates the preferences of a sample of the general public over decision making (health policy) principles (used to underpin decisions). The study is predominantly a methodological one, exploring the differences between the decision making perspective used across three sub-groups of the sample. The study considers a personal, a social and a combined personal and social decision making perspective.

Results are presented for models covering each of the perspectives considered. In all three models more patient choice, less waiting time, and less taxation were preferred. All models showed a similar pattern for the equity consideration (fairness towards lower social class), with responses indicating that there was not a preference for giving 'top priority' to lower social class groups (the coefficient for this was negative, compared to 'no priority'), however, there was a preference to afford 'limited priority' to the lower social class groups (the coefficient for this was positive, compared to 'no priority'). There was no reported difference between the preferences for equity between the personal and social perspectives. In the combined personally inclusive social model the pattern of response was the same, but some differences were noted in the equity coefficients, with a greater priority/preference given for affording a limited priority to the lower social class groups. In this study implicit prices were estimated based on the marginal willingness to pay taxation for an additional unit of other attributes. In this analysis there is a consistent finding across models that the two most important principles were (i) a concern for patient choice, and (ii) a limited concern for equity. Overall the study concludes that the evidence did not support the view that perspective was important, or at least not to any significant extent.

Bayoumi A, Hoch J (2005): The study by Bayoumi and Hoch is currently work-in-progress and has only been presented as a conference abstract (SMDM poster presentation). In this study a convenience sample of patients and their associates completed (self-complete) a computer based DCE survey. The design covered 7 attributes with 2 or 3 levels each, and the context was a hypothetical referenda for new health care programmes (with descriptions covering the average patient to benefit). Methods are not described in the abstract in any detail, and the attributes and levels are not presented in the abstract. The attributes included a quality-of-life scale of 0 to 100 (higher scores are better), baseline life-expectancy, increased life-expectancy, financial endowments (pre-programme), and age, as well as an equity function (from 0 to 100; worst to best). Trade-offs between attributes and the equity scale were used to consider the preferences across programmes. For example, each 1-point decrease in the quality-of-life attribute was associated with an increase in the equity scale of 0.61 points. The study reports that respondents preferred to allocate resources to groups with poor baseline life-expectancy and quality-of-life, the young, those with greatest potential for increased survival, and those who had less access to health care resources. However, detail is not provided on which to assess these findings, and to interpret the results in the context of methods used.

Baltussen et al (2006): This study considers the relative importance of different priority setting criteria. It is a study to consider preferences from a convenience sample of policy

makers in Ghana (Africa). The study was administered opportunistically during a health workshop. The design used was deliberately small, and simple, as the authors wanted to limit complexity and informational overload, and test the framework in a health policy setting. It is one of the only such studies to explicitly use cost-effectiveness as one of the DCE attributes. Alternatives in the choices presented were either 'cost-effective' or 'not cost-effective' (2-levels). Other attributes (Table 12) were selected based on country specific policy context (accepted as relevant and important in Ghana). The survey used paired choices, that were unlabelled, with respondents completing 12 choices.

Interestingly, as with the study presented in this thesis, Baltussen *et al* have sought to use the findings from the discrete choice data modeling to determine a general index for a range of the intervention scenarios in the factorial design. However, their approach is not based on the statistical properties of the full discrete choice approach (model methods and output), and it does not address all of the alternatives in the full factorial design. Instead the study uses a sub-set of the attributes from the DCE and limits the general index to a derivation of the logit style utility function.

Baltussen *et al* present the paper predominantly as evidence that the DCE approach can be used to consider priority setting choices, accepting that the specific study design (attributes) and the findings are not generalisable to other setting such as Europe and USA.

Schwappach & Strasmann (2006): This study considers social preferences for priority setting scenarios. It is presented as a methodological study, examining the issue of reliability, in an internet panel representing a sample of the general public in Germany. Other aspects to DCE methodology are also examined (transitivity and dominance) and the survey instrument is regarded by the authors as a complex design. The main purpose of the study is to specifically examine the hypothesis that respondents (especially in an internet based survey) make quick responses to DCE questions without much thought or deliberation. Using a survey with two stages (i.e. completion at two time points) the study examines preference patterns and reliability, and reports no evidence that 'easy' choices are being made.

Interestingly, the study uses two response formats, a binary response and an 'allocation of points' methodology, although only the latter of these is presented in the paper reported here. In the survey respondents are told to consider themselves as members of a national priority setting committee, that has to make decisions over the introduction of health care programmes where all programmes can not be funded. The respondents are told to

'compare descriptions intensively and to trade off the differences'. The allocation task used asks respondents to allocate a fixed number of 10 available points to each of the two programmes presented in the choices, in order to indicate a strength of preference. There are different types of this allocation (or budget pie) type task in the DCE literature (e.g. see Ratcliffe 2000). Such an approach is thought to be more appropriate in surveys and samples who are thought to be reluctant to discriminate between options (to make a choice). In general the results support the DCE format being used via internet based methods, with moderate to good reliability reported by the authors, and evidence of a low number of preference reversals across time points. However, there was evidence of non-trading behaviour and a suggestion that this could have been due to cognitive overload, caused by the complex survey design.

Tappenden et al (2006): This study has been published as work-in-progress as part of a University of Sheffield (UK) series of discussion papers. It collects discrete choice (binary) data from a number of (n=37) the members of the NICE Appraisal Committee, to explore how five attributes (data on incremental cost-effectiveness, degree of economic uncertainty, baseline health status, age, availability of other therapies) may impact on decisions made by the Appraisal Committee. The choice for the respondent was to recommend the described health technology or not (binary choice of yes/no).

There was a 46% response rate in this specific sample (motivated health policy decision-makers), with a self-complete format used. The nature of the questionnaire may have been appropriate for an informed decision-making group, but it may not be a feasible approach/instrument for use in a less experienced sample, or for members of the general public. Results presented from the discrete choice model indicate that greater levels of cost per QALY, a greater degree of economic uncertainty, and the availability of other treatments were associated with statistically significant reductions in the odds of the health technology being recommended for adoption/funding by the respondents. Age did not have a significant impact on the decisions made (stated preferences), and only where the health-related quality-of-life (HRQL) was at the highest level was there an impact on preference, with higher HRQL (0.75) associated with a reduction in odds of recommending the technology (compared to 0.25, and/or 0.50).

6.7 Summary

This chapter has introduced and described the framework for the discrete choice experiment, covering theoretical underpinnings, modeling methods, stages in the conduct of

a DCE, and current practice. The chapter leads in to the use of DCE in the next chapter to investigate preferences over 'social values'.

The chapter has demonstrated that the DCE approach has been used frequently in the health care literature, with a growth in the number of applications from the late 1990's. The majority of studies have been investigating individual preferences, often for health benefit valuation, but a small number of studies have been used to consider social preferences.

A summary review (above) (Table 12, plus summary text) provides an outline of the methods used and general findings from the 'social' DCE studies reported to date. It does not represent a critical review of the studies, and it is acknowledged that there are limitations with all studies. However, it represents a developing and growing literature on the use of the discrete choice approach for the elicitation of preferences surrounding social choices in a health care context. Only one of the studies identified has taken a priority setting approach in a sample of the general public (Roberts *et al* 1999, Bryan *et al* 2002), and none have used cost-effectiveness (value for money) as an attribute in a sample of the general public. Where Bryan *et al* (2002) have reported a DCE in a sample of the UK general public, the study was primarily a methods study considering the consistency of public preferences with a QALY maximisation hypothesis.

The QALY, or quality of life considerations, have been represented within attribute sets in a number of the social DCEs identified (6 of the 11 studies reported in Table 12). These have been presented in a variety of ways, using either specific qualitative categories of quality of life after treatment (Schwappach 2003; very good, limited, or bad), EQ-5D health state descriptions (which can be linked to derived health state values) (Bryan *et al* 2002, Gyrd-Hansen 2004), movements across quality of life levels with treatment (Schwappach & Strasmann 2006), or a qualitative assessment of quality of life in the pre-treatment health state (Tappenden *et al* 2006). In the study by Bryan *et al* (2002), and the earlier pilot study from Roberts *et al* (1999), the quality of life for each of the DCE scenarios presented to respondents is described using the EQ-5D health state classification (descriptive system). However, for data analysis a quantitative value for health status (health-related quality of life) is mapped to these descriptions.

Level of health gain, or health improvement, is considered in 5 of the 11 DCEs (Table 12), although in some cases it is considered indirectly via quality of life attributes (Bryan *et al* 2002, Schwappach & Strasmann 2006). Age of patient groups is also a frequently applied attribute (5 of 11 studies). Cost effectiveness is addressed in two recent studies, in

contrasting styles i.e. a very general approach, and a very specific approach (Baltussen *et al* 2006, Tappenden *et al* 2006). The recent study by Tappenden *et al* (2006), has considered the availability of other treatments, and as indicated in Chapter 4 this appears to be a relevant attribute for NICE (UK NHS), although not referred to in any detail in the empiricial literature reviewed in this area.

This review of the 'social' applications of DCE methods has been undertaken (and presented) to inform the approach taken for empirical work in this thesis (studies published up to 2004 informed the design of the DCE reported in Chapter 7). Importantly, this chapter is also used to highlight how the empirical study presented in Chapter 7 takes forward an elicitation approach that is currently in its early development for investigating social preferences. This thesis applies the DCE approach, in a policy relevant priority setting context, and makes an empirical contribution, it also makes a theoretical contribution through interpretation of findings in a probability context across the complete factorial design used (it is the first time this has been done to my knowledge in the health care literature). None of the social context DCE studies listed above have reported results using the probability scale, as a transformation of the logit or probit function. The thesis makes further contributions, using the DCE data, through a demonstration of the policy relevance of such data to NHS health policy decision-makers.

7 Empirical Study 2
A DISCRETE CHOICE SURVEY TO EXAMINE PRIORITY SETTING PREFERENCES IN A SAMPLE OF THE GENERAL POPULATION

7.1 Introduction

The current evidence base on social values and preferences, that may be relevant in difficult priority setting decisions, is relatively sparse and undeveloped. Yet, as shown in earlier chapters it does offer some clear insights on potential key social values that may be appropriate when setting priorities. The empirical evidence largely reflects findings against one particular social value at a time, compared to efficiency arguments. In practice, priority setting dilemmas often involve a number of competing arguments. The review undertaken as part of this thesis, and previously published reviews (Sassi *et al* 2001, Schwappach 2002, Dolan *et al* 2005), have highlighted an important need for studies that investigate the relative importance of key social values. The study presented here seeks to do just that.

The discrete choice experiment (DCE) discussed in this chapter was designed to elicit the preferences of citizens living in Southampton (UK) for the funding of health interventions, described using simple generic multi-attribute descriptions of interventions. Discrete choice methodology has been discussed in Chapter 6. Using the DCE framework it is possible to investigate the relative values for each of the attributes used, and to obtain estimates of the utility scores for the alternative health interventions, described using attributes and levels. The results can be used to quantify the potential priority setting preferences (distributive preferences) that may be expected from the general population.

The DCE was conceived around the context of health technology appraisal, used in the thesis as an example of an allocation problem (priority setting dilemma) in health care. The health interventions are described using attributes for differences in the patient groups they treat (either severely affected by their health condition or not), the health improvement they offer (from very small improvements to large improvements), their value for money (cost-effectiveness), and whether the intervention is the only treatment available for the patient group (Table 13).

The study presented here makes a contribution to both a health economics and a health policy literature. It extends the evidence base on the use of DCE methods in surveys

regarding social priority setting choices (i.e. choices regarding the health of others). It considers the use of DCE data to estimate 'strength of preference', over alternative health care interventions, and provides information to assist health policy makers faced with difficult priority setting decisions.

The DCE is used to consider the following research questions:

1. Is it feasible to conduct DCE surveys including social values (such as cost effectiveness) for the settting of health care priorities in samples of the general public?
2. How do members of the general public value the different attributes used in the survey?
3. How can results from the DCE be used to inform decision making?

7.2 Methods

The design and development of the DCE presented here has been informed by the detailed review of the empirical literature on social values in health care (Chapter 4), and the review of DCE studies in health care, specifically of a 'social' context (Chapter 6). The survey instrument has been developed through a series of pre-piloting activities, and through the conduct of a formal pilot study. The final version of the DCE survey instrument used is presented in Appendix 4.

7.2.1 Attributes and levels

The DCE uses four attributes (see Table 13). The selection of these attributes was informed by a review of the empirical literature on distributive preferences (Chapter 4), and policy documents on health technology appraisal in the NHS (discussed in Chapters 2, 3 and 4). The literature was reviewed in detail, including experimental studies, attitudinal studies, and the subsets of literature which presented detailed interview based, or focus group based, qualitative research. As discussed in Chapter 4, much of the literature is context-specific, and represents research that primarily considers methodological approaches for the elicitation of preferences, across a range of social values (in isolation). It is important to consider context. But, a key challenge for the current study was to examine the general relationship between key social values and public preferences for the allocation of resources across wide ranging health technology interventions. This was done using generic attributes and levels and a non-labelled experimental design. The health policy literature provides some guidance on the specific considerations by NICE (UK NHS) in the process of health technology appraisal. The selection of the attributes was also informed by expert opinion,

through discussions with methods experts and health care decision makers, and through explorative pre-pilot work and a formal pilot study of the proposed final DCE methods (see summary below).

Table 13. Discrete choice experiment: Attributes and Levels

Attribute	Description *(summary)*	Level 1	Level 2	Level 3	Level 4
Severity	Whether patients are severely affected by their condition.	Yes	No		
Health Improvement	The average health improvement expected from treatment.	Large	Moderate	Small	Very Small
Value for Money	Cost-effectiveness of treatment - the value for money expected from the treatment.	Very Good	Fairly Good	Fairly Poor	Very Poor
Other Treatment	Whether other effective treatments are available for the patient group.	Yes	No		

Note: See 'showcard' used in the survey in Appendix 4

In choosing the attributes, and respective levels, it was important to consider the fact that the respondent group was to be drawn from the general public. Only a small number of studies (DCEs) have used samples of the general public to elicit preferences (15 of 59 studies reported by Fiebig *et al* 2005), and only a very small number of DCEs have elicited social preferences from the general public over priority setting choices (Table 12). The literature on the feasibility and acceptability of the DCE approach is encouraging (e.g. Ryan *et al* 2001, Ryan & Gerard 2003, Hensher *et al* 2005), and one of the advantages of the DCE approach is that it is able to present choices that are relevant to the respondent (Louviere *et al* 2000). Although the priority setting context for health care is not an everyday setting for members of the general public, the DCE studies by Roberts *et al* (1999), and Bryan *et al* (2002) demonstrated a good level of acceptability in a large sample of the UK general public. Therefore, it seemed reasonable to expect respondents to be able to consider the nature of the choices presented, and to have preferences over the different scenarios presented (and this was tested in pilot study research). However, it was not reasonable to expect members of the general public to be familiar with the language (jargon) of health policy (health services research).

Given the small evidence base available around 'social' DCE experiments, in samples of the general population, it was decided that the number of attributes and levels in this study

should be kept to a minimum. This reflected one of the prime objectives of the study which was to test the methodology in the general population. A further consideration was to demonstrate usefulness of the empirical evidence in a policy setting. With the policy context in mind it was also an objective to keep the presentation of research, and findings, at as simple a level as possible (a parsimonious approach). Given these considerations, four attributes were used, across either 2-levels or 4-levels, resulting in a full factorial experimental DCE design of 64 health care descriptions ($4^2 \times 2^4$). A rationale for each of these attributes is presented below.

As discussed in Chapter 6, attributes need to be described in terms that respondents can relate to, which vary across plausible levels, and most importantly, are trade-able (Ryan & Gerard, 2003). The rationale, definition and description for each attribute and associated levels used in the study are discussed below.

Attributes

Severity: Whether patients are severely affected by their condition

Severity of health, the pre-treatment health state of patients, is identified in the current literature as a social value that is supported by respondents in a number of empirical studies reporting experimental data (Nord 1993, 1995, Dolan 1998, Ubel 1999, Oddsson 2003, Cookson & Dolan 1999, Gyrd-Hansen 2004, Oliver 2004, Shmueli 1999, Edwards *et al* 2003, Wiseman 2005); and attitudinal data (e.g. Myllkangas *et al* 2003 [Finish Kuopio Study], Bowling *et al* 2002). In addition, feedback from decision makers, and from pre-pilot findings, supported the use of severity of health in priority setting decisions. Severity of health condition is not specifically referred to by NICE, in its report on social values, but in the commentary literature around NICE decision making it is indicated that severity of health condition may be a consideration in the decision making process (e.g. Devlin & Parkin, 2004).

There are a wide variety of ways to describe severity of health. In this study it was important to keep the meaning generalisable across health states. Therefore, the descriptive system of the generic quality of life instrument, the EuroQol (EQ5D) (EuroQol Group, 1990), was used. The EQ5D describes health states using 5 dimensions of health (mobility, self-care, usual activities, pain or discomfort, and anxiety or depression) with 3 levels possible for each dimension. These 3 levels are analogous to mild, moderate and severe levels for each dimension. The DCE study used these descriptions to explain the notion of severity. Respondents were presented with the following text on a showcard describing the attributes:

> When considering severity we have judged that on at least one of the following areas patients have severe problems: (i) self-care (e.g. unable to wash or dress themselves), (ii) usual activities i.e. work, study, housework, family or leisure activities, (e.g. unable to perform usual activities), (iii) pain or discomfort (e.g. extreme pain or discomfort), (iv) anxiety or depression (e.g. extreme anxiety or depression).

Two levels were assigned to the attribute, 'Yes' which indicated a description which concerned patients severely affected by their condition and 'No' which indicated descriptions which were not. It was expected *a priori* that health interventions for those patient groups who were more severely affected by their health condition (all else equal), would be associated with a greater level of utility.

Improvement in health: The average health improvement expected from treatment.

The evidence base reviewed (Chapter 4) demonstrated that health gain, and level of health gain, are important considerations for respondents when eliciting preferences for the allocation of health care resources. The attribute was described to respondents as follows:

> Improvement in health refers to the benefits that the patient feels following treatment e.g. improvements in their mobility, improvements in their ability to perform usual activities, reduced pain, reduced anxiety.
> In this questionnaire treatments offer one of the following levels:
> - large improvement in health
> - moderate improvement in health
> - small improvement in health
> - very small improvement in health

It was expected *a priori* that health interventions with larger health improvements (all else equal) would be associated with a greater level of utility.

Whilst this attribute is linked with the attribute of 'value for money', discussed below, it is clear from the empirical literature reviewed, and from the literature around the cost effectiveness of health technologies, that it is possible to have all of the above levels of health improvement, at any of the levels used for value for money. For example a very small health improvement that can be either very good, or very poor value for money. Similarly, it

is possible to have a large health improvement that can be either very good value for money, or very poor value for money (e.g. to save a life, and return a person to full health, but at a cost that is very large i.e. in excess of £1 million).

Value for money: The value for money expected from the treatment.

Efficiency is a well supported motive in the allocation of health resources (e.g. Sassi *et al,* 2001), and the use of cost-effectiveness analysis in health care is now widespread. The UK Department of Health has to ensure that public money is spent wisely and efficiently, with the expenditure on health and social care expected to represent value for money (www.dh.gov.uk). It follows that the NHS places the cost effectiveness of treatment as a prime consideration for both the assessment and appraisal of interventions respectively (e.g. NICE, 2004).

Cost effectiveness is a stated objective in the NICE technology appraisal process (NICE 2004), and NICE has offered some guidance on cost effectiveness, in terms of the cost per QALY (NICE, 2005). However, the terminology of cost effectiveness and the efficiency concept of the cost per QALY, are not common place for the general public, with respect to health care. Therefore, it was decided that, unless a great deal of time could be spent with respondents in order to explain these concepts, and their common use in the NHS, the use of economic jargon was inappropriate. Rather the term 'value for money' was used to express the notion of cost effectiveness and efficiency. This term was regarded as a commonly understood term, and very much related to efficiency and cost effectiveness. Public health decision makers indicated that the terminology of value for money had been used successfully when dealing with the public. It was also tested in pre-pilot work, the formal pilot study, and all indications were that value for money was a reasonable and acceptable term to use to elicit preferences surrounding efficiency.

In the DCE, respondents were presented with the following text on a showcard describing the attributes:

> 'Value for money' is a common consideration within the NHS. Value for money is about the efficient use of resources (e.g. doctor's time, hospital beds, healthcare funds).
>
> In this questionnaire we use the following levels:
>
> - very good value for money
> - fairly good value for money

- fairly poor value for money
- very poor value for money

It was expected *a priori* that health interventions with better levels of value for money (all else equal) would be associated with a greater level of utility.

Other treatments: Whether other effective treatments are available for the patient group.

This attribute needed to reflect the importance of general circumstances surrounding the availability or otherwise of alternative treatments, when considering the allocation of resources to a health technology being appraised. The literature does not inform on this attribute in any detail. However, there is support for the use of this attribute from a mixed-methods study undertaken to investigate the use of economic evaluation in the process of health technology appraisal in the UK (within NICE), and the health technology appraisal process more broadly (Williams *et al,* 2005). Williams *et al* state that the availability of alternative treatments, or not, was a potential modifying factor in the health technology process witnessed at NICE (Williams *et al,* 2005). A recent DCE from Tappenden *et al* (2006) has also stressed the potential value of this attribute, and the study used this form of attribute in a priority setting DCE study (see outline detail in Chapter 6). A firmer foundation for the use of this attribute, and its potential importance, is found in the health policy literature (discussed in Chapter 4). When reviewing the published NICE technology appraisal guidance, the availability of other treatments, or not, was information frequently presented in the final reports of the appraisal considerations. For example, with reference to the appraisal of imatinab for the treatment of chronic myeloid leukaemia (NICE 2003), Rawlins and Culyer (2004)[15] state that although the cost-effectiveness of treatment in this case was outside of the range usually acceptable to NICE, "in the absence of *any effective alternative treatment* ... imatinab was considered to be cost effective in the chronic phase after interferon alfa" (p225, italics not in original).

Whilst there have been few empirical insights into how to present the notion of available 'other treatments', in general the use of this attribute in a priority setting context seems to be supported intuitively, from discussions with decision makers and from pre-pilot feedback, but also from the developments within NICE surrounding guidance on social values judgments. NICE have, in earlier consultation documents, considered the specific issue of 'previously

[15] Professor MD Rawlins has been Chair of NICE since its inception in 1999. Professor Culyer was Vice-Chair of NICE from 1999-2003.

untreatable conditions' (NICE, April 2005), and whether special consideration should be given in such cases. No formal statement on this issue has been made, but the NICE Citizens Council are expected to consider the topic of 'orphan treatments' (i.e. a treatment for a disease for which no alternative curative treatment for patients exists) in the later part of 2007.

In the DCE respondents were asked to consider this attribute at a simple two-level 'yes' or 'no', i.e. whether other effective treatments were available or not. Respondents were presented with the following text on a showcard describing the attributes:

> Whether other effective treatments are available:
> Yes available or No not available
> We assume that all patients will have usual and best supportive care available within the NHS (e.g. GP services, specialist outpatient appointments, and best supportive nursing care). Where referring to treatment options we are thinking about drugs, surgery, or specially developed services (such as physiotherapy, support services, occupational therapy, specialist education, preventative therapies, etc).

It was expected *a priori* that health interventions introduced when there were no other effective treatments available (all else equal) would be associated with a greater level of utility.

7.2.2 Framing of the choices

The choices were framed in the context of an NHS priority setting choice/dilemma. Respondents were introduced to the fact that decision-makers in the NHS are often faced with difficult priority setting decisions. They were asked to consider a situation where they were in the position of a health care decision maker and were faced with some of these difficult decisions on how best to use its limited budget for the provision of health care services. In each of the DCE questions respondents were asked (in the context of health care decision-maker) to consider two treatment options where there was only funding to support one of the options ("Given that only one of the options can receive funding, which

Table 14. Example of DCE question

Q:	SHOWCARD A health care decision maker is faced with difficult choices on how to allocate its budget. Imagine a choice where there are two options for the use of available funds. Given that only one of the options can receive funding, which option would you support?
Option K	Patients are severely affected by their condition. With treatment the average patient has a very small improvement in their health. The treatment is regarded as being very poor value for money. There are no other effective treatment options available.
Option L	Patients are not severely affected by their condition. With treatment the average patient has a large improvement in their health. The treatment is regarded as being very good value for money. There are other effective treatment options available.

Option K	1
Option L	2
Other	3
Don't Know/Unable to choose	4
None	5
Refused	6

Figure 6. Example of Survey Showcard (for the above question)

Health condition: Long-term health condition

	OPTION K	OPTION L
Severity of patients	Severely affected	Not severely affected
Improvement in health	Very small improvement in health	Large improvement in health
Does the treatment offer Value for money	Very poor value for money	Very good value for money
Other effective treatment options available	No	Yes

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

option would you support?"). This framing of the DCE questions was similar to that used successfully in a sample of the general public by Bryan and colleagues (2002) in their study examining the validity of the QALY maximisation objective.

The full interview format (versions 1 & 2) and the showcards used are presented in Appendix 4. An example of a DCE choice is presented in Table 14, with the related showcard in Figure 6. The survey used a multiple choice format, where respondents were asked to choose between two descriptions. The format was a forced-choice, with no alternative openly offered to the respondent, although interviewers had categories available to them for cases of 'non-response', as in Table 14.

Choices were presented in a general 'unlabelled' format, to reflect the research questions being examined, and the general policy context of the DCE.

7.2.3 Experimental design

This study uses recent and widely supported recommendations on experimental design for DCE studies (Street and Burgess, 2003, 2004). These recommendations stress the requirement of orthogonality, balance and statistical efficiency, in the design of DCE studies. Appendix 5 presents detail on the experimental design methods used. In summary, there are 4 attributes in the design, two each with four levels and two levels, therefore the full factorial is 64 combinations of attributes and levels ($4^2 \times 2^4$). A suitable main effects fractional factorial design was used from a web-based catalogue (Sloan, 2003). This entailed the use of a resolution 3 fractional design with 16 scenarios, which Street and Burgess (2003, 2004) have proved to have the properties of near optimal (94.4%) statistical efficiency for parameter estimation for a model of main effects. This statistical efficiency is maintained with the pairing of scenarios into choice sets, using the 'foldover' design recommendations of Street and Burgess (2004). See Appendix 5 for further detail.

7.2.4 Main survey interview schedule

The interview schedule for the main survey comprised:

i) Introduction by interviewer
ii) Background provided on attributes to be used (at this stage interviewees not aware of the DCE format), including presentation of the attribute showcard
iii) Respondent ranked the 4 attributes from most important to least important
iv) DCE question format introduced
v) Example (warm-up) DCE question

vi) Consistency question *[where respondent chose the dominated option they were asked why they had done so]*
vii) 8 DCE questions presented (from the 16 combinations in the efficient design set, see above)
viii) Respondent asked question on the difficulty of the questions / task
ix) Respondent asked the priority setting question (severity/disadvantage; as Chapter 5)
x) Collection of background data on socio-demographics
xi) End of Interview

As above, the full DCE interview format (versions 1 & 2) and the showcards used are presented in Appendix 4. The target time for interviews was between 25-30 minutes.

7.2.5 Development of the survey instrument and the interview schedule

Pre-pilot work: Pre-pilot research involved informal testing of early formats of background text to be used, framing of questions, potential attributes and levels. Pre-pilot interviews were undertaken in an informal way in early stages of development, and thereafter in a series of full pre-pilot interviews (in people who were not associated with the NHS or health services research i.e. through occupation or family members, with interviews including some elderly people aged over 70-years). Pre-pilot work also involved discussions with experts (methodological specialists, clinical specialists in public health, and health policy makers). Pre-pilot work considered (and rejected) a range of potential attributes, including (i) a contextual attribute for life-threatening (acute) conditions versus longer term chronic conditions, (ii) treatment related descriptions for life-saving versus life-enhancing, (iii) potential attributes to describe the values used by a decision-making committee (i.e. different moral positions – emphasis on efficiency, versus emphasis on equality). Attributes for budget impact, and the numbers of people treated, may be relevant for other decision making contexts, but were deemed inappropriate when wishing to inform on the NICE health technology appraisal process. Pre-pilot work also considered different terminology and text that could be used in the survey (e.g. use of 'value for money' rather than cost-effectiveness, use of 'health improvement' rather than health gain, consideration of use of QALY information), and it was very helpful in determining the level of background and introductory text that could be used in the interview schedule (within a target time of 25-30 minutes per interview), and the number of DCE questions that could reasonably be presented to respondents to answer in the time available.

A previous DCE study published by Bryan *et al* (2002) (see Chapter 6), and discussion with one of their research team were helpful in determining the nature of the final survey

instrument. The study by Bryan *et al* was an interview based study, examining the validity of a QALY maximising decision rule. It used face-to-face interviews with a sample of the general public (data collection by MORI). In pilot work undertaken to develop their survey instrument (published by Roberts *et al*, 1999) Bryan and colleagues found that the use of clinical descriptions to accompany a fairly general DCE scenario (using only attributes and levels) was not required (it made no difference to response data). This supported the use of general DCE descriptions in the current survey.

Early DCE design was more complex than the final version (where only 4 attributes are used, across either 2 or 4 levels), but pre-pilot work suggested that a simple format was preferable, especially as this was one of the first such surveys (e.g. priority setting survey including cost-effectiveness / value for money) in an interview based sample of the public. The survey was restricted to the 4 attributes used, even though additional attributes may be of interest (i.e. differences in preference between life-saving treatments versus life-enhancing ones, or treatment for acute versus chronic conditions), as the survey used was still expected to provide valuable information on the feasibility of methods and also had the potential to provide useful preference data. In pre-pilot work there seemed to be a strong focus by respondents on any options where life-saving treatments or life-threatening conditions were stated, regardless of other factors or considerations. Although respondents in informal discussions were aware of the need to take the alternative descriptions into account and to trade-off attributes regardless of these contextual labels, they indicated that they found it difficult to do so in the questions when the context was directly stated (it seemed they felt 'duty bound' to select these options, regardless of any wish to treat the alternative patient group). The final survey avoided the use of text such as life-saving or life-threatening, although this is acknowledged as a potential limitation, and suggested as an area for future research.

The final survey instrument used text on showcards to indicate that the respondent should consider "longer-term health conditions". This was determined via pre-pilot work, with full interview schedules undertaken where this additional information was provided on the showcards and in other instances where it was not provided. Discussions with pre-pilot respondents (and success in the formal pilot study) led to the use of the text in the final survey.

Pilot Study: A formal pilot study was undertaken using a 'near final' version of the questionnaire (using 8 of the 16 DCE combinations from the experimental design used – as above). A total of 25 full face-to-face in-home pilot interviews were undertaken between 15th

and 19th September 2005 (by MORI interviewers). Two versions of the questionnaire were used in the pilot study, each containing a ranking task, DCE test (warm-up) question, DCE consistency question, 8 DCE questions, one (of two) versions of the priority setting question (discussed in Chapter 5), questions on the level of difficulty of the questions/tasks, and collection of socio-demographic data. Different text was used in each version of the pilot survey to gather information on description of levels on the value for money attribute. Other key objectives of the pilot study were to time the interviews, and to test respondent understanding of the questions presented. Pilot interviews were in a convenience sample of the Southampton general public. No formal random or quota sampling was used, and interviewers were asked to use their expertise and local knowledge of Southampton to select appropriate area(s) / street(s) in the city in which to conduct pilot interviews. Interviewers were asked to achieve a good spread of interviews covering age, gender, work status and social grade. Interviewers were provided with a detailed briefing note, and following the pilot survey interviewers were debriefed by the MORI project executive. The pilot survey was judged to be a success with very few reported problems or suggestions, and only minor edits for the final survey instrument (small text changes/edits only).

7.2.6 Sample

The main survey instrument was administered on a sample of the adult general population (aged 18 and over) in the Southampton (UK) City Council area. Interviews were carried out in a face-to-face format 'in-home' by the MORI Social Research Institute between 19th September 2005 and 9th October 2005.

A random location quota sampling approach was used (as in Chapter 5). This sets fixed quotas of people to be interviewed in a number of randomly selected sampling points. Sampling points were based on 'Output Areas' (OAs) in the Southampton City Council area, the smallest building block of the Census. For the sample used in this survey 32 OAs were randomly selected (by MORI) proportionate to population size, controlling for socio-demographic composition. Quotas were set by MORI – individually at each sample point – to reflect the socio-demographic profile of Southampton City Council residents, on gender, age and work status.

The candidate developed the survey instrument, and the agency was briefed in detail on the nature of the survey instrument, its design, format, and the study questions that were the motivation for the survey. The candidate worked through the formal pilot study in detail with the data collection agency, providing a full and complete survey instrument, and through the testing of the survey instrument and interview schedule in the formal pilot study. Thereafter,

specific interviewers were briefed by the agency (MORI) on the survey instrument and the interview schedule – covering survey aims and background, contact procedures, detailed explanation of the questionnaire. Interviewing was conducted by MORI's own fieldforce of face-to-face interviewers. Interviewing was subject to MORI's in house quality assurance procedures, including a back check of at least 10% of all interviews by telephone during fieldwork. All MORI interviewers had extensive experience, and were members of the Interviewer Quality Control Scheme (IQCS) and were recognised by the Market Research Society (MRS). The candidate received a sample of the completed questionnaires from the final survey for quality assessment and auditing of the data set provided.

7.2.7 Assignment of questionnaire

There were two versions of the main survey instrument, each contained 8 of the 16 DCE combinations from the efficient experimental design used. Respondents were randomly assigned one of the two main survey instruments.

7.2.8 Sample size

There are no formal sample size calculation methods for this type of public preference study. The nature of the preferences and the level of variation in responses expected in this survey were difficult to predict, therefore an arbitrary judgment on sample size was necessary. The use of face-to-face interview methods was a limiting factor in the sample size used here. The survey aimed to interview 250 people. This target sample size was determined on the basis of available time and resources. Pearmain *et al* (1991) have reported that for DCE studies sample sizes circa. 100 are able to provide a basis for modeling of preference data. Respondents were not offered any financial, or other, incentive to participate.

7.2.9 Methods of analysis

The survey findings inform on the initial ranking of attributes, the discrete choice data, the data collected on the degree of difficulty respondents found in understanding the DCE questions, and in responding to them. Each of these will be reported separately. Ranking data and data surrounding difficulty with the DCE questions is reported using simple descriptive analysis, applying *Chi*-Squared statistical tests for differences between groups. Methods for analysis of the DCE data are outlined below.

7.2.9.1 DCE data analysis

The DCE data are modelled using a random utility maximisation framework (Louviere *et al* 2000, Hensher *et al* 2005), and STATA 8.1 software. A series of random utility models are fitted to the data (main effects, main effects with interactions) and results presented. Data is binary choice data, with predicted values constrained to either 0 or 1: the value 1 represents the option being chosen, with 0 where not chosen. Classical linear regression models are not suited to this form of data. As discussed in Chapter 6, a class of model known as categorical dependent variable models (CDVMs) are used to model binary choice data. Logit and probit models are the most widely used models in the case of binary dependent variables. Both logit and probit models are used to model the DCE data in this study. CDVMs adopt the maximum likelihood estimation method, which requires assumptions about probability distribution functions. Logit models use the standard logistic probability distribution, while probit models assume the standard normal distribution. In a binary choice setting the logit and probit models, whilst working on different scales, are expected to provide similar conclusions (Greene, 2003).

In the current analysis the logit model is chosen *a priori* to be the most appropriate, given the decision making context; (i) the convenience of using logit regression coefficients to directly reflect the impact and relative effect of each attribute (level) e.g. logit regression coefficients reflect a log of the odds and transform directly into an odds ratio, (ii) and as the most appropriate approach for the estimation of probabilities, through the use of the conditional logit model and the 'S-shaped' logistic probability distribution (Greene, 2003). This latter point is based on the use of the DCE data to estimate probabilities across the full factorial of 64 scenarios, and the potential for some probabilities to be very small and therefore located in the wider tails of the logistic distribution. Whilst the probit model has been more widely used in the health economics literature this may reflect the fact that the majority of studies have sought to estimate the elasticities between attributes (i.e. ratio of regression coefficients), rather than consider the direct impact of attributes on the utility function, and the relative probabilities across the factorial design employed. Details on the analytical approach of the logit model have been discussed in Chapter 6.

The logit model involves the restrictive assumption of homoscedasticity (discussed in Chapter 6), and it is desirable to test the validity of this assumption when using the logit model. To relax the assumption a random parameter model can be used (Adamowicz, 2003). In the current analysis, although the logit model is seen as the preferred approach, the assumption is relevant to the data, used in the logit model, and it is tested using the probit model in a fixed-effect and random-effect form, to consider the difference if any

between the two approaches. To further assess the modeling methods used in the study, a third party (KG) used a different statistical software package (LIMDEP) to replicate the analysis undertaken using the STATA software.

The model set out in the design of the study is of the form:

$V = \beta_0 + \beta_1$ SEVERITY $+ \beta_2$ HEALTH IMPROVEMENT $+ \beta_3$ VALUE FOR MONEY $+\beta_4$ OTHER TREATMENTS

where the deterministic component of the utility function (V) is a function of the attribute levels between options and where the coefficients (part-worth utilities) β_1 to β_4 and constant β_0 are estimated in the model. This is a linear in parameters utility function, which has typically been used in the DCE studies in the health economics literature.

Given the way the attributes are coded, the constant (β_0) is used to reflect what is expected to be the least desirable option in the factorial design, across attributes and levels. That is, it captures the attribute levels (through use of dummy variables) that are expected to be the least desirable for each of the attributes. This worst case, or least desirable scenario, defines an intervention for 'non-severely' affected persons, where 'other treatments' are available, where the health improvement is 'very small', and the value for money is 'very poor'. The effect on utility (the desirability) of other scenarios will be relative to this base case option (constant). Dummy variables are used (n-1 dummy variables for each attribute) to account for this approach. This is helpful in interpreting the findings of the study. In the analysis undertaken results are investigated using an alternative approach (applying effects codes instead of dummy variables) to examine the component parts of the constant (reference case). This work, using effects codes, supports the use of dummy variables, as it demonstrates that the attribute levels are non-linear with respect to utility (i.e. not a constant linear relationship across levels 1 to 4), (see Appendix 6).

A priori the coefficients from the discrete choice model are expected to have a positive sign, indicating an increase in utility (probability of being chosen for funding) with increases in the level of attributes, from the base case scenario. Where *a priori* relationships between choice and independent variables (attributes) hold, this is a good sign of theoretical validity. If the attributes and the constant are important the coefficients estimated will show statistical significance, and therefore have an influence on respondents' likelihood of choosing the option. Coefficients are used to consider the relative importance per unit of change in the attributes. It is important that this is done carefully, and in the context of the qualitative changes in the attribute levels used.

Data from the logit model are transformed into a set of probabilities, to show the relative probabilities of being chosen/supported by the general public (based on the sample used) for each of the 64 scenarios of the full factorial experimental design. This approach is consistent with the conditional logistic regression approach, discussed in Chapter 6 (see 6.3.4). The probability model is non-linear in parameters (using the 'S-shaped' logistic distribution).

In the current analysis the model is used to consider main effects. It is important to consider *a priori* where subgroups, defined according to sample characteristics, are thought to have a potential impact on choice. These subgroup impacts are termed interaction effects, i.e. the characteristics interact with the attribute or main effects. A common consideration in discrete choice models is effect of the income of respondents on the choice made, especially where attributes consider willingness to pay. However, *a priori* income level was not considered to be an interaction term in the current analysis as the choice is unrelated to respondent income. In the analysis here, *a priori* there is a view that the age of respondents, their health status and their experience of illness, may have an impact on the response data (interaction effect). These factors have been identified in previous DCE studies (see review in Chapter 6) as potentially important. These factors are investigated using sub-group analysis, applying a log likelihood ratio test.[16] Where this indicates a significant interaction further investigation is carried out in the discrete choice model.

In the current analysis a fixed-effects conditional logit model is used (in STATA 8.1 software) for the presentation of results. However, the probit model is also used to explore the panel data and the validity of the homoscedasticity assumption; a fixed-effect and random effects probit model are estimated. In the analysis undertaken summary statistics are presented to provide information on the quality of the estimated model; these statistics represent a standard set of measures to reflect overall model significance (likelihood ratio test), goodness of fit (pseudo R^2 e.g. McFaddens R^2), and predictive capability.

[16] Although the analysis here has used a simple sub-group approach to investigating interaction effects, other techniques are available. For example, Scott *et al* (2006) have recently demonstrated the use of a backward stepwise regression technique to explore interaction effects.

7.3 Results

7.3.1 Response rate

The DCE survey was presented to a sample of 263 respondents. There were 4 respondents from this sample who did not provide DCE responses for any of the 8 DCE questions, indicating 'don't know' (or with response not stated) on the questionnaire. These 4 respondents are excluded from analyses. The useable sample included 259 respondents.

In the survey each respondent was asked to complete 8 DCE choices from the experimental design used, giving a potential data set of 2072 pairwise observations (n=259). As shown in Table 14, whilst respondents were faced with a 'forced choice' format between the two options described, the interviewer was also able to record a number of other response categories where a respondent was not able to make a choice between the options presented. A small number of respondents (n=26) did not provide a preference between the two options presented in each of the 8 questions, and the resulting data set was 2027 observations (98%). Data were missing to some degree from 26 respondents. In 16 of these cases respondents had 1 data point missing (1 from 8), 7 had 2 data points missing, and 3 respondents had 3-5 data points missing. In all but three cases (choices) the missing data were recorded as "don't know/unable to choose".

In response to a 'consistency check' question in the survey (presented after a 'warm-up' question, but before the main DCE survey questions) 12 respondents 'failed' this question, choosing the option that was dominated across all attributes by the alternative available. In two further cases the response data stated "don't know" or "not-stated". Therefore in 5% of cases (14/259) the sample did not answer the consistency question as expected. Although such respondents were asked to explain the reason for their choice the qualitative data collected did not prove helpful. The data from these respondents has been included in the analysis.

7.3.2 Characteristics of the sample

Table 15 presents a summary of the socio-demographic characteristics of the sample. These data indicate a reasonably representative sample of the Southampton City Council geographic area (Southampton City Council, 2004), as discussed in Chapter 5.

The sample did have a larger proportion of retired and home workers, than in the population at large, which may be expected in an 'in-home' interview survey of this type. There were no

Table 15. Sample characteristics (numbers) by questionnaire version

	Number of People	%	Questionnaire (numbers)		
			Version 1	Version 2	
Total (number)	259	100	128	131	*Chi-sq*
Number by:					
Gender					*0.587*
Male	117	45.2	60	57	
Female	142	54.8	68	74	
Age					*0.114*
18-34	90	34.7	45	45	
35-54	89	34.4	38	51	
55+	80	30.9	45	35	
Social grade					*0.443*
AB	36	13.9	19	17	
C1	68	26.3	30	38	
C2	65	25.1	35	30	
DE	84	32.4	40	54	
Work Status					*0.296*
Working full time	92	35.5	44	48	
Working part-time	41	15.8	18	23	
Retired	60	23.2	36	24	
Other	66	25.5	30	36	
Household income					*0.249*
Below £17,500	91	35.1	51	40	
£17,500 - £29,999	46	17.8	23	23	
£30,000 and above	54	20.9	21	33	
Refused/not stated	68	26.3	33	35	
Ethnicity					*0.441*
White British	218	84.2	110	108	
BME[1]/other	41	15.8	18	23	
Home ownership					*0.212*
Owner occupier	147	56.8	68	79	
Social renter	60	23.2	32	28	
Private renter	46	17.8	25	21	
Illness/disability					*0.430*
Yes[2]	103	39.8	58	53	
No	148	57.1	70	78	
Household composition					*0.109*
With children	88	34.0	35	49	
Without children	170	65.6	89	91	
Health in general					*0.981*
Good or very good	185	71.4	91	94	
Fair	46	17.8	22	24	
Bad or very bad	24	9.3	12	12	
Health Insurance					*0.190*
Yes	36	13.9	14	22	

[1] Black and minority ethnic
[2] Respondent and/or someone else in household

differences in socio-demographic characteristics across the two versions of the questionnaire used (Chi^2 tests) (Table 15).

Table 16. Percentages of respondents ranking attributes 1st to 4th

Attribute	Rank 1st	2nd	3rd	4th	*Mean Rank*
Severity of Patient Group	52.7	18.4	18	10.9	*1.87*
Health Improvement	40.3	41.1	14	4.7	*1.83*
Value for Money	9.8	13.7	21.1	55.5	*3.22*
Other Treatments Available	8.6	24.7	46.7	20	*2.78*

Note: it was possible to have equal ranking for attributes

7.3.3 Ranking data

Table 16 presents the findings from the ranking of attributes by respondents (ranked in order of importance 1st to 4th). The ranking task was primarily a warm-up task to ensure respondents had familiarity with the attributes/levels used, and to offer an opportunity for respondents to ask questions on the attributes where there may have been any misunderstanding. There were no statistically significant differences (Chi^2 test) in ranking by questionnaire version (1 & 2), or by respondent characteristic (e.g. social class, work status, household income), other than for ethnic status. With ethnic status there was a significant difference in the mean rank for severity of health (Chi^2 p-value=0.01), however there was a large difference in numbers in each of these groups (n=216 vs n=40).

Table 17. Difficulty understanding and answering DCE questions (self-reported)

Level of Difficulty with Questions	Difficulty "Understanding"	Difficulty "Answering"
Very difficult	10.8%	20.8%
Fairly difficult	29.7%	47.5%
Not very difficult	31.7%	23.6%
Not at all difficult	26.6%	6.6%
Not stated/missing	1.2%	0.8%
Total	100% (n=259)	100% (n=259)

7.3.4 Data on Respondent difficulty with DCE

Table 17 presents the data on respondents self-reported difficulty in understanding and answering the DCE questions posed. The majority of respondents reported that they found the questions not very or not at all difficult to understand. The majority of respondents did report that the questions were very or fairly difficult to answer. This indicates that respondents did engage with the choice context, and suggests that they did weigh up the difficult choices presented, offering some confidence in the face validity of the experiment.

There were no statistically significant differences (Chi^2 test) in responses to the difficulty questions by questionnaire version (1 & 2). Generally there were no significant differences in response to difficulty questions by respondent characteristics. There were some significant differences in understanding the questions by age, with the elderly likely to find questions more difficult to understand, and differences by self reported health status. For health status groups (self-reported), those in bad or very bad health were more likely (16 of 24) to have difficulty understanding the questions (p-value=0.011).[17] There were no statistically significant differences (nor a tendency towards a difference) between subgroups in terms of difficulty answering the questions.

7.3.5 DCE data analysis

As discussed above, *a priori* the conditional logit model (CLM) was the preferred modeling approach. Results from the main effects CLM are presented in Table 18. Model A reports findings for a main effects only model, whilst Model B includes one significant interaction variable.

To investigate interaction effects, sub-group analysis was undertaken using the likelihood ratio (LR) test, in the CLM. Subgroups of interest were age, experience of illness, and self-reported health status. The rationale for these interaction effects has been discussed above. For age and experience of illness subgroups analysis showed no significant difference between findings. For self-reported health status there was a significant difference (in sub-group) and those respondents who stated a bad health status (either bad or very bad) appeared to place less weight on the 'other treatments' attribute, (the attribute indicating that no other effective treatments were available). The number of respondents with a self-reported health status of bad/very bad was small, only 9.3%. This reflected a similar proportion to that reported in the Southampton City Council area in the 2001 Census (8.6%). There were no significant interaction effects when health status was considered against the

[17] This may, in part, explain the interaction effect noted in Table 19.

other attributes in the model. The introduction of the interaction effect (bad health status *x* other treatment) to the CLM showed it had a significant impact, and it improved the specification of the model. When introduced to the model (Model B) it resulted in the attribute for 'other treatments' reaching statistical significance (p=0.029), and it marginally improved the robustness of the model (psudeo R^2 from 0.2051 to 0.2066 in Model B).

To test for the absence of serial correlation (the assumption on homoscedasticity), relevant for the CLM, and to further explore the data, a fixed-effects and random-effects probit model were also used to analyse the data (using STATA 8.1). Results from the probit analyses are presented in Appendix 6, and show no difference between the fixed and random-effects model. The probit analysis finds no statistically significant differences in the the correlation parameter (*rho*). These results indicate the absence of serial correlation in the multiple choices made by each respondent (Appendix 6), offering some reassurance of the appropriateness of the conditional logit model for the current study. The coefficents in the probit model are different to the logit (as expected), however the general relationship between the attributes is similar. The differences between health care scenarios (factorial design) using the logit and probit models (functions) were very similar on the intervals scales used, although each approach had differing endpoints (Appendix 6).

The rationale for preferring the CLM over the probit model has been discussed above. Data analysis provided evidence of the acceptability of the assumption of homoscedasticity, in this instance, and also showed a superior psudeo R^2 statistic (goodness of fit) for the CLM, indicating the CLM was a better fit than the probit model (0.2066 vs. 0.1350). On this basis the conditional logit model, with the one interaction effect (Model B in Table 18), is regarded here as the most appropriate to use in the discussion of the discrete choice data. Therefore this is the model used in the remainder of this chapter, and for inference thereafter.

As shown in Table 18, all main effects coefficients are positive (as per *a priori* expectations), showing an increase in utility (desirability) where there is a change in the attribute level, from the base case. Where dummy variables have been used for health improvement and value for money, the staging of the impact is as expected over the incremental changes in the attribute (e.g. $\beta_3 < \beta_4$). All of the main effects except the 'other treatment' attribute are statistically significant (at the p=0.01 level or below).

The findings presented in Table 18 provide support for the theoretical validity of the model. When an interaction is introduced for health status, the 'other treatment' attribute is also statistically significant (p=0.029). The interaction effect against self-reported health status

(and the 'other treatment' attribute) is introduced to the model using the proportion of persons in the sample/population with self-reported 'bad' health (i.e. coefficient multiplied by proportion).

Table 18. Discrete Choice Model Results (CLM)

Input / Attribute (level)	MODEL A Main Effects Coefficient	MODEL B Main effects with interactions Coefficient (OR)
Constant (β_0)	0.2689 *	0.2731 * (1.31)
Severely affected (β_1)	0.5393 *	0.5314 * (1.70)
No other treatments (β_2)	0.0781	0.1243 ** (1.13)
Small health improvement (β_3)	0.4917 *	0.4959 * (1.64)
Moderate health improvement (β_4)	1.0428 *	1.0443 * (2.84)
Large health improvement (β_5)	1.3773 *	1.3756 * (3.96)
Fairly poor vfm (β_6)	0.2889 *	0.3150 * (1.37)
Fairly good vfm (β_7)	1.0121 *	1.0314 * (2.80)
Very good vfm (β_8)	1.1655 *	1.1744 * (3.24)
Interaction: Health status (bad) x No other treatments (β_9)	N/A	-0.4289 ** (0.65)
Summary Statistics:		
Log-likelihood	1116.81	1097.10
Model Chi^2 (df)	576.40 (9)	571.45 (10)
Pseudo R^2	0.2051	0.2066
% correct predictions	73%	73%

* p=0.01, ** p=0.05

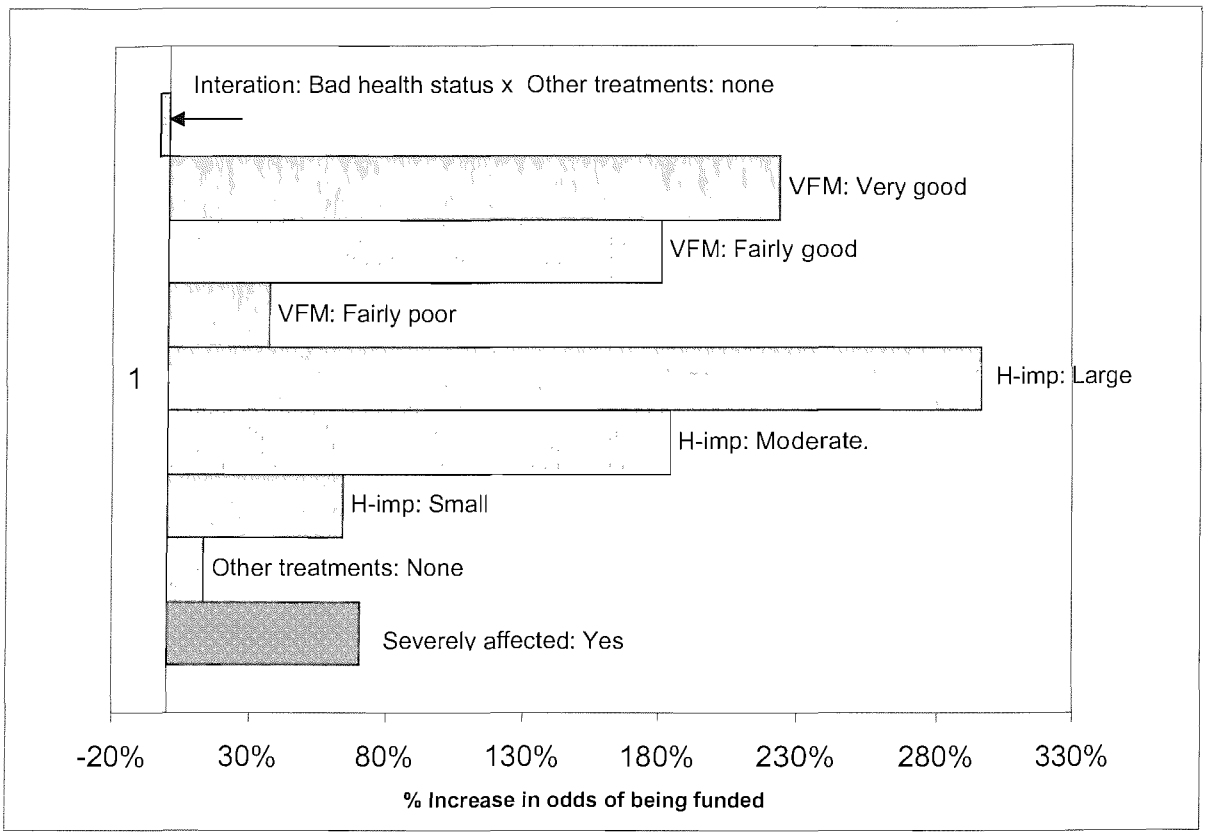
The most important attribute is the 'level of health improvement', followed by 'value for money', with 'severity of health' the next important, and 'other treatments' being the least important of the attributes. The most important single increment in utility is that from (fairly) poor value for money to (fairly) good value for money, with an increase in utility of 0.716. The interpretation of the importance of each attribute by level is important relative to the base case (constant). In terms of utility (logit function), severity (treating a severely affected patient group), whilst not as important overall as the attributes for health improvement or value for money, is seen to have an impact that is greater than some of the incremental impacts across differing levels of these attributes. For example the impact of treating a severely affected patient group is seen to be more important than a change (a) from 'very poor' value for money to 'fairly poor' value for money, (b) from a 'very small' health improvement to a 'small' health improvement, (c) from a 'moderate' health improvement to a 'large' health improvement, or (d) from 'fairly good' value for money to 'very good' value for money. It can also be interpreted as being of similar importance to a change from a 'small' to a 'moderate' health improvement. The impact on choice of the attribute covering the

availability of other treatments would appear to be small, although statistically significant (coefficient 0.1243). It is smaller than the other incremental changes in the model, although similar to the impact of moving from 'fairly good' value for money to 'very good' value for money.

The results (Table 19) indicate that as long as an intervention represents good value for money (fairly or very) and offers a health improvement that is at least moderate, it will have a relatively good utility value (be desirable), regardless of the other two attributes. Table 19 shows that scenarios ranked 8th, and 10th to 12th are for patient groups who are not severely affected by their health condition (and in two cases there are other treatment options available). However, these scenarios have either a large health improvement or represent very good value for money. The scenarios ranked 19th and 21st, still retaining a relatively high rank order, are for health interventions that are for non-severely affected patient groups (and with other treatments available in the latter scenario), with moderate health improvements and fairly good value for money enabling them to reflect a desirable funding scenario. In scenarios ranked 22nd to 27th the severity attribute is important in the utility function, with the severity attribute being influential in the rank order in the face of either small (very small) and/or poor or very poor value for money attribute levels. In 21 of the top ranked 32 scenarios the severity attribute reflects a severely affected patient group, with only 11 of the bottom ranked 32 scenarios reflecting a treatment for severely affected patients.

Interpreting the coefficients as odds ratios (Table 18) gives a clear view of the relative importance of incremental changes from the base scenario (worst case scenario), see Figure 7. The odds ratio is defined as the ratio of the odds of an event occurring in one group to the odds of it occurring in another group. In this instance it is the odds of success (of a health care scenario being chosen for funding). An odds ratio of 1 indicates that being chosen (for funding) is equally likely in both cases. An odds ratio greater than 1 indicates that the scenario is more likely to be chosen than the base case scenario. And an odds ratio less than 1 indicates that the scenario is less likely to be chosen than the base case scenario. For example, an odds ratio of 1.70 (severity) indicates an increase of 70% in the odds of that scenario being chosen relative to the base case (base case scenario, versus same scenario with one change from non-severe patient group to severely affected patient group, *ceteris parabis*). These incremental differences shown in the figure are similar to those in the utility (logit) function, but not the same.

Figure 7. Model results as the impact on the odds of 'success' (being funded) relative to the worst case scenario.



Key: VFM = value for money; H-Imp= level of health improvement

The base case scenario is judged *a priori* to be the worst case (least attractive description in the design used), and the constant is used (via the dummy variable approach) to capture the relative value of all other scenarios, compared to the worst case. A different analytical approach would be to use effect codes. Effects codes are a useful alternative to the use of dummy variables. In contrast to dummy variables, effects codes are uncorrelated with the constant (β_0) in the model (i.e. the constant is not used to explore the relative difference between each scenario and the base reference case). Effects codes are equivalent to calculating the marginal means of the levels of each attribute (Louviere *et al,* 2001). Effects coding allows the analyst to investigate the incremental differences in attributes across all levels, this does not use the constant to capture the base case or reference case impacts for each attribute. For example, the utility associated with a 'very small' health improvement, or a 'very poor' value for money. In order to consider this approach, analysis was undertaken using effects coding. This was undertaken to consider the component parts of the constant term used, and the relationship across the levels of the attributes (those attributes with greater than 2 levels). This analysis indicated that the constant comprised almost entirely of

the reference case attribute values, and that the relationship across attributes for health improvement and value for money was a non-linear one (see Appendix 6 for detail).

7.3.6 Using the DCE results to predict utility levels and the probability of success

The utility functions for each of the 64 scenarios (full factorial) have been estimated. Table 19 presents the utility estimates, from the logit function, showing utility (preference) scores range from a high of 3.44 for the best case scenario to a low of 0.23 for the worst case scenario, and provides the basis for considering the differences between each of the alternatives.

Using the conditional logit model (see section 6.3.4) it is possible to estimate a probability of success (i.e. being selected for funding) for each of the scenarios, relative to one another (i.e. probability of being selected from the total group of 64 options). The probabilities for the 64 scenarios sum to 1.00. Table 19 presents the 64 scenarios in a ranking according to their probability of being chosen/funded (highest probability ranked at the top). This table shows, for example, the top ranked item has a probability of 5.39% of being chosen from the set of 64 scenarios. More importantly it shows that one from any of the top 10 scenarios would be selected from the full factorial in around 40% of cases, and that in over 70% of cases 1 scenario from the top 25 scenarios would be chosen. A number of scenarios are shown to be very unlikely candidates for funding from the scenarios available (i.e. a less than 5% probability that one from bottom 10 scenarios would be chosen).

7.4 Discussion

This chapter has presented a policy application of the DCE framework. It adds to the sparse literature informing on the use of DCE methods to explore social values, and is (to my knowledge) the first interview-based study in a sample of the general public (a quasi-random sample) to consider the relative weights that respondents place on a range of competing social values; where cost-effectiveness (using the value for money attribute) is considered in the range of social values. The findings from the study provide important and useful insights to the relative importance of competing social values amongst the general public, and persuasive evidence on the feasibility and acceptability of the DCE approach in a sample of the general public.

The study demonstrates that it is feasible to present priority setting choices using the DCE framework to samples of the general public. Whilst there is no objective data on the

Table 19. Estimated utility score and probability of success, by scenario (full factorial), rank order from 'most preferred/desirable' to 'least preferred/desirable'.

Ranked by prob./utility	Attributes (scenarios): Value for Money	Health Improvement	Severity	Other Treatments	Utility (logit)	Probability	Cumm. Probability
1	very good	large	severe	No other Tx	3.44	5.39%	5.39%
2	very good	large	severe	other Tx avail	3.31	4.76%	10.16%
3	fairly good	large	severe	No other Tx	3.30	4.68%	14.83%
4	fairly good	large	severe	other Tx avail	3.17	4.13%	18.96%
5	very good	moderate	severe	No other Tx	3.11	3.87%	22.83%
6	very good	moderate	severe	other Tx avail	2.98	3.42%	26.25%
7	fairly good	moderate	severe	No other Tx	2.96	3.36%	29.61%
8	very good	large	not severe	No other Tx	2.91	3.17%	32.78%
9	fairly good	moderate	severe	other Tx avail	2.84	2.96%	35.75%
10	very good	large	not severe	other Tx avail	2.78	2.80%	38.55%
11	fairly good	large	not severe	No other Tx	2.76	2.75%	41.29%
12	fairly good	large	not severe	other Tx avail	2.64	2.43%	43.72%
13	fairly poor	large	severe	No other Tx	2.58	2.28%	46.00%
14	very good	moderate	not severe	No other Tx	2.58	2.28%	48.28%
15	very good	small	severe	No other Tx	2.56	2.24%	50.52%
16	fairly poor	large	severe	other Tx avail	2.46	2.02%	52.54%
17	very good	moderate	not severe	other Tx avail	2.45	2.01%	54.55%
18	very good	small	severe	other Tx avail	2.43	1.98%	56.52%
19	fairly good	moderate	not severe	No other Tx	2.43	1.97%	58.50%
20	fairly good	small	severe	No other Tx	2.42	1.94%	60.43%
21	fairly good	moderate	not severe	other Tx avail	2.31	1.74%	62.18%
22	fairly good	small	severe	other Tx avail	2.29	1.71%	63.89%
23	very poor	large	severe	No other Tx	2.26	1.67%	65.56%
24	fairly poor	moderate	severe	No other Tx	2.25	1.64%	67.20%
25	very poor	large	severe	other Tx avail	2.14	1.47%	68.67%
26	fairly poor	moderate	severe	other Tx avail	2.12	1.45%	70.12%
27	very good	very small	severe	No other Tx	2.06	1.36%	71.48%
28	fairly poor	large	not severe	No other Tx	2.05	1.34%	72.82%
29	very good	small	not severe	No other Tx	2.03	1.32%	74.14%
30	very good	very small	severe	other Tx avail	1.94	1.20%	75.34%
31	very poor	moderate	severe	No other Tx	1.93	1.20%	76.54%
32	fairly poor	large	not severe	other Tx avail	1.92	1.19%	77.72%
33	fairly good	very small	severe	No other Tx	1.92	1.18%	78.90%
34	very good	small	not severe	other Tx avail	1.90	1.16%	80.07%
35	fairly good	small	not severe	No other Tx	1.88	1.14%	81.21%
36	very poor	moderate	severe	other Tx avail	1.81	1.06%	82.26%
37	fairly good	very small	severe	other Tx avail	1.80	1.04%	83.31%
38	fairly good	small	not severe	other Tx avail	1.76	1.01%	84.31%
39	very poor	large	not severe	No other Tx	1.73	0.98%	85.29%
40	fairly poor	moderate	not severe	No other Tx	1.72	0.96%	86.26%
41	fairly poor	small	severe	No other Tx	1.70	0.95%	87.20%
42	very poor	large	not severe	other Tx avail	1.61	0.87%	88.07%
43	fairly poor	moderate	not severe	other Tx avail	1.59	0.85%	88.92%
44	fairly poor	small	severe	other Tx avail	1.58	0.84%	89.76%
45	very good	very small	not severe	No other Tx	1.53	0.80%	90.56%
46	very good	very small	not severe	other Tx avail	1.41	0.71%	91.27%
47	very poor	moderate	not severe	No other Tx	1.40	0.70%	91.97%
48	fairly good	very small	not severe	No other Tx	1.39	0.69%	92.66%
49	very poor	small	severe	No other Tx	1.38	0.69%	93.36%
50	very poor	moderate	not severe	other Tx avail	1.28	0.62%	93.98%
51	fairly good	very small	not severe	other Tx avail	1.26	0.61%	94.59%
52	very poor	small	severe	other Tx avail	1.26	0.61%	95.20%
53	fairly poor	very small	severe	No other Tx	1.20	0.58%	95.78%
54	fairly poor	small	not severe	No other Tx	1.17	0.56%	96.34%
55	fairly poor	very small	severe	other Tx avail	1.08	0.51%	96.84%
56	fairly poor	small	not severe	other Tx avail	1.04	0.49%	97.34%
57	very poor	very small	severe	No other Tx	0.89	0.42%	97.76%
58	very poor	small	not severe	No other Tx	0.85	0.41%	98.16%
59	very poor	very small	severe	other Tx avail	0.76	0.37%	98.54%
60	very poor	small	not severe	other Tx avail	0.73	0.36%	98.90%
61	fairly poor	very small	not severe	No other Tx	0.67	0.34%	99.23%
62	fairly poor	very small	not severe	other Tx avail	0.55	0.30%	99.53%
63	very poor	very small	not severe	No other Tx	0.36	0.25%	99.78%
64	very poor	very small	not severe	other Tx avail	0.23	0.22%	100.00%

* utility (logit) = conditional logit function; probability = conditional logit equation [see section 6.3]

numbers who refused to participate in the study, given the quota sampling approach and the door-to-door nature of the survey, the feedback from interviewers in the pilot study and the main survey was that respondents were keen to participate, seemed engaged in the survey, and generally had few problems completing the survey (accepting the fact that choices were difficult ones). There were only a small number of respondents who 'failed' the consistency question, and the majority of respondents reported little or no difficulty in understanding the questions posed (although circa. 30% did report it was fairly difficult to understand the questions). Whilst respondents indicated that the questions were difficult to answer, it would have been more of a concern if they had not indicated such difficulty.

The study provides data on the relative value of the different attributes, when considering changes across the different healthcare scenarios presented. The conditional logit model used demonstrates a good fit in terms of the pseudo R^2 statistic, and the predictive power of the model. The coefficients across all attributes and levels are highly significant indicating that, in general, the choices respondents made were sensitive to variation in the levels for the attributes. In addition the positive (+) sign on the coefficients for all main effects is as expected, and supports the theoretical validity of the model used.

The most important attribute/level changes involved level of health improvement and value for money. The findings show that the general public hold 'severity of health condition' and 'the availability of other treatments' as important social values, not to be ignored, and this provides some support to those arguing against an often dominant 'health maximising' decision maker perspective. However, the findings also indicate that in many instances the level of health improvement and value for money arguments provide a strong indication of the social value (preference) associated with a health care intervention. It is important to consider the qualitative levels of the attributes, and when considering the relative importance of the severity attribute it is notable that it has a relatively big (important) impact, compared to a number of the incremental level changes in the attributes for value for money and health improvement. This issue has been highlighted above.

It is important to consider how the data presented can be interpreted when looking at the relative desirability (attractiveness) of alternative health care funding scenarios. The logit scale, used in the utility functions, is linear in parameters (an additive model) and it provides values on an interval scale (interval scale properties) i.e. it is the difference between alternatives that is important (e.g. difference of 3.21 units between the top ranked and bottom ranked scenarios). The data provides a relative measure of the desirability of the differences between alternative scenarios. It is not theoretically possible to interpret the

utility (logit) values on a ratio scale, and it is therefore not possible to state how much one alternative may be regarded as better or worse than another (e.g. ten times better, or five times worse). For example, the data show that the scenario ranked 10th is preferred to the scenario ranked 30th, and that the difference between the two scenarios in utility terms is 0.84 (a difference similar to the difference between the 2nd ranked scenario and the 16th ranked scenario), but it is not possible to quantify how much better ('betterness') the 10th ranked scenario is compared to the 30th ranked scenario. It is possible to state that the utility gain from the 30th to the 10th ranked scenario (0.84 units) is approximately a quarter of the utility gain from the worst to the best ranked scenario (3.21 units).

The probability scale has ratio level properties, and when estimating the probability that a scenario will be chosen for funding, from the full factorial (64 scenarios), these probability values are arguably more suitable for use as an absolute measure of desirability across the alternatives (in the experimental design used). That is, when comparing estimates of the probabilities that each option will be chosen, from the full factorial, it is possible to state how much worse (or better) one probability is compared to another. For example, using the probability scale it is possible to state that the scenario ranked 10th in Table 19 is 2.33 times more likely to be chosen than the scenario ranked 30th in Table 19 (the preference for the 10th ranked item may be regarded as over twice as great than that for the 30th ranked item). Therefore, it is suggested here (discussed in Chapter 8) that the probability estimate can be used as a more appropriate measure of the "strength of preference" across scenarios, for the purpose of health care decision making.

As well as the above discussion on the interpretation of the data, a number of other issues arise for discussion from the study. These include the observed interaction term for health status in the DCE model, the observations from the comparison of ranking data and discrete choice data, the limitations in the current study design, and the way this study fits with the current literature on the elicitation of social values using the DCE framework.

The discrete choice modeling indicated that there was one statistically significant interaction term amongst the respondent characteristics identified *a priori* as having the potential for interactions effects. Table 18 reports the impact of self-reported health status on the choice model, when health status is interacted with the 'attribute covering the availability of other treatments'. The finding that persons with a self reported 'bad' health status (bad or very bad) placed less weight (priority) on health care treatment scenarios where there were 'no others treatment options available', compared to persons with a self-reported health status of good (or very good), was unexpected. It may have been more likely *a priori* to consider

those in a worse health status would put greater value on health interventions which were the only option available for patient groups (i.e. supporting those patients who had no alternative treatment available). To speculate as to why there is such an interaction, the finding may be a true reflection of respondent preferences, taking into account the various scenario attributes. An alternative explanation may be that it is a result of persons with a bad health status either misunderstanding the questions posed, or failing to consider all of the trade-offs present in the DCE questions, when choosing which of the two scenarios to support. There is some evidence from the sub-group analysis, of the data on difficulty with the understanding of DCE questions, that those with a bad health status were more likely to have difficulty understanding the DCE questions presented; 16 of the 24 respondents in this group indicated a degree of difficulty (13 stated fairly difficult, 3 stated very difficult). This group (bad health status) had a mean age of 60-years (SE 16.6 years).

Of some interest is the difference we see in the ranking of some attributes compared to their impact on the discrete choice model. There is a notable difference in the ranking of the 'value for money' attribute, by level of importance (ranked least important on a mean ranking score), and the impact of the attribute on the discrete choices made (the attribute has a big impact on choice). There is also a suggested difference in the ranked level of importance of the severity attribute (ranked most important) and the role it plays in the discrete choice model. Whilst the ranking task was primarily a warm-up task, and it is not suggested that differences in stated preferences can be analysed with any certainty, the results suggest that some thought should be given to the relative merits of the two different tasks i.e. ranking versus a series of pairwise choices. Other studies have also reported such differences. For example, Schwappach (2003) reports differences between ratings of importance and the actual impact of attributes on choices. Schwappach reports that the majority of respondents in his study stated that social class was of "no or little" importance, but this attribute contributed significantly to the allocation decisions in the discrete choice study reported. Such evidence suggests that where respondents are asked direct questions, such as 'is value for money important in priority setting decisions?' they may give the answer they feel is expected (i.e. "no" or "of course not"). Yet, when faced with choices, which involve a sacrifice (and may force some deliberation), they do indeed see value for money as important. It is clear that this may be the case with an issue like social class in some instances (political correctness). It may also be an explanation of the conflicting findings in some of the 'empirical ethics' literature around issues such as age and desert (as discussed in Chapter 4).

7.5 Limitations

Whilst the study design has a number of positive characteristics - i.e. it used a representative sample of the general public, professional interviewers, recommended statistical techniques to define the experimental design, a relatively small factorial design and a fractional factorial design that comprised 50% of the scenario descriptions in the full factorial (this will minimise any potential bias in the main effects from interaction terms), no response data are excluded - it is also important to consider the limitations of the study.

The first of the design issues considered is the potential limitation reflected by the small number of attributes, and subsequent levels, used in the experimental design. With four attributes (2 *x* 2-level, 2 *x* 4-level) the simple design employed in the study was regarded as the basis for a parsimonious model capable of addressing the key generic social values (attributes) present in a wide range of health care allocation decisions. The small factorial design used (n=64) was a deliberate choice. Firstly, because this study was one of the first interview-based studies of this type (first to include value for money in a sample of the general public), it was regarded as important to test the methods in a fairly simple context. Secondly, the simple experimental design was chosen in order to present an acceptable set of results to health policy decision-makers (discussed Chapter 8); to test the framework for priority setting choices, and to get some general messages on the relationship between key generic social values. However, there may be other attributes that are important and that have not been presented in the design used. A greater number of attributes and levels would create a larger experimental design and a need for a larger sample and/or more complex questionnaire format, but could lead to greater precision in the estimates from the choice data.

A further design limitation is the use of the fractional factorial main effects design used and the simple additive logistic regression model used, as this is likely to account for only around 70-90% of the explained variance (Louviere *et al,* 2000). The inclusion of two-way or higher order interactions between attribute levels could potentially account for further explained variance. Yet, the literature suggests that DCE models derived from main effects only designs generally predict well in the attributes of greatest interest (Louviere *et al,* 2000).

In terms of the presentation of the scenarios, in the questionnaire, the study used a simple generic description of attributes and levels, and the questionnaire was set up with limited description/explanation of the survey and the attributes and tasks. The interview format, and interviewer briefings, enabled any serious misunderstandings to be identified and tackled as part of the interview schedule, but there were no reports from the interviewers of responders

unable to understand or engage in the survey. Where four interviewees failed to provide a response to the DCE questions (see under response rate) these same people did provide some information for the ranking task, and a subsequent priority setting question (unrelated to the DCE survey). There was no indication that they did not understand the questions (two of these people stated they did not find the questions at all difficult to understand, whilst two did not respond to this question), they just did not wish to respond to the questions presented. The DCE questionnaire format employed here was informed by the successful survey methods of a previous DCE study, in a sample of the general public (Bryan *et al,* 2002). In the study by Bryan *et al* (2002) limited information was presented to the sample, and pilot work negated the need for clinical contextual descriptive information to inform the DCE questionnaire used. To inform the DCE study reported in this thesis, pre-pilot work and a formal pilot study provided a good level of support for the questionnaire format used in the main survey.

The use of generic attributes and generic (unlabelled) scenarios may be regarded as enabling the preferences to be considered in a general health care context, transferable across health policy areas, given that the preference data are less likely to be confounded by supporting information (e.g. age of patients, or specific health condition). However, it may be that a greater amount of information could be presented, and collected, in future studies of this nature. But such an approach may have to be at the expense of a smaller number of DCE questions being presented to each respondent (to limit responder burden and to limit the length of the interview). In such a way future studies could explore the impact on contextual factors and framing effects, which are limitations in all empirical studies eliciting preferences in the health care literature. For example, in the survey materials used here the health care scenarios were presented as being relevant to "long-term health conditions", and the study did not investigate any differences in response to different health conditions i.e. life-saving vs. life enhancing, and/or acute vs. chronic conditions.

The study has not considered test-retest reliability. The focus was on maximising the sample size, and the response data in the main survey rather than trading-off these factors to allow a smaller follow-up study to test reliability. Other studies using DCE in a health care context have reported positive findings against reliability (e.g. Bryan *et al* 1998, Schwappach & Strasmann 2006). In addition there are a number of factors which may suggest that the data collected are reliable and robust. The data are from a sample which is representative of the general public in Southampton, the sample is quite large, the study used face-to-face interviews rather than a postal questionnaire format, thereby allowing the opportunity to assess whether responders understood the nature of the DCE task they were presented

with. However, in future studies reliability could be considered to further support the DCE framework in a social choice context.

Of some concern in DCE studies is the fact that a high number of respondents are often excluded from data analysis, due to specific response patterns (i.e. non-traders, those with monotonic preferences), or due to respondents failing specific choices which were set out to test consistency (where one scenario from the choice presented was dominant against all attributes). In the present study it was important to include as much preference data as possible, and not to be seen to be excluding data on the basis of assumptions over potentially irrational choice data. This study is collecting preferences on social choices from a sample of the general public and no *a priori* assumptions were made on potential exclusion criteria. Whilst a DCE is based on the premise that responders trade-off attributes against one another, it is also important in DCE studies (especially those of a social value context) to take into account data from those with preferences that are not fully consistent with this approach (Lanscar & Louviere 2006). There has been no analysis of preference patterns in the current study, rather the strategy was to use all available data in the analysis. Where respondents were deemed to have failed the 'consistency question' in the current study (n=14) their response data was included in the analysis wherever possible. A question was placed in the interview schedule to qualitatively explore the reasons behind inconsistent responses to the consistency check question, however the findings from the survey were unhelpful in explaining these inconsistent responses. This may be due to the largely quantitative nature of the survey design, or the poor recording of verbatim responses by the interviewers. In the present study there were very few inconsistent responders (to this specific question), only 5% of the sample, therefore it is not expected to present a serious limitation with the data or the survey findings. In future studies it is important to consider this issue further, with the use of more comprehensive qualitative techniques (or possible tape recording of interviews).

In a survey of this nature there could be some concern over the level of engagement and reflection of responders, and it would have been preferable to have some qualitative insights to this aspect of the study. However, in the current study this was not possible. There is a lack of objective data on response rate (due to sampling method used), but feedback from interviewers (and in the pilot work) was positive indicating a high level of interest and engagement from responders. It may be that the observed differences between the stated ranking of the importance of attributes and their impact in the discrete choice model could indicate that responders did consider the attributes and weigh up the different scenarios presented. For example, finding that value for money may have been more important than

stated in the ranking task, when other attributes were allowed to vary across levels. In the DCE choices responders were faced with the outcome of their choice (i.e. one scenario was not chosen). The presentation of scenarios with varying levels may have forced some deliberation, rather than the case in the ranking task where no opportunity cost was present in the ranking and a more immediate preference (without deliberation) may have been provided. In future work, more deliberative and qualitative interviews could be used to assess the basis for differences in the two tasks, and more appropriate ranking tasks (not warm-up exercises) could be introduced to future surveys to consider the two response formats.

7.6 Conclusions

The current evidence base on social values is diverse and complex to interpret, and a number of commentators have highlighted the need for studies to consider the relationship between key social values. The study presented here has explored such a relationship, and is able to provide an important insight to the relationship between the social values used in the survey. There have been no other DCE studies using cost effectiveness in a sample of the general public, therefore it is difficult to consider the findings in the context of findings from other studies. The broad conclusions from the study support the existence of an equity versus efficiency trade-off, with equity characterised as a balance between the competing social values. The level of health improvement, and the value for money attributes are presented as prominent social values in the discrete choice modeling of the choice data from respondents, but the severity of health attribute is also shown to be an influential social value in finding a balance between social values.

This study differs from many others, as it is based on a robust sampling technique, it reports findings from a reasonably representative sample of the general public in Southampton, and uses interview-based techniques. Data indicate that respondents were willing to trade across attributes, and as such were willing to trade efficiency gains (health improvements, value for money) to support other attributes of a 'non-health maximising' nature (severity, other treatments). Whilst it is not possible to consider whether responders chose in line with a health maximising approach (given the nature of the attributes and levels employed), it is possible to see that health gains could be sacrificed to support non-health maximising objectives (Table 19). However, the trade-offs are of a qualitative nature, and trade-offs may be complex and may depend on the combination of attributes and levels in the different scenarios. But it would appear that the general public (sample used) would give priority to interventions that deliver large health improvements to patient groups (compared to small

improvements), and those that offer good value for money (compared to poor value for money).

This study provides a useful empirical insight, in a policy-relevant (real-world) context. This insight is explored in more detail in the following chapter.

8 MAKING USE OF THE DISCRETE CHOICE SURVEY DATA: AN APPLICATION TO UK NHS HEALTH POLICY

8.1 Introduction

This chapter demonstrates how data from the DCE presented in Chapter 7 may be used to assist in a health policy decision-making environment.

Firstly, and generally, findings from the empirical work offer a broad insight to the views of the general public. As discussed in Chapter 7, the findings suggest that the public are able to consider competing social values, and are able to trade-off these values against one another. The results from the survey presented do not support the use of health maximisation (obtaining the greatest health gain from available resources) as the dominant decision-making criterion. The findings from the survey do indicate that health maximisation is a strong influence on social preferences, but it is clear that this is not equivalent to the dominant use of cost-effectiveness analysis. That is, value for money is suggested as a strong social value, but the level of health improvement (regardless of cost-effectiveness) is also a strong social value. Level of health improvement here has been considered independent of efficiency concerns. These findings, together with a strong preference to treat the more severely affected patient groups, and a less strong indication that the public prefer to fund treatments that provide health gain to those who would otherwise not have an effective treatment available to them, run counter to the prominent use of efficiency arguments to determine health care funding decisions. They also run counter to the assumption made by NICE that a QALY is of equal value to all persons (NICE, 2005). These general insights contribute to the literature in this area, discussed earlier in Chapter 4.

However, the findings from the DCE can be used more directly to provide guidance on the social preferences that may be assigned to specific categories of health technologies. In this chapter such an application of the DCE data is suggested, and demonstrated.

Given the context of the study undertaken, this specific discussion on the application of DCE data is focused specifically around the NICE health technology appraisal programme. The NICE health technology appraisal process has been introduced in earlier chapters (Chapters 2, 3 & 4), and it is used here as a vehicle for demonstrating the application of findings from the empirical work undertaken.

8.2 Relative utility value for health technologies

The DCE data analysis presents estimates of the relative desirability of the alternative health care scenarios presented (via the experimental design used). This relative preference measure is based on the four attributes used. There may be other factors that are important in the appraisal of health technologies, but the appraisal of a health technology can take into account information from various other sources (as well as including this measure of public preference). The DCE survey, presented in Chapter 7, was designed, using the attributes and levels previously described, to provide a broadly specified matrix of scenarios to capture the full range of possible health technology appraisal scenarios at a general level, using the key social values represented by the attributes. The expectation being that a decision-maker, at a policy level, should be able to make a well informed judgment (against the four attributes) on the health technology under review. In such a way it should be possible to assign the health technology to one of the 64 scenarios in the DCE design employed. Where this is possible the decision-maker is able to draw some general view on the preferences of the general public (as reflected through the sample used in the survey).

The NICE appraisal process is comprehensive and is able to facilitate such judgments against the DCE attributes and levels. In the NICE appraisal process a number of documents are presented to inform the appraisal (e.g. manufacturers submission on clinical and cost-effectiveness, submissions from clinical and patient groups, an independent assessment report on the clinical and cost-effectiveness of the technology), and the NICE secretariat, together with the NICE Appraisal Committee will be able to draw conclusions from the documents, and discussions, on the nature of the technology with respect to the attributes used in the DCE. For example, they will be able to make a context specific judgment as to whether the condition severely affects the patient group, and whether other treatment options are available. The decision-maker (NICE Appraisal Committee) will also be able to arrive at a judgment on the level of health improvement (e.g. large, moderate or small), and the placement of the health technology with respect to the cost-effectiveness estimates available (e.g. very good value for money, or very poor value for money). These judgments will be influenced by the specific decision making context and the guidance used by the organisation undertaking the appraisal of the health technology. Using the data presented in the DCE on level of utility for health care interventions, comparisons can be made between the different health care interventions, as described using DCE attributes. This will be on the basis of the level of utility expected from a specific health technology relative to a base case, i.e. the worst placed option in the experimental design.

Whilst not the focus for current illustrative discussions on the use of DCE findings, it is also anticipated that in other decision-making forums (e.g. hospital drug formulary groups, regional and local NHS organisations), deliberation and judgment of health technologies against the four attributes and levels (in the DCE) would be a reasonable expectation. With such judgments being made without being too onerous or time-consuming (over and above currently held discussions over health policy decisions).

8.3 Deriving a measure of 'strength of preference'

Chapter 7 has discussed the DCE data on utility (logit function), and the estimated probabilities presented to show the chance of each health scenario being chosen from the choice set (n=64). The probability data is introduced as a basis for presenting a measure of the 'strength of preference' across the health care scenarios presented in the experimental design of the DCE. It is argued here that the experimental design should be able to capture the general nature of all health technologies under consideration, given the limits on the attributes used.

It is suggested here that the probability data (Table 20) can be interpreted quite broadly, as a measure of 'strength of preference', to place the health care scenarios into a limited range of useful (policy-relevant) preference categories. To my knowledge there are no such measures of 'strength of [public] preference' in the current literature. In this thesis a very simple and general matrix of 'strength of [public] preference' is proposed (Table 20), applying the data from the DCE presented in Chapter 7.

Table 20 presents a simple categorisation of health technologies (by descriptive scenario). To characterise strength of preference four categories are used, from 'very strong' public (social) preference to 'very weak' public preference. This qualitative and ordinal range of preference categories is derived and presented here based on a wish to avoid a specific numeric measure of preference (or giving the impression of being too scientific). At the outset, the DCE study undertaken has aimed to provide empirical insights to a 'general relationship' between key social values, using the relative value of attributes. The thesis does not present the DCE as a means to capture the specific or definitive magnitude of differences between the scenarios described. However, to illustrate its potential value in health policy these suggested categories of preference are proposed here.

Using the estimates of probability in Table 20 it is clear that there are a number of strongly favoured scenarios, and a number of unlikely scenarios (unattractive options). The scenarios ranked 1st to 15th would be chosen most commonly from those presented in the

full choice set. One scenario from these 15 would be selected in over 50% of cases, therefore these scenarios are categorised here as 'very strongly preferred'.

A number of the scenarios are clearly unattractive (those ranked lowest). In those scenarios ranked from 46th to 64th (the bottom ranked scenario) the cumulative probability that one of these 18 scenarios is selected from the choice set is less than 10%. Therefore these scenarios are described as 'very weakly preferred'. In the scenarios ranked 16th to 25th, neither very strongly preferred or very weakly preferred, a judgment on preference is made by judging each against what may be regarded as a mean probability, or preference weight. A mean probability across all 64 scenarios would be a 1.55% chance of being chosen (1÷ 64), a mean preference weight of 15.5. Those scenarios with a preference weight greater than the average weight, but outside of the very strongly preferred category, are judged to be 'fairly strongly' preferred for funding (9 scenarios). Those scenarios below an average probability and outside of the very weak preference category are judged to be 'fairly weakly' preferred (21 scenarios).

This characterisation of public and social preference is presented for illustrative purposes, and to allow some further investigation of the application of such a preference measure, as below. However, it is accepted, in the context of the current thesis that this may be viewed as a simple and crude categorisation. Yet, it is seen as an advance on the currently available empirical evidence to inform on social preferences of the general public with respect to health care priority setting decisions. Further qualitative work, outside of the scope of the current thesis, would be useful in informing the process of deriving a measure of 'strength of preference', possibly feeding back the results of such an exercise to the sample used for elicitation of preferences (DCE sample).

8.4 Applying the preference data to health policy decisions

To illustrate how the data, and the preference measure, may be transferred to a decision making perspective, the experimental design (attributes/levels) and resulting findings are mapped onto a selection of health technologies that have been appraised by NICE.

Ten of the health technologies that have been appraised by NICE have been selected, for use in the demonstration of the use of the DCE data. Topics have been selected, in a purposive way, on the basis of the candidates familiarity with the appraisal topic[18], and/or

[18] Through the candidates involvement in the appraisal of the health technologies through some level of involvement with the independent assessment reports submitted to NICE as part of the technology appraisal process (e.g. insulin pumps for diabetes, drotrecogin alfa for severe sepsis, bortezomib).

Table 20. Derived Preference Weights and Categories for Health Care Interventions (as described by attribute scenarios)

Ranked by prob.	Attributes (scenarios): Value for Money	Health Improvement	Severity	Other Treatments	Probability	Cumm Probability	Pref Weight*	Public Preference (Illustrative judgement)
1	very good	large	severe	No other Tx	5.39%	5.39%	53.94	Very Strong
2	very good	large	severe	other Tx avail	4.76%	10.16%	47.63	Very Strong
3	fairly good	large	severe	No other Tx	4.68%	14.83%	46.75	Very Strong
4	fairly good	large	severe	other Tx avail	4.13%	18.96%	41.29	Very Strong
5	very good	moderate	severe	No other Tx	3.87%	22.83%	38.73	Very Strong
6	very good	moderate	severe	other Tx avail	3.42%	26.25%	34.20	Very Strong
7	fairly good	moderate	severe	No other Tx	3.36%	29.61%	33.57	Very Strong
8	very good	large	not severe	No other Tx	3.17%	32.78%	31.70	Very Strong
9	fairly good	moderate	severe	other Tx avail	2.96%	35.75%	29.64	Very Strong
10	very good	large	not severe	other Tx avail	2.80%	38.55%	28.00	Very Strong
11	fairly good	large	not severe	No other Tx	2.75%	41.29%	27.48	Very Strong
12	fairly good	large	not severe	other Tx avail	2.43%	43.72%	24.27	Very Strong
13	fairly poor	large	severe	No other Tx	2.28%	46.00%	22.84	Very Strong
14	very good	moderate	not severe	No other Tx	2.28%	48.28%	22.76	Very Strong
15	very good	small	severe	No other Tx	2.24%	50.52%	22.38	Very Strong
16	fairly poor	large	severe	other Tx avail	2.02%	52.54%	20.17	Fairly Strong
17	very good	moderate	not severe	other Tx avail	2.01%	54.55%	20.10	Fairly Strong
18	very good	small	severe	other Tx avail	1.98%	56.52%	19.76	Fairly Strong
19	fairly good	moderate	not severe	No other Tx	1.97%	58.50%	19.73	Fairly Strong
20	fairly good	small	severe	No other Tx	1.94%	60.43%	19.40	Fairly Strong
21	fairly good	moderate	not severe	other Tx avail	1.74%	62.18%	17.43	Fairly Strong
22	fairly good	small	severe	other Tx avail	1.71%	63.89%	17.13	Fairly Strong
23	very poor	large	severe	No other Tx	1.67%	65.56%	16.67	Fairly Strong
24	fairly poor	moderate	severe	No other Tx	1.64%	67.20%	16.40	Fairly Strong
25	very poor	large	severe	other Tx avail	1.47%	68.67%	14.72	Fairly Weak
26	fairly poor	moderate	severe	other Tx avail	1.45%	70.12%	14.48	Fairly Weak
27	very good	very small	severe	No other Tx	1.36%	71.48%	13.63	Fairly Weak
28	fairly poor	large	not severe	No other Tx	1.34%	72.82%	13.42	Fairly Weak
29	very good	small	not severe	No other Tx	1.32%	74.14%	13.15	Fairly Weak
30	very good	very small	severe	other Tx avail	1.20%	75.34%	12.04	Fairly Weak
31	very poor	moderate	severe	No other Tx	1.20%	76.54%	11.97	Fairly Weak
32	fairly poor	large	not severe	other Tx avail	1.19%	77.72%	11.85	Fairly Weak
33	fairly good	very small	severe	No other Tx	1.18%	78.90%	11.81	Fairly Weak
34	very good	small	not severe	other Tx avail	1.16%	80.07%	11.62	Fairly Weak
35	fairly good	small	not severe	No other Tx	1.14%	81.21%	11.40	Fairly Weak
36	very poor	moderate	severe	other Tx avail	1.06%	82.26%	10.57	Fairly Weak
37	fairly good	very small	severe	other Tx avail	1.04%	83.31%	10.43	Fairly Weak
38	fairly good	small	not severe	other Tx avail	1.01%	84.31%	10.07	Fairly Weak
39	very poor	large	not severe	No other Tx	0.98%	85.29%	9.80	Fairly Weak
40	fairly poor	moderate	not severe	No other Tx	0.96%	86.26%	9.64	Fairly Weak
41	fairly poor	small	severe	No other Tx	0.95%	87.20%	9.48	Fairly Weak
42	very poor	large	not severe	other Tx avail	0.87%	88.07%	8.65	Fairly Weak
43	fairly poor	moderate	not severe	other Tx avail	0.85%	88.92%	8.51	Fairly Weak
44	fairly poor	small	severe	other Tx avail	0.84%	89.76%	8.37	Fairly Weak
45	very good	very small	not severe	No other Tx	0.80%	90.56%	8.01	Fairly Weak
46	very good	very small	not severe	other Tx avail	0.71%	91.27%	7.08	Very Weak
47	very poor	moderate	not severe	No other Tx	0.70%	91.97%	7.03	Very Weak
48	fairly good	very small	not severe	No other Tx	0.69%	92.66%	6.94	Very Weak
49	very poor	small	severe	No other Tx	0.69%	93.36%	6.92	Very Weak
50	very poor	moderate	not severe	other Tx avail	0.62%	93.98%	6.21	Very Weak
51	fairly good	very small	not severe	other Tx avail	0.61%	94.59%	6.13	Very Weak
52	very poor	small	severe	other Tx avail	0.61%	95.20%	6.11	Very Weak
53	fairly poor	very small	severe	No other Tx	0.58%	95.78%	5.77	Very Weak
54	fairly poor	small	not severe	No other Tx	0.56%	96.34%	5.57	Very Weak
55	fairly poor	very small	severe	other Tx avail	0.51%	96.84%	5.10	Very Weak
56	fairly poor	small	not severe	other Tx avail	0.49%	97.34%	4.92	Very Weak
57	very poor	very small	severe	No other Tx	0.42%	97.76%	4.21	Very Weak
58	very poor	small	not severe	No other Tx	0.41%	98.16%	4.06	Very Weak
59	very poor	very small	severe	other Tx avail	0.37%	98.54%	3.72	Very Weak
60	very poor	small	not severe	other Tx avail	0.36%	98.90%	3.59	Very Weak
61	fairly poor	very small	not severe	No other Tx	0.34%	99.23%	3.39	Very Weak
62	fairly poor	very small	not severe	other Tx avail	0.30%	99.53%	3.00	Very Weak
63	very poor	very small	not severe	No other Tx	0.25%	99.78%	2.48	Very Weak
64	very poor	very small	not severe	other Tx avail	0.22%	100.00%	2.19	Very Weak

*preference weight = probability multiplied by 1,000

due to the topics being controversial ('hot') topics and subject to general discussion and debate in the health care media (e.g. treatment for multiple sclerosis, drugs for Alzheimer's disease). The topics have been selected to provide a basis for discussion of the NICE appraisal process, and the application of the derived social preference data. Table 21 below outlines the topics selected for discussion; providing a judgment for the health technologies against the DCE attributes and levels. Appendix 7 presents a brief description of each of the selected health technologies, the relevant disease area, and a rationale for the judgments made to place the health technologies in the attribute matrix. For the purposes of this exercise, the following guide was used to assess the level of health improvement offered by the health technologies, and the judgment on the cost-effectiveness of the interventions:

Health Improvement: mean health improvement (QALY gain) per patient (as used in the estimated incremental cost-effectiveness ratio [ICER])

Large	> 1.00 QALY
Moderate	> 0.10 ≤ 1.00 QALY
Small	> 0.05 ≤ 0.10 QALY
Very Small	< 0.05 QALY

Cost-effectiveness (value for money): cost per incremental QALY

Very good	≤ £20,000 per QALY
Fairly good	> £20,000 ≤ £30,000 per QALY
Fairly poor	> £30,000 ≤ £40,000 per QALY
Very Poor	> £40,000 per QALY

It is accepted that this is a judgment made solely for the purposes of this thesis (by the candidate), to assist with the illustrative analysis, and may be regarded as a limitation of the analysis. In practice, the decision making body (organisation) will set out the basis upon which to assess health technologies against the categories/attributes used. NICE (2004) have provided some guidance for the consideration of cost-effectiveness, but do not offer any insight to the judgment needed against the level of health improvement. NICE state that where the ICER is below £20,000 per QALY judgments about the acceptability of the technology as an effective use of NHS resources are based primarily on the cost-effectiveness estimate, and above £20,000 per QALY technologies will need to refer to other factors (e.g. innovative nature of technology, particular features of the condition and population receiving the technology, wider societal costs and benefits) to inform the judgment over the acceptability of the technology as an effective use of NHS resources. NICE states that where a technology has an ICER above £30,000 per QALY the case for

supporting the technology, on these additional factors, has to be increasingly strong (NICE, 2004). NICE (2005) have stated that "the consensus amongst the Institute's economic advisors is that the Institute should, generally, accept as cost effective those interventions with an ICER of less than £20,000 per QALY and that there should be increasingly strong reasons for accepting as cost effective interventions with an ICER of over £30,000 per QALY" (p22-23). There is no such guidance on the level of health gain that may be regarded as large, moderate, small or very small, therefore the above judgments have been made for the purposes of the current analysis.

When mapping the health technologies to the attribute descriptions, many of the judgments are very straight forward, however, on a number of occasions it was necessary to consult with experts (public health consultants, lead researchers from the assessment teams involved in NICE appraisals) to inform judgments. The most challenging attribute, for assessment of the levels against the NICE reported guidance (and related assessment report), is the level of health improvement. This issue is not covered in any specific detail in many of the NICE guidance reports, and in a number of instances it has been difficult to determine an estimate of the level of health improvement considered most appropriate by NICE (as part of the appraisal process).

Table 21 presents a mapping of 10 appraisal topics, across a total of 13 appraisal recommendations. Appendix 8 presents a table covering current published NICE health technology appraisal guidance (up to June 2006). It draws from Raftery (2006) to categorise the guidance issued by NICE (either yes or no, and if yes whether with major, minor, or no restrictions); applying a similar approach to categorise guidance presented in the table but not included in the analysis by Raftery. In 10 of these 13 recommendations the recommendation made by NICE (yes or no) is consistent with the derived measure of public preference from the DCE data. In two cases the NICE recommendation is not consistent with the public preference shown via the DCE data. In one case, the use of insulin pumps in Type 1 diabetes, the public preference is shown to be very weak across the attributes, but the NICE guidance recommends use of insulin pumps (with some restrictions in use). In the case of riluzole for motor neurone disease, the preference data suggests a weak preference, and NICE recommend the drug for use in the NHS. In this latter case the NICE guidance does stress that the nature of the condition was a consideration in the appraisal judgment. In one instance, drugs for moderately to severe Alzheimer's disease, the mapping of the technology across the value for money attribute is unclear, given the discussion of cost effectiveness estimates presented in NICE guidance (£23,000 to £35,000 per QALY, or worse), and the public preference is sensitive to this input.

8.5 Reflections on the NICE health technology appraisal process

In the NICE health technology appraisal process, NICE are legally obliged to take account of both clinical and cost-effectiveness when developing guidance. The general directions to NICE from The Secretary of State for Health, require that they have regard for the broad balance of benefits and costs, the degree of clinical need of patients with the condition or disease under consideration, guidance from the Secretary of State, and the potential for long-term benefits to the NHS of innovation. It is suggested here that the NICE decision making process does not make explicit all of the considerations in arriving at the appraisal decision (e.g. Dakin *et al,* 2006). Guidance is issued which provides a summary on the main considerations over clinical and cost-effectiveness. Although NICE have developed their process over time and have introduced a section in their public guidance for discussion of 'considerations' of the evidence, they rarely discuss how various value judgments have been taken into account. Devlin & Parkin (2004), note that in their analysis of the first 42 NICE appraisal topics, in only four instances do factors that may be related to 'equity' (other social values, distributive arguments) get mentioned.

It is clear that cost-effectiveness plays a key role in the NICE appraisal process (NICE, 2004 & 2005, Dakin *et al* 2006, Devlin & Parkin 2004, Bryan *et al* 2006). The use of a strong efficiency objective (health maximisation) has been clear in the NICE appraisal process (NICE, 2004 & 2005). As it has developed its processes, NICE (2004, 2005) have become more explicit over the importance of cost-effectiveness, not stating explicitly that a cost-effectiveness threshold is used, but offering some guide as to what might be regarded as cost-effective, and what level of cost-effectiveness technologies will need to be supported by 'other factors'. NICE have come under increasing public scrutiny over their decisions not to recommend technologies on the grounds that they are not cost-effective, and the role of cost-effectiveness in the appraisal process has been the subject of general debate and analysis. NICE have been questioned over their treatment of factors such as age, gender, and social status, and have recently presented a statement on the social value judgments in the NICE appraisal process (NICE, 2005). However, in most cases social values are linked to consideration of the cost-effectiveness of the technology. The treatment of other decision making factors, such as severity of health, level of health improvement, have not come under such specific public scrutiny, but can be related back to cost-effectiveness in many instances through the calculation of health benefits, and the cost per QALY data often used in the technology appraisal process.

This thesis seeks to inform the debate in this area through discussion of, and empirical enquiry around factors that influence the type of social decision characterised by NICE

TABLE 21: Case-studies of health technologies appraised by NICE and the application of social preference data

Health Technology /	NICE Recommendation	Patients severely affected	Other treatments available	Health improvement *	Value for money **	Judgment on public preference[#]
Drugs for Alzheimer's disease *(TA111))*						
- Mild AD	No	Yes	No	Small	Very poor	Very Weak
- Moderate to Severe	Yes	Yes	No	Small	Fairly good/ fairly poor	Fairly Strong/Weak
- Mod-sev to Severe AD (memantine)	No	Yes	No	Small	Very poor	Very Weak
Insulin Pumps for Type 1 diabetes *(TA057)*	Yes	No	Yes	Small	Fairly poor	Very Weak
Drotrecogin alfa (activated) for severe sepsis *(TA084)*	Yes	Yes	No	Moderate	Very Good	Very Strong
Riluzole for Motor Neurone Disease *(TA020)*	Yes	Yes	No	Small	Fairly/Very Poor	Fairly/Very Weak
Bortezomib for treatment of multiple myeloma *(in progress)*	No	Yes	Yes	Moderate	Very Poor	Fairly Weak
Trastuzumab (Herceptin) for early breast cancer *(TA107)*	Yes	Yes	Yes	Large	Very Good	Very Strong
Beta interferon and glatiramer for treatment of multiple sclerosis *(TA032)*	No	Yes	No	Small/Moderate	Very Poor	Very/Fairly Weak
Anakinra for rheumatoid arthritis *(TA072)*	No	Yes	Yes	Small	Very Poor	Very Weak
Photodynamic therapy for macular degeneration *(TA068)*						
- classic ARMD with no occult CNV	No	Yes	No	Small	Very poor	Very Weak
- pred. classic ARMD with subfoveal CNV	Yes	Yes	No	Small	Fairly good	Fairly Strong
Imatinib for treatment of chronic myeloid leukaemia *(TA070)*	Yes	Yes	No	Large	Fairly good	Very Strong

* Judgments on the level of health improvement are based around the mean QALY gains reported in the NICE guidance and the assessment reports commissioned to support the NICE appraisal process. A topic-by-topic judgment is needed in this area. Appendix 7 provides summary detail on the rationale for the judgment. Most cases are reasonably clear, but where some expert judgment was required the lead researcher for the assessment group report was contacted for advice.

** Judgments have been made here on the basis of: Very Good Value for Money (VFM) – where cost per QALY is ≤£20,000; Fairly Good VFM – where cost per QALY is >£20,000 ≤£30,000; Fairly Poor VFM – where cost per QALY is >£30,000 ≤£40,000; Very Poor VFM – where cost per QALY is >£40,000.

[#] See discussion above for the basis of this judgment (made by the candidate, for illustrative purposes).

health technology appraisal guidance. It is suggested here that the empirical insights from the DCE presented in this thesis, and future research building on the findings presented here, can inform the public debate over the relative importance of cost-effectiveness in health policy decisions of this type.

A number of commentators have stressed the important role played by cost-effectiveness data (specifically cost per QALY estimates) in the NICE appraisal process. Devlin and Parkin (2004) have suggested the importance of cost-effectiveness data, and the use of some form of cost-effectiveness threshold in the NICE appraisal process; indicating some form of probability-based threshold in the region of £30,000 to £45,000 per QALY. In their analysis, Devlin & Parkin suggest that in addition to cost-effectiveness, uncertainty around the cost per QALY estimate (ICER), and burden of disease appeared to contribute to NICE decisions. However, their analysis had a number of limitations (e.g. covered a small number of appraisals, covered early NICE appraisals, found a lack of cost per QALY data) and it is presented as illustrative, suggesting further research is important.

Dakin *et al* (2006) present a model of the NICE decision-making process, characterising it as a simple, single-step decision between the three outcomes of 'routine use' (a yes decision), 'restricted use' (a yes, but... decision), and 'not recommended'. Dakin and colleagues provide some evidence on the importance of cost-effectiveness analysis in the NICE appraisal process. They present logistic regression analysis around NICE decisions, and suggest that several of the factors considered by NICE play an important role in NICE decision-making. These were primarily the nature and extent of the clinical evidence (RCTs, systematic reviews, meta-analyses) and the cost-effectiveness ratio of the intervention. Dakin *et al* also discuss the importance of intervention type (pharmaceutical or other), and budget impact. But suggest that intervention type may be correlated to the evidence on clinical effectiveness (RCT data, and standards), and highlight that budget impact did not have a significant impact in any of the regression analyses, only suggesting that budgetary constraints may have been taken into account alongside the evidence on clinical and cost-effectiveness (despite explicit statements to the contrary). The analyses also suggested that the presence of a patient group submission may also have been an important factor in the NICE appraisal decision, although the authors noted that limited data were available and the findings on this factor may have arisen by chance.

Bryan *et al* (2006) have explored the role of cost-effectiveness analysis in the NICE appraisal process. They use a qualitative approach, and their findings reinforce the view that economic analysis is of central importance in the appraisal process. They present case-

study analysis to demonstrate how new economic evaluations are commissioned for all appraisals, how in many cases new economic models are developed and presented as part of the process, and how the evidence is used and interpreted within the appraisal process.

Raftery (2006) presents some discussion on the importance of the cost-effectiveness data in the NICE appraisal process; questioning what may be an 'acceptable' cost effectiveness. He highlights a number of instances where NICE have considered cost per QALY estimates in excess of £30,000. Raftery also highlights instances where the cost effectiveness estimates from independent assessment teams have been subject to additional analysis, typically resulting in a much reduced estimate of the cost per QALY.

Some of the appraisal topics that have been contentious (e.g. drugs for treatment of people with multiple sclerosis, and those with Alzheimer's disease) and have taken a long time to complete, have involved a series of reports and additional analyses. For example, Raftery (2006) refers to the appraisal of drug treatments for multiple sclerosis, commenting that it took around 2-years to complete, with 338 documents listed on the NICE website for the appraisal. It is not clear from NICE guidance, and the supporting documentation, what factors have been taken into account when decisions are made surrounding the commissioning and use of additional cost-effectiveness analysis, where such analyses often takes a more conservative view (e.g. different cost and benefit assumptions, use of additional clinical data, use of retrospectively defined sub-groups) and results in lower cost effectiveness estimates than those initially presented by the independent analyses commissioned by the Department of Health to assist NICE in the appraisal process. This scenario can be observed in many of the 'hot' topics covered by NICE (e.g. drugs for multiple sclerosis, drugs for Alzheimer's, imatinab for CML, anakinra for rheumatoid arthritis, photodynamic therapy), many of which are included in the illustrative analysis above (Table 21), and summarised in Appendix 7. It may be that where NICE has a strong desire to recommend the use of a technology, even though it is not regarded as cost-effective at the normal level (i.e. below circa. £30,000 per QALY) there is a reluctance to make a statement that a technology at a high cost per QALY is regarded as a good use of NHS resources in that particular instance. It is clear in a number of the guidance documents that positive recommendations ('yes', or 'yes but') have been made following a reassessment of cost-effectiveness which has resulted in a cost per QALY estimate below, or around the £30,000 per QALY level. The suggestion in this thesis is that rather than the use of a specific re-assessment of cost-effectiveness estimates, using potentially overly optimistic scenarios, to support a positive recommendation for use of technology, such decisions could be supported on other grounds, regardless of cost-effectiveness. The data presented in the DCE study

undertaken as part of this thesis could offer some preliminary support (or substantial support where further research supports such findings) to decision makers wishing to provide health technologies that offer a moderate, or large, health improvement (as defined by the appraisal process) even where they are regarded as poor value for money. Such support can be mediated on the basis of the other value judgments considered in the DCE design (i.e. severity, other treatments).

From another perspective, where NICE feels it appropriate not to recommend a technology for routine use, or to recommend it with restricted use, on the grounds that it is not cost-effective, and/or that it offers only a very small health improvement, the data presented in the DCE may provide some form of social mandate or support for such decisions. The level of support will depend on other value judgments concerning the role that public preference data should play in health care priority setting decisions, however that is beyond the scope of the current thesis.

It is clear that the NICE appraisal process involves complex deliberations, covering a number of social value judgments, and that these may not be clearly documented. However, it is suggested that the type of public preference data presented in this thesis may offer support to such a process in a number of different ways, both general and specific.

9 DISCUSSION: EXPLORING THE SOCIAL VALUE OF HEALTH CARE INTERVENTIONS

This chapter draws together the findings from the thesis. It discusses the contributions made by the research reported in the thesis, and also discusses the limitations of the research reported. The chapter also presents suggested research recommendations.

9.1 Introduction

The thesis started from a position where the social value of health care interventions was expected to be a function not only of the health benefits (e.g. health gains) from health care, but also of the distribution of health benefits across different individuals in society. It has explored the social value of health care interventions within the analytical framework of health economics, and in the context of health technology appraisal within the UK NHS.

The aim of the thesis, outlined in Chapter 1, was to explore the social value of health care interventions. It has addressed two prime research questions:

1. What are the social values that can be used to set health care priorities over the funding of health care interventions in the UK NHS?

2. Where social values are identified, for use in decisions over the funding of health interventions, what are their relative values?

In addressing these primary research questions, the research has explored the social value of health care interventions. Research has placed the social evaluation of health care technologies in the broader literature of moral philosophy, examining theories of justice and fairness and how such theories are able to inform on the social value of health care. The complex and varied literature reporting empirical research on social values, and distributive preferences, i.e. preferences to rival a strong efficiency argument in the UK NHS, has been examined to identify key generic social values, that could be subject to further detailed scrutiny.

Empirical research has examined social values, in isolation (severity of health hypothesis) and in a multi-attribute format using a stated preference discrete choice experiment with four social values as attributes. This empirical research has been placed in a policy context, using the NICE health technology appraisal process, and has been set out as empirical

research that can guide health policy decision makers, although at the present time it is sill early explorative research. Future research recommendations are suggested.

9.2 Key findings and contributions

9.2.1 Theories of justice and fairness

The background survey on theories of justice (Chapter 3) provides a firm foundation on which to explore the social value of health care interventions. It sets out the social evaluation of alternative actions, as a multi-layered process, involving guiding principles, process, and some basis for evaluation. It provides a clear rationale for needing to examine and specify the basis upon which actions are evaluated (e.g. social values).

Whilst inferences have been made over the process of health technology appraisal and its underpinnings, in editorial and commentary style papers, there has been no published exposition (none identified in the current literature search) of the health technology appraisal process against general theories of justice. This thesis has considered the health technology appraisal process against a detailed description of each of the theories of justice covered in the thesis. The thought experiment used to consider health technology appraisal (e.g. NICE process) against the general theories of justice indicates that the process of health technology appraisal is not guided by any one approach to justice and fairness. Although there are signs that various elements of a number of theories of justice may be applied in a pluralist approach (Culyer 2001, NICE 2005), the NICE appraisal process (described in published reports) is not consistent with any of the theories described in this thesis.

Whilst not aligned to any specific theory of justice, it does appear that the NICE appraisal process has a firm procedural basis. However, the basis for evaluation, the 'evaluative space' (Sen, 1992) is not explicit across a range of distributive preferences, other than putting a heavy emphasis on the efficient use of public funds.

All theories of justice examined as part of the thesis state that the criteria for evaluating alternative actions should be clearly stated. The theories vary in their approach to the 'focal variable' (Sen, 1992) for evaluation, with some placing all issues in one general measure of social welfare (e.g. utilitarians and 'utility'), and others advocating principles which are able to operate when there are numerous possible variables (objectives, values) (e.g. Rawlsian justice, extra-welfarism, Nozian theory). Considering the broader assessment of social

welfare, through the theories of justice literature, provides a rationale for exploring the social values that may be used to guide decisions over the provision of health care.

A secondary outcome from a formal consideration of theories of justice is that the discussion of the theories is able to 'put right' a number of common 'misinterpretations' around theories of justice in the health care literature.

Firstly, whilst a number of commentators may suggest that a utilitarian approach is common place (i.e. seeking the greatest good for the greatest number) (e.g. Burls and Bradley, 2003), the utilitarian approach to social evaluation, using utility as a measure of well-being is very rarely, if ever, applied in the approaches documented to the allocation of health care resources. The common consequentialist approach in health care i.e. seeking to maximise the production of health outcomes, is often confused with a utilitarian approach, but this does not fit with the presentation of utilitarianism (e.g. Bailey, 1997). The consequentialist approach (looking at health outcomes), common in health care, is not consistent with the requirements for a utilitarian approach, as it does not have a theory of 'good' (i.e. utility and well-being) which places 'value' on consequences.

Secondly, the theory of justice proposed by Rawls (1971) is often cited in health care in the context of a general 'maximin' approach to the evaluation of welfare (allocation decisions) (e.g. Olsen, 1997) and this misrepresents Rawlsian theory. Rawls has distanced his theory from micro-level decisions, and has stated that the Rawlsian theory of justice is not aligned to the economists 'maximin' rule for decision making under uncertainty (Rawls, 1999).

Furthermore, whilst egalitarianism, and equality, are commonly stated general principles within the health care literature, it is not clear (from the literature reviewed) how they can be presented as a coherent theory of justice (Arneson 1989, Konow 2003). Egalitarian goals of equality are more likely to be a special case of other approaches to justice and fairness (Campbell, 2001), and to be higher level principles rather than a basis for distributive preferences.

The discussion of theories of justice and fairness sets out a stylized hierarchy of the inputs for a health technology appraisal process. The thesis concentrates on specific social values (criteria) for the evaluation of alternative actions, this area being the most neglected in the literature reviewed.

9.2.2 'Empirical ethics' literature review

The literature review presented (Chapter 4) has extended the current evidence syntheses in this growing area of 'empirical ethics'. It has taken a broader perspective than previous published reviews, and has presented a broader evidence base. The review finds that the evidence base is varied and complex, with many of the published studies being simple experimental studies, often of a potentially low methodological quality (e.g. small sample size, convenience samples). Whilst earlier reviews (e.g. Dolan *et al,* 2005) have commented on this fact, the current thesis has taken an additional step and characterised the literature against the quality of the methods employed in the studies. The review has provided a summary of the nature of the literature reviewed, and has quantified the level of studies that may be classed as low, moderate or high quality, in terms of the methods used in the published studies. The review suggests that over 50% of the identified empirical studies only satisfy a low quality threshold (for methods used). In less than 15% of studies the methods are judged to be of a high quality threshold (methods).

A number of important points have emerged from the literature review of empirical studies. Firstly, although there is a large degree of uncertainty in the results presented in the evidence base reviewed, the finding that respondents are prepared to trade-off efficiency gains (health gains) in order to allocate resources according to other distributive preferences (social values) is consistent across all forms of study, across different samples, different context-specific studies, and different methodological approaches. It is clear from this background research (literature review) that the commonly suggested maximand of health, and the maximisation of health gain subject to budget constraints (most frequently used for economic evaluation), is not an objective that is supported, as a dominant or sole objective, by the current empirical evidence base on distributive preferences (for health care allocation problems). The empirical evidence reviewed does strongly suggest that respondents to surveys (e.g. general public, decision-makers, clinical samples) are prepared to sacrifice efficiency gains to select what is interpreted as a fairer, or more equitable, distribution of resources. This leads to consideration of the social values that may be potentially useful for setting priorities when allocating limited resources across a number of competing demands. The empirical literature is less clear on this issue.

When wishing to identify social values that may inform a fair allocation of resources, the empirical literature provides firm support for an efficiency argument (i.e. health maximisation, within budget constraints), 'level of health improvement (gain)', and 'severity of health condition' as potential key generic social values. There is also some support for the importance of 'other treatments' being available (or not), this support being largely from the

health policy literature. Whilst there is a broad literature across a number of different potential distributive preferences (e.g. social role, desert, age), there is no strong support for these alternative priority or social preferences. This is especially the case when put in the context of the UK NHS, and the NICE health technology appraisal process. However, current research is largely simple and preliminary research, and further more in-depth research may lead to more informative findings against the range of social preferences that have been subject to research to date.

The review of the literature has indicated that one way much of the literature could be drawn together is under a general preference for 'fairness', which presents itself as a preference to give at least equal, if not preferential, priority to the patient group thought to be in the 'worst off' scenario. For example, the patient group who are in the most severe health condition and unable to benefit greatly from treatment (but do benefit). Other examples are comparisons of patients groups with similar health conditions, where one group can benefit a little (the worst off group) and one can benefit 'a lot', and comparisons of patients who can get a similar health benefit from treatment, but one group (the worst off) also has comorbid conditions, or is thought to be from a more deprived social group, or to have a shorter life-expectancy. There are many ways through which to interpret the worst off from two competing patient groups. But, across a wide number of empirical studies there is an indication that respondents are prepared to give equal treatment to those in a worse off group, at the expense of efficiency gains, or to give them greater priority. Even where there is a very marked difference in the health gains available between two competing groups, there is evidence that respondents do not wish to 'abandon' patient groups that would otherwise receive no treatment (even if the likely health gain is small). The review of the empirical literature has been used here to set a hypothesis that where evidence is reported against specific social values, e.g. to treat the most severely affected patient groups, that this may be a representation of a general preference to treat the worse off groups in choices presented to respondents. This hypothesis has been tested in an experimental study in the thesis (Chapter 5).

9.2.3 Examining a general preference for 'fairness'

In Chapter 5 of this thesis empirical evidence has been presented which extends the evidence base in 'empirical ethics'. It builds on the evidence available to support a preference counter to the maximisation of health, as a dominant social objective, and builds on the evidence base which supports the use of 'severity of health condition' as a basis for setting health care priorities. The study reported in Chapter 5 uses robust sampling methods, resulting in a representative sample of the general public, and face-to-face

interview techniques, rather than the simple convenience samples and self-complete methods used in earlier studies by Nord (1993) and Ubel (1999). Importantly, the study extends the current empirical literature to consider a hypothesis that evidence presented to support the presence of specific social values, such as the 'severity hypothesis', may be interpreted as a proxy preference for a broader preference for 'fairness', in support of the worse off patient groups in health care choices. Empirical evidence is presented to support this hypothesis.

The study considering severity of health, and the preference for fairness hypothesis, has also examined the response to give equal priority to the two competing groups. It has explored a hypothesis that a preference registered for 'equal' priority may be used to avoid a difficult choice, and that it may not be a true egalitarian preference. Through the use of additional response categories (i.e. 'to let others choose') the study is able to support an alternative hypothesis that a preference for equal priority is indeed a true egalitarian preference in this study.

9.2.4 'Weighing goods': Exploring the social value of health interventions in a discrete choice experiment

The DCE reported in Chapter 7, with results applied to a policy context in Chapter 8, provides a contribution to both a health economics and a health policy literature. The study has addressed a research need highlighted by a number of commentators (e.g. Sassi *et al* 2001, Schwappach 2002, Dolan *et al* 2005), and considered the relationship between numerous social values in a priority setting context. Using the terminology of Broome (1991), the empirical study considers numerous 'goods' and 'weighs' these goods, against one another in a decision making framework. The study explores the social value of health care interventions, when described using the attributes (social values) and levels applied in the DCE design.

The DCE approach has not been widely tested in the 'social' context, and especially not in samples of the general public, using concepts such as 'value for money' (cost effectiveness). Chapter 6 has set out the DCE framework, and outlined current practice for DCE in health care. There have been some recent applications of DCE in a social context, but none similar to the design used here, and none in a sample of the general public employing value for money as an attribute. This thesis has further tested the DCE framework in a social context, in a representative sample of the UK general public.

The study reported here demonstrates that it is feasible to use a DCE approach for the assessment of social value in health care decisions, and that the methods used were acceptable to the general public. It has also demonstrated that it is possible to use general data collection agencies (e.g. MORI) to administer survey questionnaires, in short face-to-face interviews.

The DCE study presents stated preference data, and output from discrete choice modeling, to provide a measure of the relative weight for the attributes used, across the attribute levels employed. All of the attributes, and levels, have proved to be statistically significant and important in the choices presented. The resulting regression coefficients are all consistent with the expected direction of effect i.e. a positive impact on utility as attribute levels move to greater (better) levels. The outputs from the discrete choice model (beta coefficients and odds ratios), provide a measure of the relative weight for each of the attribute level differences (from the reference case, i.e. worst case), indicating that 'level of health improvement' and 'value for money' are the most important attributes, with 'severity of health condition' the next most influential attribute. However, across the attribute levels for 'health improvement' and 'value for money' the results show that for some changes the attribute of 'severity of health' is of equal or greater weight. Whilst the attribute describing the availability (or not) of other treatment options was statistically important, it had a small impact on the utility of alternative health care interventions (as described by attributes), and can be seen as the attribute with the smallest weight in the decision making process (in the design used).

As well as an immediate interpretation against the attributes and levels, the DCE model is used to estimate the level of utility (desirability) for each of the health care interventions described using the full factorial in the DCE design (64 scenarios). This provides a relative measure, a ranking, of the utility associated with each of the interventions described (calculating a utility function for each of the 64 scenarios). Such data allows a judgment of which scenario is better than another, in a simple ranking, and it allows consideration of alternative competing scenarios using the utility estimate. It is possible to address 'betterness' in some way, but the utility estimate is against an arbitrary interval scale, and it is therefore not possible to say that one scenario is 'twice' as good as another, for example.

The findings from the DCE contribute to the economists and policy makers understanding of trade-offs across social values. The study offers a considerable addition to a sparse evidence base. At one level, the findings indicate that level of health improvement and value for money may provide a useful summary, and proxy, measure of social value. However, at

a further more detailed level the results show how other important attributes interact to give a more complex presentation of overall social value from health care interventions, described using the simple matrix of attributes and levels.

9.2.5 Health technology appraisal and 'betterness': Looking at 'strength of preference'

The thesis takes the DCE approach a step further by moving the analysis from the relative weights for attributes and levels indicated by the 'logit' function (i.e. regression output/coefficients), and applying a transformation of the logit function to estimate a set of probabilities (for the dependent outcome) across the full factorial design. This approach is consistent with the statistical presentation of the conditional logit model, applied in the analysis (McFadden, 1973). This is thought to be the first time that such a transformation, across a full factorial design, has been employed in a DCE in health care. Through the use of the probability scale, a mathematical scale with ratio level properties, the thesis has presented an approach to consider the strength of preference between potentially competing alternative health care scenarios (interventions).

The thesis has presented the estimation of a set of probabilities, and a demonstration of how such a set of probabilities can be used in a health policy context (NICE health technology appraisal programme) (Chapter 8). Whilst the estimation of the 'probability set' is clear and follows statistical techniques, the policy application requires judgments to be made against the attributes for specific health technologies. This thesis has demonstrated how such judgments could be made, and presented the use of the social value data against a range of health technologies. The thesis suggests that the data, or similar data from future studies, could be used by health care decision makers as part of what is already a considered appraisal process; using the current process to inform against the attributes for each health technology subject to appraisal. Through the data presented and the demonstrated policy application, and context, it is argued that the thesis makes a contribution to the health economics literature, and the health policy literature, bringing both environments together in an application of the theory underpinning discrete choice techniques.

9.2.6 Towards a better understanding of 'equity' in health care (health technology appraisal)

This thesis has risen to the challenge highlighted by a number of commentators, to seek out empirical information on the relationship of key social values in a health care priority setting context. In exploring the social value of health care interventions it has provided some

information on how the general public in the UK may trade-off across four potentially key generic social values. By looking at numerous social values the research presented here is able to offer a greater understanding of what it is that is meant by equity in the allocation of health care resources.

Equity is a much used word in the health economics and health care literature. However, clear and operational definitions of equity are not presented in the literature. Whilst the 'trade-off' between efficiency gains and distributive ideals is widely acknowledged as an 'equity versus efficiency' trade-off, such an understanding of equity is simple and largely unhelpful. Definitions of equity as the absence of inequity, or as a fair allocation of resources, or as the application of the definition of vertical or horizontal equity, are not operational definitions. For example, they are not helpful to those faced with frequent decisions over who should or should not have access to specific health technologies.

In the literature on theories of justice and fairness the word equity does not get used. There is discussion of the evaluation of alternative actions against that impact on social welfare, but there is no discussion of the state of the world that may be characterised as 'equity'. The thesis offers up a view of what it is that might be meant by equity in health care priority setting choices, specifically health technology appraisal in the UK.

The research presented here is an early step in a currently undeveloped area of health care research. It puts forward a view of equity in the allocation of health care resources as 'a state of the world in which key social values are balanced against one another'. As with the economists notion of 'equilibrium' as a state of 'balance', so through the use of multiple social values the state of 'equity' can be practically represented as a state of balance, between competing social values, and/or other objectives. This presentation of equity is consistent with the views of Young (1994) on equity as a "balance or compromise between competing principles" (p9).

In economics, equilibrium is simply a state of the world where economic forces are balanced. For example, market equilibrium is a condition where market price (e.g. established through a competitive market) is a balance between supply and demand, i.e. the market price is the equilibrium or market clearing price. The economics notion of equilibrium is generally fully understood, as a concept, by all who learn and/or practice economics. It is accepted as a theoretical ideal where multiple equilibria are possible, where equilibria are unstable and where market forces are always inhibited in some way from a simple economic equilibrium (e.g. externalities, market interference). Through the use of explicit social values, and data

(at some level) on the relative value of social values when traded-off against one another in a priority setting context, it is possible to have a greater understanding of what balance of competing values and objectives may move a health care decision maker towards the most equitable allocation of resources, given limited information.

This thesis supports the following definition of equity as:

A state of the world (in health care) where the opportunity costs of alternative uses of resources are balanced against competing stated social values, so that the potential gains offset the potential losses.

To make such a definition operational, and useful for practical decisions, it is clear that the relevant social values should be 'stated' and that any equity position has been arrived at after consideration of opportunity costs. The argument here (the thesis) is that the social values which are the basis for the assessment of equity must be explicitly stated, and the trade-offs across them made known, at least in a general way.

The definition set out here is general, in so far as it is possible to set out the assessment of equity in a number of different decision making contexts. The stated social values and trade-offs between them are not specific to this definition or statement, and can vary according to decision making context. The health policy approach presented, and the above working definition for equity, are both flexible and can be applied to different decision making settings (e.g. life saving versus life-enhancing), and different perspectives or viewpoints (e.g. societal, or decision maker perspectives). This general (to specific) definition of equity fits with the economists notion of 'equilibrium' and balance amongst competing forces, and it also fits with the general notion of a social welfare function set out by Nath (1973) i.e. a general statement of the objectives of a society, with some rough and ready idea of the relative weights of these objectives.

The concept of equity in health care will remain an ambiguous and much contested concept, with a wide variety of views and contributions. This point has been made by Young (1994), who has argued in the preface to his book titled '*Equity*', that the way the concept of equity is used, and the way the word equity is widely employed,[19] is very unclear, and he argues that in most instances it has no intrinsic meaning at all. He puts forward an entertaining discussion on the 'non-existence' of equity, against objective criteria and against academic

[19] Young (1994) considers equity in a broad way, and not specifically against health and/or health care. Although health care (e.g. allocation of kidneys among transplant patients) is part of his detailed consideration of equity.

theory. Young argues that where equity is used it is a subjective concept and cannot be analysed scientifically, and that there is no sensible theory about it ("certainly none that is compatible with modern welfare economics" *p.*xi). However, Young does see equity as a useful concept, aligning it to what he terms "a just social order" (*p.*xi), and he goes on (a whole book is thereafter devoted to the concept of equity) to characterise equity, as he wishes to address it, as an 'everyday' allocation problem, an issue of fair allocation. Young argues that it is possible to consider the meaning of equity 'in the small', as an everyday problem, without needing to solve the larger problems around social justice. He uses 'sharing rules' as a notion of equity in the division of jointly produced goods, stressing that equity is about a given situation and the importance of context. The context specific nature of equity, put forward by Young (1994), is supported by Konow's detailed consideration of theories of justice and fairness (Konow, 2000), which has stressed the importance of putting emphasis on the context of the evaluation of social welfare.

This thesis has set out in a systematic way to depict the concept of equity as a practical concept, an 'everyday distributive problem' (as Young has called it) distancing it from the more theoretical and philosophical notions of justice and fairness. The thesis argues that justice and fairness (the large) should not be neglected, and that these related issues should form the basis upon which a practical notion of equity (the small) can be addressed, in common and recurrent real-time decisions. All theories of justice noted in this thesis, whilst neglecting the practical necessities of decision-making, argue for a clear basis for evaluation. This thesis, through the emphasis on identifying social values to guide decisions seeks to make the basis for evaluation of health care priorities more explict; clearer to the communities affected by such priorities. Young (1994) has stated persuasively that every day problems and decisions need to be considered, and they cannot be delayed through a need to solve the broader problems of society. Whilst society must have an ongoing debate about justice and fairness, in various fields of social activity, it is necessary to proceed with a practical approach to everyday decisions. Equity sits in this latter activity, and this thesis offers a greater understanding of equity through exploring the social value of health care interventions.

9.3 Limitations with the research presented

The research methods used in the thesis have been clearly described in each Chapter. The research undertaken has been the sole work of the candidate, undertaken in the context of the thesis alone, and not for other purposes, or through other externally funded research projects. There are limitations with the scope of the research undertaken, and the methods

employed. These issues have been addressed in each of the relevant Chapters, however a summary is provided here, to accompany the above discussion of findings and contributions.

The introductory chapter highlighted that the scope of the thesis was limited to exploring the social value of health care interventions using the analytical framework of health economics, and the policy context of health technology appraisal in the UK NHS. It was acknowledged that the subject area was wide-reaching and that the thesis would be unable to cover all of the interesting, and deserving, areas of research involved. The thesis does not tackle issues of justice and fairness in a detailed way, other than where they set the context for consideration of social values in health technology appraisal, and consideration of equity in this context.

The introductory chapter presented a brief introduction to many of the key concepts involved in the framework used for the thesis. It is acknowledged that there is often debate and uncertainty around some of these issues (e.g. definition of health), and the rapidly developing nature of the health research, and policy advice, in some areas (e.g. process of health technology appraisal), and the changing organisational structure of the UK NHS. It is acknowledged that the theoretical and conceptual components of health economics (consumer theory, welfare economics) are often axiomatic and abstract in nature, and difficult to apply in a practical way in health policy decisions. Furthermore, the thesis begins with an open acknowledgement of the ambiguities of the terminology and language used around the issue of the social valuation of health and health care. All of these issues create difficulties when undertaking, and reporting research in this area.

Where the thesis explores general theories of justice and fairness it considers a range of alternative theories, and includes the most commonly referred to theories. However, the thesis does not have the opportunity to present a detailed description of these theories, with coverage limited to a summary description, and the presentation of summary findings from a thought experiment to relate them to the process of health technology appraisal. It is possible that other contributions from moral philosophy, and other disciplines, could inform on the place of justice and fairness in the area of health care.

Where the thesis undertakes and reports against a thought experiment, to consider health technology appraisal against theories of justice, it is accepted that there are limitations in the documentation available to inform on the health technology appraisal process used by NICE. While NICE present guidance documents which make much of their work relatively clear and transparent, there is an absence of information around the use of competing scientific and

social values. This has been noted by a number of other commentators (e.g. Dakin *et al*, 2006, Tappenden *et al, 2006*). During the course of this thesis (from 2002- present) there has been a welcome growth in the literature and documentation from NICE around social values, with some clear and specific guidance issued (NICE, 2005). This area of activity remains an area of current development within NICE (with other reports expected from the NICE Citizens Council, and further updates from NICE expected later in 2007 on the issue of social values).

Linked with the above point on theories of justice, is the scope of the literature review presented in Chapter 4. The literature review covers studies that present empirical findings against a range of social values reflecting distributive preferences. The growing analytical literature that seeks to inform on the evaluation of social welfare in health care has not been presented in the thesis. The thesis has placed importance on the empirical literature, as the most useful area from which to draw conclusions over distributive preferences. But, it is acknowledged here that the literature developing theoretical and/or analytical observations related to health care resource allocation decisions, may also be useful in exploring the social value of health care interventions.

The literature review presented in Chapter 4 has applied rigorous search methods, but it is also acknowledged that in such a broad and complex methodological review of this type, where trade-offs have to be made around sensitivity and specificity when searching, and where an element of judgment is needed when making decision over inclusion and exclusion of references, there is the potential for reviewer bias. It is possible that there are some studies that have not been identified, or included, in the review presented. It is also possible that there are informative research studies that have been published in journals in non-English language formats, and these are excluded from the literature reviewed. Again, it has not been possible, given the extent of the literature covered, to provide a detailed narrative review against all of the studies identified as part of the literature search, and included in the general review presented. However, the review stands as at least of an equivalent methodological standard to other previous reviews of a similar nature, where researchers have also had to contend with similar challenges.

Within the findings presented from the literature review of empirical studies, the thesis presents an approach to consider the quality categories of studies with respect to the research methods used. Whilst this is presented in the thesis as a useful addition to the literature characterising the evidence base, it is accepted here that such an approach is simple and open to criticism. For example, for its focus on only the methods employed in the

studies, its neglect of some subjective assessment of transparency in the studies identified, and for the use of only a small range of categories for assessment of methods (i.e. low, moderate, high).

The chapters presenting empirical research (5, 7 & 8) have been clear on the specific limitations for each of the areas of work presented. All empirical research of this type is open to some degree of criticism due to design issues, framing and context effects, as well as choices over methods for analysis, and subjectively driven elements of discussion and/or conclusions. In summary, the study presented in Chapter 5 has used earlier study designs, and applied and extended these designs to explore stated hypotheses. The study has limitations in terms of the general text used to frame the hypothetical choices presented to the respondents, and in so far as the differences between alternative scenarios in the choices are qualitative, and offer no opportunity to generalise findings in a quantitative manner. Discussion of research methods and the sampling process, and the sample used, are presented in the chapter. One limitation with this study, in common with the study presented in Chapter 7 is the absence of objective data on response rate in the sampling frame approached. This is a by-product of using a random selection of sampling areas, and a subsequent quota approach to obtain a representative sample of the general public.

The DCE presented in Chapter 7 has limitations in terms of the experimental study design used, the small number of attributes and levels, and the nature of the discrete choice framework employed. These have all been discussed in Chapter 7. DCE studies, like other approaches to the elicitation of preferences, are open to the challenge that the stated preferences are potentially a product of issues such as framing effects and heuristics (e.g. Lloyd, 2003, Bryan & Dolan, 2004). The design was specifically kept simple, given the immediate policy context. That other social values may be important, and are not included in the current design, should not detract from the contribution made by the research findings presented. But it is accepted that other issues may be relevant and are deserving of attention in future research (discussed below), together with methodological issues surrounding response rates, and tests of reliability, for example. Whilst the DCE was deliberately conceived as a generic design, taking an unlabelled approach to the choices presented, it may be that a more detailed design with specific labels could have provided further context specific evidence to inform particular areas of decision making (e.g. life saving conditions, compared to long term health conditions). However, this is one of the first studies of this kind and the current study has tested the framework, enabling future research can be undertaken with greater confidence.

The design used for the DCE was informed by a literature review, analysis of policy documents/reports, discussion with experts (DCE methods, decision-makers), and extensive pilot work. However, one of the limitations with the study is the fact that no in-depth interviews, or focus groups, have been used to inform the design, and subsequent survey. In this thesis it could be argued that there was recourse to other published focus groups (e.g. Cookson & Dolan 1999, Dolan *et al* 1999, Dicker & Armstrong 1995), however, it may also be a limitation with the approach reported here.

It is always difficult to bridge the gap between research results and policy relevance. Whilst this thesis has sought to address this issue, there are limitations with the approach developed and demonstrated in Chapter 8. The application of research findings around the policy context of the NICE health technology appraisal process, is based on a selected group of health technology appraisals already undertaken by NICE. The selection is based on the involvement of the candidate in certain technology appraisals, and the prominence of the appraisal topics (as contentious appraisal judgments). The examples are used to demonstrate policy relevance for research findings, and policy context, and are thought to offer a general scenario of relatively difficult health technology appraisals (i.e. difficult for various reasons, such as lack of demonstrated cost-effectiveness, or difficulties establishing QALY benefits).

In order to apply the findings from the DCE modeling, it has been necessary to make judgments across the attributes and levels for a range of health technologies, reported in NICE guidance. These judgments are of a subjective nature, and have been made by the candidate here for demonstration purposes. The judgments made are explicit, and open to challenge, but the important point is the suggestion that the organisations (or groups) responsible for making health technology appraisal decisions should be in a position to make an informed judgment across attributes and levels for each technology being appraised, and that the matrix presented should be broad enough to capture most, if not all, health technologies (albeit in a simple and generic way).

General, and broader, limitations with the research presented may involve the core elements comprising the framework used for the thesis. These are the analytical framework of health economics, and the policy context of health technology appraisal. It may be that findings are limited through the use of this specific research framework. Other disciplines, such as medical sociology, health management, psychology and wider contributions from moral philosophy and political economy, could help with a clearer view of the social value of health care interventions. It may also be that the health technology appraisal context used may not

reflect the general problems of health care resource allocation in the UK NHS. For example, using the technology appraisal context of NICE has enabled the research here to bypass considerations of 'budget impact', and 'numbers of people affected', when it comes to consideration of an 'allocation problem' (i.e. health technology appraisal). Such considerations may be more relevant in other organisational settings within the NHS. However, given the confines of this PhD thesis, it is argued that the use of the stated framework allows a focus on a topical and policy relevant area of the UK NHS.

The above limitations, often given more detailed coverage in the body of the thesis, should be taken into consideration in the context of the research area covered in the thesis (a broad and complex area of research), and the techniques available for the conduct of research (all techniques are subject to limitations). They should also be considered against the contributions suggested above, and the following suggestions for future research.

9.4 Future research

There area a number of future research suggestions arising from the thesis.

9.4.1 Further examination of theories of justice

The thesis has given only general coverage to the translation of the general theories of justice to the area of health care, and health technology appraisal specifically. It is suggested in the current research that there may be benefits from further consideration of general theories of justice in the context of the evaluation of resource allocation in health, to assist in the definition of guiding principles, and for insights around procedural justice and fairness. Further research, of an analytical and empirical nature, is recommended in this area of enquiry, together with further research to establish a clearer interpretation of the theory of utilitarianism and Rawls' theory of justice. There are currently research initiatives in progress (e.g. Professor J Coast, University of Birmingham) to further develop theories of extra-welfarism (or similar) in health care.

9.4.2 Empirical research on distributive preferences, and a general preference for 'fairness'

Whilst the current thesis, and previously published reviews, have described a growing empirical literature around social values in health care, many studies are simple experimental studies and there is still a research need for further empirical evidence using more rigorous methods with a greater emphasis on policy relevance. A particular emphasis should be placed on research studies that consider priority setting preferences involving a

range of social values, and offering the potential to inform on the relationship between key social values when trade-offs are necessary.

There are currently research initiatives in progress (e.g. Professor Cam Donaldson, University of Newcastle upon Tyne, Professor Paul Dolan, Imperial College London) to provide further empirical and analytical research findings on the potential basis for weighting QALYs (health benefits) in health care. Findings from these research projects, funded by the UK Department of Health, and UK NHS (NICE), are expected to be disseminated from Autumn 2007.

Of specific interest, is the research in the thesis that has pointed to a hypothesis for a broad preference for fairness, reflected as a preference for the worst off group (regardless of the specific reasons for being the worse off). More work is needed to explore this hypothesis. The thesis has considered the use of a general non-health description of 'worse off' (disadvantaged), together with treatment related health benefits, and further research is recommend against this category of group description, but also against other descriptions, preferably in single large studies. Such studies may be able to develop the hypothesis towards a general focus on fairness, which can be applied in specific decision making contexts.

9.4.3 Developing a basis for assessing the quality of empirical studies in 'empirical ethics'

Further research is recommended on the assessment of the quality of empirical studies in the area of social values and preferences. Such research should take a more rigorous approach to developing some simple categories through which the quality of research can be graded, for example through workshops and/or consensus methods amongst active researchers in this area. Policy makers, analysts, and journal editorial staff may find such research helpful, and it may also provide a firmer scientific basis for research in the area of social values and preferences, and a better foundation from which to inform health policy.

9.4.4 Development of a discrete choice study in a sample of the general public in the UK

To build on the DCE research reported in the thesis, it is recommended that a similar study be undertaken in a broader sample of the UK general public. Future research should use a random sample of the general public. It is recommended that any future study should use a large enough sample to allow investigation of response data for different decision making

contexts, for example in life-saving treatments, and in what may be regarded as non-life threatening or longer term chronic conditions. The current research provides a good foundation for future work, but future work may also wish to consider other social values for inclusion in the design. The inclusion of attributes surrounding 'budget impact' and 'numbers of people to be treated', may be expected to pose some problems in a sample of the general public, as they are attributes that would involve absolute numbers (e.g. £1 million, or 1,000 people) and the general public would not be expected to have a reference level for such information. The provision of information with a future survey may lead the respondent, and care would need to be taken over potential for bias and framing effects. However, research to investigate such issues is encouraged in a broader programme of research around social values in health care.

Future research is recommended across other relevant respondent groups of interest (e.g. clinical sample, decision-maker sample), in order to consider differences in response data, and to explore and contrast any differences.

In any future DCE research it is recommended that at some level (e.g. a sub-sample) a qualitative investigation of respondent choices is undertaken, to assess the reasons that choices are made against attributes, and attribute levels. This research would involve serious challenges, as respondents are often not able to articulate the reasons for choice patterns, but such research would add a greater understanding to DCE data.

In the DCE presented here the design was informed by literature, policy analysis, pilot work, and expert opinion, and in future research it may be helpful to use group interviews, or focus groups, to explore the design of the experiment (presentation of survey).

9.4.5 Consideration of the health policy customer

The thesis has referred to the health technology appraisal process in the UK (specifically NICE), as the decision making context, however there are other decision making contexts that are relevant for a discussion of social values, and there use in everyday allocation problems. Research is recommended to explore the health policy customers and contexts in which any future research is thought to be potentially helpful. There is currently research being undertaken against the NICE appraisal process (e.g. Bryan *et al,* 2006), and research at a hospital trust level (e.g. Jenkins & Barber, 2004). Further research in this area will assist in making future empirical studies into social values more useful, and policy relevant.

Research of a mixed methods form is recommended, given the often complex organisational settings for decision making.

The research presented here has been against the context of the UK NHS, however other countries may also find research of this type helpful to guide their decision making processes (e.g. New Zealand, Australia, Canada).

10 CONCLUDING THOUGHTS

The many challenges to the UK NHS health care budget are not going to diminish in the coming years. For example, there will be pressures from more frequent technological advances, often resulting in the emergence of beneficial but expensive health technologies, and also pressues from a changing demographical profile, and from a better informed population with greater expectations of the health service. There is already heightened awareness, across all groups in society, of the need for the organisations responsible for the delivery of NHS and social care to make difficult decisions over the coverage of health care interventions. The public debate, as witnessed over some high profile media stories during the last 12-months (e.g. drugs for Alzheimer's disease, and drugs for breast cancer), will no doubt become more public, and there may be a greater need for the organisations responsible for spending public funds to be more explicit over the basis for decisions which will result in some people being denied access to clinically effective health care.

Whilst the NHS at a general level, and organisations such as NICE at a more devolved decision making level, have made clear statements that they will be guided by the views of the general public, such statements will need to be supported by a demonstrated willingness to collect the views of the public, and to apply such views to health policy decisions. These issues pose challenges for the health policy and health services research communities, but such challenges can be met. This thesis seeks to inform research in this area.

In summary, the main message from this thesis is that there are a number of clear social values (social goods) that are important when setting priorities in health care. From the information currently available from the general public, the relationship between social values is not fixed, and social values may interact with one another according to the choices to be made, and the decision making context. The thesis has demonstrated that it is feasible and acceptable to elicit the views of the general public around health care priorities, and the basis for setting priorities, and it has demonstrated that the discrete choice framework offers an opportunity to elicit such values. The thesis adds to the growing body of evidence that supports the view that the distribution of health care resources across different groups of people is important i.e. 'distribution matters', and that the maximisation of health gains (e.g. QALY gains) within budget constraints is not consistent with social values and preferences. The research presented in the thesis does support a strong influence, on the setting of priorities, from social values reflecting 'level of health improvement', and 'value for money'. Findings suggest that these social values can offer a great insight to the general societal

value to be attached to health care interventions. Findings also demonstrate that other factors are also important and that achieving an 'equitable' solution is about balancing competing social values, amidst consideration of the opportunity costs.

The thesis has also demonstrated that it is possible to bridge the gap between research findings and health policy, using DCE data and the NICE health technology appraisal process. It has done this using a simple research design, and it may be that there are also trade-offs between scientific complexity and the acceptability of methods to policy makers, when wishing to bridge the 'research to policy' gap.

Research into the areas of social welfare evaluation, the social valuation of health outcomes and consequences, priority setting criteria in health care, social values and equity in health care, is relatively new and undeveloped. A large number of interesting research activities are developing in these areas of research, with a considerable growth in activity seen during the time spent on this thesis. The thesis has contributed to some of these research areas, and there are a large number of research initiatives emerging to explore the theoretical and empirical aspects around health-related social welfare, and the societal value of health care. It is hoped that in the years to come there is a clearer picture over the use of the term equity, and the use of equity versus efficiency trade-offs in health care allocation problems.

APPENDICES

APPENDIX 1

APPENDIX 1

THOUGHT EXPERIMENT: THEORIES OF JUSTICE AND THE NICE HEALTH TECHNOLOGY APPRAISAL PROCESS

A1.1 A 'Thought Experiment'

To examine the theories of justice presented in Chapter 3 in the context of health care and the process of health technology appraisal, a "thought experiment" has been undertaken by asking:

(a) What would the health technology appraisal process look like if it was firmly based on [each of the theories]?

and,

(b) Does the present process for health technology appraisal address (to any extent) the concerns of [each of the theories]?

In the above thought experiment each of the theories of justice are considered in the context of observations on the conduct of health technology appraisal in the UK NHS.

In the UK the NICE health technology appraisal programme is the most public and prominent process for health technology appraisal, and the NICE technology appraisal programme is used here to consider the theories of justice in the above thought experiment.

A1.2 The National Institute for Health and Clinical Excellence (NICE) Health Technology Appraisal Process

NICE has been introduced in Chapter 2 of this thesis, and is discussed throughout the thesis (Chapter 3, Chapters 4, and 8). It was set up as a special health authority for England and Wales in 1999. It has three main functions: to appraise new technologies, to produce and approve guidelines, and to encourage improvement in quality (Raftery, 2001). As a special health authority it is part of the Department of Health. It's technology appraisal programme is charged with the task of making recommendations on use in the NHS of particular health technologies (selected new and established technologies) based on appraisal and assessment of their 'clinical and cost-effectiveness'.

NICE is becoming an increasingly important element within the UK Department of Health, and a prominent source of guidance for the NHS. The NICE health technology appraisal process, offers an opportunity to consider the procedural issues related to the operation of

NICE (as an authoritative committee) and also the basis upon which recommendations are made (i.e. decision variables).

The NICE appraisal committee(s) appraises technologies on the basis of submissions and commissioned independent assessments. Companies who manufacture the relevant technologies and professional and patient groups are invited to submit evidence. The form of evidence is specified as clinical outcomes, cost per quality adjusted life year (QALY), and impact on the cost of NHS and personal social services. Academic centres are commissioned to provide independent assessments to NICE. The views of both professional and patient groups are taken into account, and all relevant groups attend the meetings of the NICE Appraisal Committee (AC).

In light of the evidence before it the NICE AC will reach a judgement as to whether, on balance, the intervention can be recommended as a cost-effective use of NHS resources in general, or for specific indications, or for defined subgroups of patients (NICE 2004). The AC will, as appropriate, estimate the net impact on both costs and benefits of any new intervention under consideration. This judgement is referred to as the appraisal determination, and once the appraisal process is complete (including any appeal) the determination is submitted to the Institute as the basis of its guidance to the NHS in England and Wales.

In reaching its judgement NICE will have regard to the factors listed in the Secretary of State and National Assembly for Wales' Directions, namely (NICE 2000):

- the Secretary of State and National Assembly for Wales' broad clinical priorities (as set out for instance in National Priorities Guidance and in National Service Frameworks, or any specific guidance on individual referrals);
- the degree of clinical need of the patients with the condition under consideration;
- the broad balance of benefits and costs;
- any guidance from the Secretary of State and National Assembly for Wales on resources likely to be available and on such other matters as they may think fit;
- the effective use of available resources.

A further factor, which the Institute will take into account in its appraisal, is the wish to be sympathetic to the longer-term interests of the NHS in encouraging innovation of good value to patients.

The detail around the NICE process of technology appraisal, used for this thought experiment, has been informed using published policy documents on process and methodological guidance (NICE 2004), an overview of NICE guidance published up to mid-2006 (discussed further in Chapter 4), and the related recent NICE guidelines on social value judgments within the technology appraisal process (NICE, December 2005, Rawlins & Culyer 2004). Published studies evaluating and observing the NICE technology appraisal process have also been consulted (Sheldon *et al* 2004, Bryan *et al* 2006).

A1.3 Theories of Justice

A1.3.1 Utilitarianism

(a) What would health technology appraisal (HTAP) look like if it was firmly based on 'utilitarianism'?

Health technology appraisal in a utilitarian framework:

- The objective of the HTAP process under utilitarianism would be to maximise overall utility, subject to the resource constraints present in the NHS.
- Utility would be based on the consequences of the alternative actions being appraised. The benefits and dis-benefits (e.g. adverse events) associated with interventions would form the basis for the assessment of the utility gained (lost). For example, benefits may be increased life-expectancy, improved overall quality-of-life, improvements in specific dimensions of emotional and physical functioning, clinical markers of disease improvement (response to treatment), life-style flexibility.
- Consequences would be considered in conjunction with a theory of the personal good.
- HTAP would be based on a notion of utility (personal good, well-being) that captures the goodness (and badness) associated with the use of health technologies being appraised, or the absence of the technology (the outcome for the next best comparator action).
- The HTAP process would not use 'non-utility' information, for example information on the socio-economic or demographic characteristics of persons, to inform appraisal decisions.
- The appraisal process would be based on an accepted premise that interpersonal comparisons could be made between the states of the world where the health intervention was made available and where it was not made available, i.e. that utility

gains (losses) from health interventions could be aggregated and compared across different actions.

- The utility measure for HTAP would be assumed to be comparable across individuals and the value of the utility would be the same for all persons regardless of descriptive and clinical characteristics (i.e. distributive indifference).

(b) Does the present process for HTAP address (to any extent) the concerns of utilitarianism?

- The objective of the NHS is not solely to act as a utility maximising organisation. The NHS has other objectives to contend with.
- The HTAP process is based around consequences (e.g. life-expectancy, QALYs, health-related quality-of-life). But other information unrelated to consequences is also considered in the appraisal process.
- Consequences are not generally linked to a theory of the personal good (i.e. to how consequences are valued). The most common interpretation of the QALY is as a measure of health-related quality-of-life (health-related welfare), yet there are good empirical grounds for questioning the use of the QALY (as commonly derived) as a measure of utility associated with health gains (e.g. Nord 1991, 1993). Some argue that the QALY merely reflects an approximate measure of health production and not the utility (or social value) it offers those persons receiving QALY gains (Nord 1991).
- The HTAP process uses 'non-utility' information in the appraisal of health technologies (e.g. information on patient characteristics, health gains, clinical effectiveness data). Although decision-makers use, as a basis for HTAP judgements, data on the cost-effectiveness of interventions, where data on outcomes (e.g. life-years gained, QALYs) is combined with data on costs, there is still a role (in the overall appraisal process) for the clinical-effectiveness data, and this is data that can be classified as non-utility data. Furthermore, there is some interest in the use of non-utility data to inform on the relative value of health outcomes, such as the QALY, across different subgroups of patients, and again such data would be regarded as non-utility data and is regarded as inconsistent with the utilitarian approach. Given the deliberative assessment process (although not explicitly documented in many cases) and the fact that a view on the benefits of health technologies is formed in a multi-dimensional manner, not just from utility data (e.g. QALYs as a proxy for utility), there are good grounds for arguing the HTAP process is inconsistent with a utilitarian framework (i.e. theory of personal good, principle of evaluative consequentialism).

- In the process of HTAP, consistent with utilitarianism, decision-makers are prepared to make interpersonal comparisons between individuals (groups) based on different courses of action, and there is evidence to support the fact that where a summary measure of outcome is used (i.e. the QALY, and/or the cost per QALY) the decision-maker is prepared to consider that the unit of benefit is of equal value to all (NICE 2004, NICE 2005).
- The process of HTAP regards the QALY as being of equal value to all persons. This could be interpreted as being consistent with the utilitarian principle of 'distributive indifference'. However, distributive indifference is based on the notion of diminishing marginal utility, whereby those who value an extra unit of output the most are the ones who should receive it. In HTAP there is no notion of the diminishing marginal value of a QALY gain. Whilst there may be an element of support for the principle of distributive indifference this is indifference across consequences rather than the value attached to them.
- Indications are that the HTAP process does take into account, in a significant way, the maximisation of health gain (health-related welfare), and this may find some association with the utilitarian ideal of maximising utility. However, as above, HTAP has at its core a number of other important objectives and it is difficult to understand how these are traded-off in the context of the appraisal decision. Furthermore, it is questionable whether a measure such as the QALY is consistent with the utilitarian requirement for a theory of the personal good (the principle of utility), and it is clear that a range of non-utility data is also present in the decision making process, and that outcomes may be based on a broad definition of consequences, or indeed on information that is not directly linked to the consequences of an intervention.
- In conclusion, there are good grounds for arguing the HTAP process is inconsistent with a utilitarian framework.

A1.3.2 Rawlsian Justice

(a) What would health technology appraisal (HTAP) look like if it was firmly based on 'Rawlsian justice'?

Health technology appraisal in a Rawlsian framework:

- The HTAP process under a Rawlsian framework would be based on a strong societal view of 'justice as fairness' with a firm emphasis on procedural justice.
- A Rawlsian framework for HTAP would recognise, and make explicit, basic liberties with respect to health (in so far as the provision of health technologies was relevant to those liberties), and acknowledge that such liberties could not be sacrificed (for example, by the few) in order to achieve a greater sum of advantages (for example, to favour a greater number of persons).
- Rawlsian HTAP would accept a priority of justice (fairness) over efficiency; with liberties, fair process, key distributive principles such as the 'difference principle', all taking priority over the consideration of efficiency.
- Following Rawls' theory of justice the institutions providing the means of applying a fair process in HTAP would be expected to offer a basis for ensuring equality of opportunity (i.e. within the process), and would be required to fulfil the requirements of decision making behind a veil of ignorance (e.g. decision makers and those affected by the decisions are 'free and equal'). The framework would provide the principles and the background process under which HTAP would operate (i.e. the 'macro' framework), yet within specific HTAP judgements (i.e. the 'micro' framework) there would be an opportunity to base decisions on numerous and varied criteria.
- Only after consideration of basic liberties and fair process (equal opportunity) would Rawlsian HTAP consider the issue of distributive shares, accepting in the first instance that any inequality should favour the least advantaged members of society (i.e. Rawls' difference principle). These 'least advantaged', or 'worst off', may be those with the severest current health conditions (e.g. greatest pain or suffering), those with the worst prognosis (e.g. conditions with very low life-expectancy), or those currently deprived of needed health care, the Rawlsian process of HTAP would be required to make it clear under what premise the least advantaged were to be defined.
- HTAP in a Rawslian framework (for both macro and micro considerations) would not view any one criterion as dominant in the evaluation of health technologies (e.g.

maximising health gain would not be the dominant basis for the consideration of technologies within the HTAP process).

- The Rawslian framework is based on a social contract model, where persons (in a well-ordered society) are presumed to act justly and to do their part to uphold just institutions (i.e. there are general obligations on citizens). However, given the complex nature of health behaviour this social contract requirement would be expected to have a 'weak' interpretation with respect to health, offering an opportunity for those who may in some part be responsible for their own ill-health (or prospect of ill-health) to have equal rights with respect to health care, unless gross negligence of some form is present (even in such cases, e.g. self-harm, there will be extenuating circumstances with respect to health care).

(b) Does the present process for HTAP address (to any extent) the concerns of a Rawlsian theory of justice?

- HTAP in a Rawslian framework (for both macro and micro considerations) would not view any one criterion as dominant in the evaluation of health technologies (e.g. maximising health gain would not be the dominant basis for the consideration of technologies within the HTAP process).
- Considering the Rawlsian theory in a health care perspective, would require a simple societal structure (institutions) with respect to health. It may be that the major institutions of relevance to HTAP would be the Department of Health, the NHS and major institutional components of the NHS such as NICE. But, from the current processes of health technology appraisal it is not possible to see the application of Rawlsian principles.
- Decision makers undertaking health technology appraisal do not appear to make any explicit policy statements encompassing the Rawlsian principles surrounding basic liberties, equality of opportunity (within process), or on the Rawlsian notion of priority of justice over efficiency, for example on priority being given to the least advantaged persons (as defined by the HTAP institutions).
- There is an absence of explicit statements surrounding the Rawlsian priority of justice over efficiency, and it would appear, through the emphasis placed on cost-effectiveness, and health maximisation, that institutions value efficiency arguments highly in the appraisal process. This reliance on dominant criteria in HTAP i.e. clinical and cost-effectiveness, and the prominent use of measures of health gain (e.g.

QALY), lead to a conclusion that the present approach to HTAP is not consistent with the Rawlsian 'pluralist approach'.

- In the approach to HTAP the prominence of arguments related to distributive justice (e.g. allocation of resources, distribution of health outcomes), rather than on the central issue of procedural justice (i.e. underlying principles) would indicate that HTAP does not address the major considerations of the Rawlsian approach of 'justice as fairness'.
- Whilst there are indications that some of the elements of the Rawlsian framework may be present in part (i.e. the institutional arrangements, and an element or process), there appears little to support a view of the current HTAP process as Rawlsian.

A1.3.3 Nozian Justice

(a) What would health technology appraisal (HTAP) look like if it was firmly based on 'Nozian justice'?

Health technology appraisal in a Nozian framework:

- A Nozian framework for justice would not see the provision of health care via an extensive state as consistent with the entitlement theory of Robert Nozick. Nozick argues for a minimal state, and in so doing would not agree with the provision of health care dominated by government intervention.
- In a Nozian framework there would need to be collective provision of health care using privately funded health care providers, such as the organisation of care in US insurance based health maintenance organisations (HMOs), and the charitable provision of health care. Such providers would still have a requirement (depending on constraints) to appraise the costs and consequences of health technologies in order to decide whether they should be included in the coverage of specific health care plans. In such circumstances HTAP may be undertaken by these provider groups, or groups of providers (and or consumers) may act together in voluntary organisations (committees) to appraise health technologies (e.g. to address principal agent problems).
- HTAP in a Nozian framework would make explicit the underlying principles of justice upon which HTAP would be founded.
- Where a Nozian framework were used in HTAP the process would be based on a set of underlying principles, there would be no dominant criteria that would lead to a 'patterned' approach to evaluation (e.g. no place for a single dominant maximand [such as QALYS, or utility], or specific weights to be attached to outcomes/arguments).
- HTAP in a Nozian world would place emphasis on the historical perspective, using a broad information base (i.e. much broader than individual utility/preference data), it would consider, for example, how a person had arrived at a state of ill-health and what health care services had been received over time.

(b) Does the present process for HTAP address (to any extent) the concerns of a Nozian theory of justice?

- The present process of HTAP does not appear to address, to any extent, the concerns of Nozick's theory of justice.
- In the UK the NHS is seen as a cornerstone of health care provision, with over 90% of health care provided publicly through the NHS (ref). Although some elements of private provision are prominent (e.g. dentistry), there is strong support for a publicly provided NHS. Such a system, with financial constraints, leads to the need for some central (regional) element of health technology appraisal, in order to limit variations in service provision and access. The UK NHS is inconsistent with Nozick's theory of a minimal state.
- The NHS, although characterised within the broader Welfare State of looking after citizens from 'cradle to grave' does not place great emphasis on what Nozick terms the historical perspective. The process of HTAP, although interested in a persons history, and circumstances takes a current-time-slice view of a persons health care needs. A historical perspective may be a consideration in some situations, but health technology appraisal would seem to consistently take a current-time-slice view of patient groups (i.e. the patient group as they are presently described), not how they arrived in their health state, and an end-state view of resource allocation. Both views are inconsistent with Nozick's theory. For example, Health technology appraisal tends to view patient groups at a population (or regional) level based on their present health state and present needs e.g. patients with mild to moderately severe Alzheimer's Disease, or patients with heart disease and a cholesterol level (or blood pressure level) above a certain threshold.
- Health technology appraisal tends to operate on a set of objectives (e.g. maximise health gain, whilst minimising cost), and could therefore be characterised as patterned. Therefore HTAP is currently inconsistent with Nozick's view that evaluation of social arrangements [Nozicks holdings of goods] should be from a historical unpatterned perspective.

A1.3.4 Egalitarianism

(a) What would health technology appraisal (HTAP) look like if it was firmly based on 'egalitarianism'?

Health technology appraisal in an egalitarian framework:

- HTAP in an egalitarian framework would be based around an over-riding principle of equality amongst persons.
- In an egalitarian framework the particular presentation of HTAP and the process of HTAP would be based on what it was that was being equalised within the framework of justice.
- A set of underlying principles and objectives with respect to equality amongst persons would be stated as a foundation for HTAP.
- As a minimum it may be expected that a weak presentation for egalitarianism in HTAP would be an acceptance of the notions of vertical and horizontal equity; that patients who are alike in relevant respects be treated in like fashion and that patients who are unlike in relevant respects be treated in *appropriately* unlike fashion (Culyer, 2001).

(b) Does the present process for HTAP address (to any extent) the concerns of an Egalitarian theory of justice?

- In the HTAP process there is no clear view or statement of what the overriding principle of equality would be i.e. what should be equal (e.g. Culyer 2001, p75).
- There is not statement or clear understanding around the common principles of horizontal and vertical equity. In the context of the NHS [and NICE] these common principles must be qualified in terms of the comprehensive-ness, and coverage of the NHS services, and some understanding of what is meant by 'like treatment' is needed (Culyer 2001).
- If equality is linked to 'need' as many state (e.g. Culyer 1998, 2001), it would appear that need itself is not clearly defined in the NHS and in the NICE appraisal process.
- would be stated as a foundation for HTAP.
- Whilst health technology appraisal may place emphasis on equality in various dimensions (equality of access, equality of a unit of health outcome e.g. QALY), within the process of health technology appraisal there seems little evidence of explicit equality objectives when considering issues related to the appraisal of technologies and the restrictions on access to (use of) technologies.

- There may be a case for arguing that equality objectives are inconsistent with a process which does not make a health technology available to all who are able to benefit from it; acting, for example, to limit access to an intervention on the basis of capacity to benefit, or by age, or disease severity.

A1.3.5 Extra-welfarism

(a) What would health technology appraisal (HTAP) look like if it was firmly based on 'extra-welfarism'?

Health technology appraisal in an extra-welfarist framework:

- An extra-welfarist framework for HTAP would use an information base that was broader than utility information alone, as a means of assessing well-being and social welfare.
- An extra-welfarsist framework for HTAP would cover the use of both utility information and non-utility information.
- The capabilities approach advocated by Sen (1992) would see the process of HTAP place functionings and capabilities as a basis for assessing well-being and evaluating alternative outcomes.
- If the extra-welfarist approach from Culyer (1989) were to be used for HTAP the evaluation of alternative actions would be based on 'health' as a maximand, subject to budget constraints. It would also see a 'decision-maker' perspective for the evaluation of alternative actions.

(b) Does the present process for HTAP address (to any extent) the concerns of an extra-welfarist theory of justice?

- The HTAP process does use an information base that is broader than utility alone, and it includes non-utility information in the assessment of health technologies.
- However, it appears that non-utility information alone may be the basis for assessment, rather than using utility and non-utility information (assuming QALYs are a measure of health production, and/or preference, but not the value attached to those consequences).
- Sen's theory is a special kind of extra-welfarist approach, but we do not see in the process of health technology appraisal any consideration of what would be regarded as a vector of functionings, nor subsequent capability sets. Although it may be argued that there are often undercurrents of opinion, or outline sources of evidence, on what the true value of a health technology may be to potential recipients of the benefits, and these could be thought of as 'functionings' and 'capabilities'.
- The HTAP process would appear to fit with the notion of extra-welfarism set out by Culyer (1989), but it is not clear how consistent the primary outcome measure for

health i.e. the QALY is with Culyer's suggested maximand of 'health'. However, Culyer does offer a broad interpretation of health, as he suggests that through a 'decision-maker' perspective the values that are important for evaluation are those specified by the decision-maker.

- It is not clear to what extent the extra-welfarism of Culyer is consistent with the foundations of welfarism, and to what extent the absence of utility information from the NICE HTAP process limits such an interpretation of extra-welfarism in practice.

APPENDIX 2

APPENDIX 2

Chapter 4 – Literature Review

Literature Searching – up to January 2005

Searching undertaken 7/01/05

Limited to English language

Databases and years searched	Date searched and search files	Number of hits
Medline (OVID) <1966 to November Week 3 2004>	1 (health adj1 (polic$ or decision$)).mp. [mp=title, original title, abstract, name of substance, mesh subject heading] (7511) 2 health care.mp. [mp=title, original title, abstract, name of substance, mesh subject heading] (113129) 3 Health Policy/ or Health Priorities/ (32049) 4 health care rationing/ (8239) 5 (health technolog$ and (appraisal$ or assessment$)).mp. (473) 6 or/1-5 (148170) 7 (equity or justice or fairness).mp. (7334) 8 (public adj3 (preference$ or attitude$)).mp. [mp=title, original title, abstract, name of substance, mesh subject heading] (983) 9 priority setting.mp. [mp=title, original title, abstract, name of substance, mesh subject heading] (507) 10 (social value$ or societal value$).tw. (503) 11 7 or 8 or 9 or 10 (9242) 12 6 and 11 (2658) 13 (letter or comment).pt. (602182) 14 12 not 13 (2586) 15 limit 14 to (human and english language) (1898)	1898
Embase (OVID) <1980 to 2005 Week 01>	1 (health adj1 (polic$ or decision$)).mp. [mp=title, abstract, subject headings, drug trade name, original title, device manufacturer, drug manufacturer name] (5167) 2 health care.mp. [mp=title, abstract, subject headings, drug trade name, original title, device manufacturer, drug manufacturer name] (78056) 3 Health Care Planning/ (12978) 4 Health Care Organization.mp. [mp=title, abstract, subject headings, drug trade name, original title, device manufacturer, drug manufacturer name] (21513) 5 (health technolog$ and (appraisal$ or assessment$)).mp. (399) 6 or/1-5 (108153) 7 (equity or justice or fairness).mp. (6099) 8 (public adj3 (preference$ or attitude$)).mp. [mp=title, abstract, subject headings, drug trade name, original title, device manufacturer, drug manufacturer name] (757) 9 priority setting.mp. [mp=title, abstract, subject headings, drug trade name, original title, device manufacturer, drug manufacturer name] (395) 10 (social value$ or societal value$).tw. (350) 11 7 or 8 or 9 or 10 (7518) 12 6 and 11 (1831) 13 (letter or comment).pt. (277956) 14 12 not 13 (1813) 15 limit 14 to (human and english language) (1220)	1220
Medline in process <January 5, 2005>	Medline search run as above	68

Databases and years searched	Date searched and search files	Number of hits
Web of Knowledge ISI *SCI, SSCI: Timespan=1981-2004*	#1 TS=(health SAME (polic* or decision*)) #2 TS=health care #3 TS=(equity or justice or fairness) #4 TS=(public SAME (preference* or attitude*)) #5 TS=priority setting #6 TS=(social value* or societal value*) #7 #1 or #2 #8 #3 or #4 or #5 or #6 #9 #7 and #8 SSCI - 1655 2003-2005 367 2001- 2002 295 1998-2000 397 1995-1997 299 1990-1994 246 1981-1989 49 SCI (not SSCI) 1981-2005 298	SCI & SSCI 1953
HMIC (ovid) 1979 – present Searched 10/01/05	1 (health adj1 (polic$ or decision$)).mp. [mp=title, other title, abstract, heading words] (3339) 2 health care.mp. [mp=title, other title, abstract, heading words] (21291) 3 (health technolog$ and (appraisal$ or assessment$)).mp. (364) 4 or/1-3 (23810) 5 (equity or justice or fairness).mp. (2634) 6 (public adj3 (preference$ or attitude$)).mp. [mp=title, other title, abstract, heading words] (234) 7 priority setting.mp. [mp=title, other title, abstract, heading words] (823) 8 (social value$ or societal value$).tw. (56) 9 5 or 6 or 7 or 8 (3636) 10 4 and 9 (1187) 11 limit 10 to article (389)	389
EconLit 1969-2004/11 Searched 10/1/05	(((social value*)or(societal value*)) or ((Equity or justice or fairness)or(public adj (preference* or attitude*))or(priority setting*))) and ((health adj (polic* or decision*))or(health care))	350 (saved as text file)
	Total refs in Refman database equity searches after de-duplication	**3471**

Hits = 3,471 + 350 (EconLit, most of which were duplicates of other database options)

Summary of First Literature Search: Following removal of other duplicates 3,404 titles/abstracts available.

***Update search below**

'Update' Literature Searching – up to February 2007
Searching undertaken 26/02/07

Limited to English language

Databases and years searched	Date searched and search files	Number of hits
Ovid MEDLINE(R) <1996 to February Week 2 2007>	1 (health adj1 (polic$ or decision$)).mp. [mp=title, original title, abstract, name of substance word, subject heading word] (19912) 2 health care.mp. [mp=title, original title, abstract, name of substance word, subject heading word] (193207) 3 Health Policy/ or Health Priorities/ (20199) 4 health care rationing/ (4559) 5 (health technolog$ and (appraisal$ or assessment$)).mp. (562) 6 or/1-5 (208428) 7 (equity or justice or fairness).mp. (8540) 8 (public adj3 (preference$ or attitude$)).mp. [mp=title, original title, abstract, name of substance word, subject heading word] (467) 9 priority setting.mp. [mp=title, original title, abstract, name of substance word, subject heading word] (443) 10 (social value$ or societal value$).tw. (337) 11 7 or 8 or 9 or 10 (9666) 12 6 and 11 (3713) 13 (letter or comment).pt. (358164) 14 12 not 13 (3556) 15 limit 14 to (human and english language) (2864) 16 limit 15 to yr="2005 - 2007" (628) 17 15 (2864) 18 limit 17 to yr="2004" (288) 19 from 18 keep 1-288 (288) 20 from 16 keep 1-628 (628)	288 (2004) 628 (2005-07)
EMBASE <1996 to 2007 Week 08>	1 (health adj1 (polic$ or decision$)).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name] (4171) 2 health care.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name] (233042) 3 Health Care Planning/ (14694) 4 Health Care Organization.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name] (26209) 5 (health technolog$ and (appraisal$ or assessment$)).mp. (537) 6 or/1-5 (234093) 7 (equity or justice or fairness).mp. (7066) 8 (public adj3 (preference$ or attitude$)).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name] (424) 9 priority setting.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name] (410) 10 (social value$ or societal value$).tw. (257) 11 7 or 8 or 9 or 10 (8053) 12 6 and 11 (3318) 13 (letter or comment).pt. (228082) 14 12 not 13 (3230) 15 limit 14 to (human and english language) (2458) 16 limit 15 to yr="2004" (344) 17 15 (2458) 18 limit 17 to yr="2005 - 2007" (771)	344 (2004) 771 (2005-07)
Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations <February 23,	1 (health adj1 (polic$ or decision$)).mp. [mp=title, original title, abstract, name of substance word] (288) 2 health care.mp. [mp=title, original title, abstract, name of substance word] (2984) 3 (health technolog$ and (appraisal$ or assessment$)).mp. (41)	82

2007>	4 or/1-3 (3242) 5 (equity or justice or fairness).mp. (341) 6 (public adj3 (preference$ or attitude$)).mp. [mp=title, original title, abstract, name of substance word] (31) 7 priority setting.mp. [mp=title, original title, abstract, name of substance word] (25) 8 (social value$ or societal value$).tw. (22) 9 5 or 6 or 7 or 8 (409) 10 4 and 9 (88) 11 (letter or comment).pt. (16989) 12 10 not 11 (88) 13 limit 12 to english language (82)	
HMIC Health Management Information Consortium <January 2007>	1 (health adj1 (polic$ or decision$)).mp. [mp=title, other title, abstract, heading words] (4184) 2 health care.mp. [mp=title, other title, abstract, heading words] (23745) 3 (health technolog$ and (appraisal$ or assessment$)).mp. (496) 4 or/1-3 (27002) 5 (equity or justice or fairness).mp. (2986) 6 (public adj3 (preference$ or attitude$)).mp. [mp=title, other title, abstract, heading words] (267) 7 priority setting.mp. [mp=title, other title, abstract, heading words] (908) 8 (social value$ or societal value$).tw. (69) 9 5 or 6 or 7 or 8 (4092) 10 4 and 9 (1364) 11 limit 10 to (yr="2004 - 2007" and article) (68)	68
	Total refs after de-duplication	**1407**

These refs (n=1407) considered for update search. Overlap of original and update search over 2004, to capture any delayed entry references in earlier search (as advice from Information Scientist – Alison Price, Univeristy of Southampton).

APPENDIX 3

EMPIRICAL STUDY 1: MAIN QUESTIONNAIRE - VERSION 1

ASK ALL

We would now like you to consider a similar but slightly different type of question. It involves setting priorities, and again it involves a difficult choice. However, please note that there is no right or wrong answers and that we are not trying to catch you out.

NOTE TO INTERVIEWERS:
RESPONDENTS SHOULD ONLY BE ASKED EITHER Q15a OR Q15b, AND NOT BOTH QUESTIONS. PLEASE ROTATE Q15 AFTER EACH INTERVIEW THAT YOU COMPLETE SO THAT YOU HAVE EQUAL NUMBERS OF PEOPLE HAVING ANSWERED Q15a AND Q15b. FOR EXAMPLE, IN ONE SAMPLE POINT, FOUR PEOPLE SHOULD BE ASKED Q15a AND ANOTHER FOUR ASKED Q15b.

Q15.a SHOWCARD M **Imagine an illness – illness A - that gives severe health problems, and an illness - illness B – that gives moderate health problems. Treatment will help patients with illness A a little, while it will help patients with illness B considerably. The cost of treatment is the same in both cases. An increase in funding is available but we are unable to treat all of the patients in both patient groups. Which of the three different options on this card come closest to your own view? Just read out the number.**

❑ ROTATE Q15a AND Q15b. TICK IF Q15a IS ASKED ☑. SINGLE CODE ONLY

		(61)
1)	Most of the increase should be allocated to treatment for illness B	1
2)	Most of the increase should be allocated to treatment for illness A	2
3)	The increase should be divided evenly between the two groups	3
	Other (PLEASE WRITE IN AND CODE '4')	4
	Don't Know	5
	None	6

(61)

Q15.b SHOWCARD N **Imagine an illness – illness A – where the patient group is disadvantaged, for example, from a low income family, and an illness - illness B – where the patients are from a more advantaged group. Treatment will help patients with illness A a little, while it will help patients with illness B considerably. The cost of treatment is the same in both cases. An increase in funding is available but we are unable to treat all of the patients in both patient groups. Which of the three different options on this card come closest to your own view? Just read out the number.**

❑ ROTATE Q15a AND Q15b. TICK IF Q15b IS ASKED ☑. SINGLE CODE ONLY

(62)

1)	Most of the increase should be allocated to treatment for illness B	1
2)	Most of the increase should be allocated to treatment for illness A	2
3)	The increase should be divided evenly between the two groups	3
	Other (PLEASE WRITE IN AND CODE '4')	4
	Don't Know	5
	None	6

(62)

SHOWCARDS FOR EMPIRICAL STUDY 1:

QUESTIONNAIRE VERSION 1

SHOWCARD M

ILLNESS A	ILLNESS B
Severe health problems Treatment helps patients a little	Moderate health problems Treatment helps patients considerably
Cost of treatment is the same in both cases *Unable to treat all patients in both groups*	

1) Most of the increase should be allocated to treatment for illness B, involving moderate health problems which improve considerably with treatment

2) Most of the increase should be allocated to treatment for illness A, involving severe health problems which improve a little with treatment

3) The increase should be divided evenly between the two groups

SHOWCARD N

ILLNESS A	ILLNESS B
Disadvantaged patient group (e.g. from a low income family) Treatment helps patients a little	More advantaged patient group Treatment helps patients considerably
Both patient groups have similar health problems without treatment. *Cost of treatment is the same in both cases.* *Unable to treat all patients in both groups.*	

1) Most of the increase should be allocated to treatment for illness B, in a more advantaged patient group, which improves patients considerably with treatment

2) Most of the increase should be allocated to treatment for illness A, in a disadvantaged patient group, which improves patients a little with treatment

3) The increase should be divided evenly between the two groups

EMPIRICAL STUDY 1: MAIN QUESTIONNAIRE - VERSION 2

ASK ALL

We would now like you to consider a similar but slightly different type of question. It involves setting priorities, and again it involves a difficult choice. However, please note that there is no right or wrong answers and that we are not trying to catch you out.

NOTE TO INTERVIEWERS:
RESPONDENTS SHOULD ONLY BE ASKED EITHER Q15c OR Q15d, AND NOT BOTH QUESTIONS. PLEASE ROTATE Q15 AFTER EACH INTERVIEW THAT YOU COMPLETE SO THAT YOU HAVE EQUAL NUMBERS OF PEOPLE HAVING ANSWERED Q15c AND Q15d. FOR EXAMPLE, IN ONE SAMPLE POINT, FOUR PEOPLE SHOULD BE ASKED Q15c AND ANOTHER FOUR ASKED Q15d.

Q15.c SHOWCARD M **Imagine an illness – illness A - that gives severe health problems, and an illness - illness B – that gives moderate health problems. Treatment will help patients with illness A a little, while it will help patients with illness B considerably. The cost of treatment is the same in both cases. An increase in funding is available but we are unable to treat all of the patients in both patient groups. Which of the four different options on this card come closest to your own view? Just read out the number.**

❑ ROTATE Q15c AND Q15d. TICK IF Q15c IS ASKED ☑. SINGLE CODE ONLY

		(61)
1)	Most of the increase should be allocated to treatment for illness B	1
2)	Most of the increase should be allocated to treatment for illness A	2
3)	The increase should be divided evenly between the two groups	3
4)	I am not able to make a decision and would prefer that the choice be made by others	4
	Other (PLEASE WRITE IN AND CODE '5')	5
	Don't Know	6
	None	7

(61)

Q15.d SHOWCARD N **Imagine an illness – illness A – where the patient group is disadvantaged, for example, from a low income family, and an illness - illness B – where the patients are from a more advantaged group. Treatment will help patients with illness A a little, while it will help patients with illness B considerably. The cost of treatment is the same in both cases. An increase in funding is available but we are unable to treat all of the patients in both patient groups. Which of the four different options on this card come closest to your own view? Just read out the number.**

❑ ROTATE Q15c AND Q15d. TICK IF Q15d IS ASKED ☑. SINGLE CODE ONLY

(62)

1)	Most of the increase should be allocated to treatment for illness B	1
2)	Most of the increase should be allocated to treatment for illness A	2
3)	The increase should be divided evenly between the two groups	3
4)	I am not able to make a decision and would prefer that the choice be made by others	4
	Other (PLEASE WRITE IN AND CODE '5')	5
	Don't Know	6
	None	7

(62)

SHOWCARDS FOR EMPIRICAL STUDY 1:

QUESTIONNAIRE VERSION 2

SHOWCARD M

ILLNESS A	ILLNESS B
Severe health problems Treatment helps patients a little	Moderate health problems Treatment helps patients considerably
Cost of treatment is the same in both cases *Unable to treat all patients in both groups*	

1) Most of the increase should be allocated to treatment for illness B, involving moderate health problems which improve considerably with treatment

2) Most of the increase should be allocated to treatment for illness A, involving severe health problems which improve a little with treatment

3) The increase should be divided evenly between the two groups

4) I am not able to make a decision and would prefer that the choice be made by others

SHOWCARD N

ILLNESS A	ILLNESS B
Disadvantaged patient group (e.g. from a low income family) Treatment helps patients a little	More advantaged patient group Treatment helps patients considerably
Both patient groups have similar health problems without treatment. *Cost of treatment is the same in both cases.* *Unable to treat all patients in both groups.*	

1) Most of the increase should be allocated to treatment for illness B, in a more advantaged patient group, which improves patients considerably with treatment

2) Most of the increase should be allocated to treatment for illness A, in a disadvantaged patient group, which improves patients a little with treatment

3) The increase should be divided evenly between the two groups

4) I am not able to make a decision and would prefer that the choice be made by others

APPENDIX 4

DCE STUDY

MAIN QUESTIONNAIRE: VERSION 1

MORI/J25567
1-5

Questionnaire No: ☐☐

Serial No
OUO (6-9)
CARD 1 | 10

Priority Health Setting in Southampton
Main Stage - Questionnaire Version 1 (01/09/05)

Sample Point Number:
☐ ☐ ☐
(11) (12) (13)
Sample point name:

_____

Gender

Male	1	
Female	2	(14)

WRITE IN & CODE EXACT AGE
Exact Age ☐☐ (15) (16) (15-16)

18-24	1	
25-34	2	
35-44	3	
45-54	4	
55-59	5	
60-64	6	
65+	7	(17)

QA. SHOWCARD O **Working Status of Respondent** SINGLE CODE ONLY

Working - Full-time (30+ hrs/wk)	1	
Working - Part-time (8-29 hrs/wk)	2	
Working (under 8 hrs/wk)	3	
Houseperson	4	
Retired	5	
Registered unemployed	6	
Unemployed but not registered	7	
Permanently sick/disabled	8	
On a training scheme	9	
Voluntary work	0	
Student	X	
Other	Y	(18)

Occupation of Chief Income Earner

Position/rank/grade

Industry/type of company

Quals/degree/apprenticeship

Number of staff responsible for

REMEMBER TO PROBE FULLY FOR PENSION AND CODE FROM ABOVE

Class

A	1	
B	2	
C1	3	
C2	4	
D	5	
E	6	(19)

Respondent is:

Chief Income Earner	1	
Not Chief Income Earner	2	(20)

QB. SHOWCARD P **Home Ownership** SINGLE CODE ONLY

Owned outright	1	
Buying on mortgage (privately)	2	
Buying on a mortgage from Council	3	
Rented from Council	4	
Rented from Housing Association	5	
Rented from private landlord	6	
Other (WRITE IN & CODE '7')	7	(21)

Interviewer Declaration
I confirm that I have carried out this Interview face-to-face with the named person of the address attached and that I asked all the relevant questions fully and recorded the answers in conformance with the survey specification and within the MRS Code of Conduct and the Data Protection Act 1998.

Signature:.....

Interviewer Name (CAPS):.....

.....

Interviewer Number:
☐☐☐☐/☐☐
(22) (23) (24) (25) (26) (22-26)

Day of Interview 1 2 3 4 5 6 7
(Mon) (Thur) (Sun) (27)

Date of Interview: ☐☐/☐☐/05 (28-31)

Length of Interview: ☐☐ (minutes) (32-33)

QC. SHOWCARD Q **Household is:** SINGLE CODE ONLY

Single adult under 60	1	
Single adult 60 or over	2	
Two adults both under 60	3	
Two adults at least one 60 or over	4	
Three adults or more all 16 or over	5	
1-parent family with child/ren, at least one under 16	6	
2-parent family with child/ren at least one under 16	7	
Other (PLEASE WRITE IN & CODE '8')	8	
Not stated	9	(34)

QD. SHOWCARD R **In which would you place your total household income from all sources before tax and other deductions?** SINGLE CODE ONLY

Per Week	Per Year	(35)
Up to £86	Under £4,500	1
£87 - £125	£4,500-£6,499	2
£126 - £144	£6,500 - £7,499	3
£145 - £182	£7,500 - £9,499	4
£183 - £221	£9,500 - £11,499	5
£222 - £259	£11,500 - £13,499	6
£260 - £298	£13,500 - £15,499	7
£299 - £336	£15,500 - £17,499	8
£337 - £480	£17,500 - £24,999	9
£481 - £576	£25,000 - £29,999	0
£577 - £769	£30,000 - £39,999	X
£770 - £961	£40,000 - £49,999	Y
		(36)
£962 - £1,441	£50,000 - £74,999	1
£1,442-£1,922	£75,000 - £99,999	2
£1,923 +	£100,000+	3
	Refused	4 (35-36)

QE. SHOWCARD S **Which of the groups on this card do you consider you belong to?** SINGLE CODE ONLY

(37)

WHITE:	
British	1
Irish	2
Any other white background (PLEASE WRITE IN)	3
MIXED:	
White and Black Caribbean	4
White and Black African	5
White and Asian	6
Any other mixed background (PLEASE WRITE IN)	7
ASIAN OR ASIAN BRITISH:	
Indian	8
Pakistani	9
Bangladeshi	0
Sri Lankan	X
Any other Asian background (PLEASE WRITE IN)	Y

(38)

BLACK OR BLACK	
Caribbean	1
African	2
Any other black background (PLEASE WRITE IN)	3
ARAB OR MIDDLE EASTERN:	
Arab	4
Iranian	5
Any other Arabic/Middle eastern background (PLEASE WRITE IN)	6
CHINESE OR OTHER ETHNIC GROUP:	
Chinese	7
Any other background (PLEASE WRITE IN)	8
Refused	9 (37-38)

MORI

QF. SHOWCARD T **How is your health in general? Would you say it was...** SINGLE CODE ONLY

	(39)	
Very good	1	
Good	2	
Fair	3	
Bad	4	
Very bad	5	
Don't know	6	(39)

QG. **Do you or any members of this household have any longstanding illness, disability or infirmity? By long-standing we mean anything that has troubled you or a member of your household over a period of time, or that is likely to affect you/them over a period of time.** MULTICODE

	(40)	
Yes, respondent	1	
Yes, someone else in household	2	
No	3	(40)

ASK ALL WHO PERSONALLY HAVE OR SOMEONE ELSE IN THE HOUSEHOLD WHO HAVE A LONG-STANDING ILLNESS, DISABILITY OR INFIRMITY (CODES 1 OR 2 AT QG). OTHERS GO TO QI.

QH. **Does this illness, disability or infirmity limit [your and/or a member of your households] activities in any way?** MULTICODE

	(41)	
Yes, respondent	1	
Yes, someone else in household	2	
No	3	(41)

ASK ALL

QI. **Do you have private health insurance cover?** SINGLE CODE ONLY

	(42)	
Yes	1	
No	2	
Don't know	3	(42)

INTERVIEWER RECORD END TIME AFTER DEMOGRAPHICS ☐☐ : ☐☐ Hours Mins

INTERVIEWER RECORD START TIME ☐☐ : ☐☐ Hours Mins

INTRODUCTION/CONFIDENTIALITY

Good morning/afternoon/evening, I'm from MORI, the independent research organisation. We are doing a survey about some of the issues that are important in the decision-making process of the National Health Service (NHS), often referred to as priority-setting.

I would like to assure you that all the information we collect will be kept in the strictest confidence, and used for research purposes only. It will not be possible to identify any particular individual or address in the results. The interview may take about 25 minutes.... would you be willing to take part?

As I mentioned, this survey is about setting priorities in health. In the NHS decision-makers are asked to consider priority-setting issues and it often involves very difficult choices over the health care treatments and services that are to be provided within a limited budget. Decision-makers have to consider a range of health conditions such as heart disease, arthritis and mental illness. They also have to consider many different patient groups: some are severely affected by their health whilst some are not severely affected, some patients are able to get large health improvements while some are not able to get large health improvements.

This interview is hoping to gather views on how the public might wish to set priorities when faced with difficult choices. I would like you to put yourself in the position of a health care decision maker and to consider what it is like to be in their shoes when they have to make some of these difficult choices on how to best use its limited budget for the provision of health care services.

ASK ALL

Q1. SHOWCARD A **The four issues on this card represent some of the key issues decision makers have to consider when setting priorities. Please take a few minutes to read about these four issues and then I would like you to rank them in order of importance, from 1st (the most important) to 4th (the least important). You may think that some are equally important. Please note that there are no right and wrong answers here, I just want you to be familiar with the issues for the purposes of this interview.** READ OUT. SINGLE CODE FOR EACH

a) **Which do you think is the most important?**
b) **And which do you think is 2nd in order of importance?**
c) **And which do you think is 3rd in order of importance?**
d) **And which do you think is 4th in order of importance?**

INTERVIEWER NOTE – IT IS POSSIBLE TO HAVE MORE THAN ONE ISSUE RANKED EQUALLY. FOR EXAMPLE, IF WE HAD TWO EQUAL FIRST RANKINGS (AND A THIRD AND A FOURTH BUT NO SECOND), YOU WOULD CODE THIS AS 1, 1, 3, 4. PLEASE NOTE THAT YOU WOULD NOT CODE '2' BECAUSE IN THIS EXAMPLE 2ND MOST IMPORTANT HAS NOT BEEN MENTIONED.

		Severity of health condition	Improvement in health	Value for money	Other treatments available	Don't know	
a)	Most important	1	1	1	1	1	(43)
b)	2nd most important	2	2	2	2	2	(44)
c)	3rd most important	3	3	3	3	3	(45)
d)	4th most important	4	4	4	4	4	(46)

NOTE TO INTERVIEWER – PRESENT RESPONDENT WITH SHOWCARD B.

I am now going to ask you to consider some of the choices that a decision maker may have to face such as the ones on this card, where there are two treatment options and there is only funding to support one of them.

As you can see we have described the choices in a very general way, so that we can consider your views across a wide range of treatments and patient groups. However, to help you when considering the choices I would like you to think of patient groups with longer term health conditions, for example, where ill-health may affect patients' daily lives. They may have problems with mobility, carrying out their usual activities, or may have pain or anxiety, or a combination of these problems.

Q2. SHOWCARD B **A health care decision maker is faced with difficult choices on how to allocate its budget. Imagine a choice where there are two options for the use of available funds. Given that only one of the options can receive funding, which option would you support?** READ OUT. ROTATE. SINGLE CODE ONLY

Option A **Patients are not severely affected by their condition. With treatment for this condition the average patient has a large improvement in their health. The treatment is regarded as being fairly good value for money. There are no other effective treatment options available for this condition.**

Option B **Patients are severely affected by their condition. With treatment for this condition the average patient has a small improvement in their health. The treatment is regarded as very good value for money. There are other effective treatment options available for this condition.**

(47)

Option A	1
Option B	2
Other (PLEASE WRITE IN AND CODE '3')	3
Don't Know/Unable to choose	4
None	5
Refused	6

(47)

Q3.a SHOWCARD C **Again... a health care decision maker is faced with difficult choices on how to allocate its budget. Imagine a choice where there are two options for the use of available funds. Given that only one of the options can receive funding, which option would you support?** READ OUT. ROTATE. SINGLE CODE ONLY

Option C **Patients are severely affected by their condition. With treatment the average patient has a large improvement in their health. The treatment is regarded as very good value for money. There are no other effective treatment options available.**

Option D **Patients are not severely affected by their condition. With treatment the average patient has a small improvement in their health. The treatment is regarded as being very poor value for money. There are other effective treatment options available.**

(48)

Option C	1	GO TO Q4
Option D	2	ASK Q3.b
Other (PLEASE WRITE IN AND CODE '3')	3	
		GO TO Q4
Don't Know/Unable to choose	4	
None	5	
Refused	6	

(48)

ASK ALL WHO PICK OPTION D (CODE 2) AT Q3a. OTHERS PLEASE GO TO Q4.

Q3.b **Can you explain why you chose option D?**
PROBE FULLY AND WRITE IN.
ANY ANSWER (WRITE IN AND CODE '1') 1

None/no answer X

Don't know Y (49)

ASK ALL

Q4. SHOWCARD D **If a health decision-maker was given these two choices, which option would you support?** READ OUT. ROTATE. SINGLE CODE ONLY

Option E **Patients are not severely affected by their condition. With treatment the average patient has a large improvement in their health. The treatment is regarded as being fairly poor value for money. There are no other effective treatment options.**

Option F **Patients are severely affected by their condition. With treatment the average patient has a moderate improvement in their health. The treatment is regarded as being very poor value for money. There are other effective treatment options available.**

(50)

Option E	1
Option F	2
Other (PLEASE WRITE IN AND CODE '3')	3
Don't Know/Unable to choose	4
None	5
Refused	6

(50)

Q5. SHOWCARD E **If a health decision-maker was given these two choices, which option would you support?** READ OUT. ROTATE. SINGLE CODE ONLY

Option G **Patients are not severely affected by their condition. With treatment the average patient has a moderate improvement in their health. The treatment is regarded as being fairly poor value for money. There are other effective treatment options available.**

Option H **Patients are severely affected by their condition. With treatment the average patient has a small improvement in their health. The treatment is regarded as being very poor value for money. There are no other effective treatment options available.**

(51)

Option G	1
Option H	2
Other (PLEASE WRITE IN AND CODE '3')	3
Don't Know/Unable to choose	4
None	5
Refused	6

(51)

Q6. SHOWCARD F **If a health decision-maker was given these two choices, which option would you support?** READ OUT. ROTATE. SINGLE CODE ONLY

Option I **Patients are severely affected by their condition. With treatment the average patient has a large improvement in their health. The treatment is regarded as being very good value for money. There are no other effective treatment options available.**

Option J **Patients are not severely affected by their condition. With treatment the average patient has a moderate improvement in their health. The treatment is regarded as being fairly good value for money. There are other effective treatment options available.**

(52)

Option I	1
Option J	2
Other (PLEASE WRITE IN AND CODE '3')	3
Don't Know/Unable to choose	4
None	5
Refused	6

(52)

Q7. SHOWCARD G **If a health decision-maker was given these two choices, which option would you support?** READ OUT. ROTATE. SINGLE CODE ONLY

Option K **Patients are severely affected by their condition. With treatment the average patient has a very small improvement in their health. The treatment is regarded as being very poor value for money. There are no other effective treatment options available.**

Option L **Patients are not severely affected by their condition. With treatment the average patient has a large improvement in their health. The treatment is regarded as being very good value for money. There are other effective treatment options available.**

(53)

Option K	1
Option L	2
Other (PLEASE WRITE IN AND CODE '3')	3
Don't Know/Unable to choose	4
None	5
Refused	6

(53)

Q8. SHOWCARD H **If a health decision-maker was given these two choices, which option would you support?**

Is it helpful for me to continue reading out the options on the showcard?
IF YES, PLEASE CONTINUE TO READ OUT (AND TO ROTATE). IF NO, PRESENT THE SHOWCARD AND LET THE RESPONDENT READ IT THEMSELVES. SINGLE CODE ONLY.

Option M **Patients are not severely affected by their condition. With treatment the average patient has a small improvement in their health. The treatment is regarded as being fairly good value for money. There are other effective treatment options available.**

Option N **Patients are severely affected by their condition. With treatment the average patient has a very small improvement in their health. The treatment is regarded as being fairly poor value for money. There are no other effective treatment options.**

(54)

Option M	1
Option N	2
Other (PLEASE WRITE IN AND CODE '3')	3
Don't Know/Unable to choose	4
None	5
Refused	6

(54)

Q9. SHOWCARD I **If a health decision-maker was given these two choices, which option would you support?** READ OUT. ROTATE. SINGLE CODE ONLY

Option O **Patients are severely affected by their condition. With treatment the average patient has a moderate improvement in their health. The treatment is regarded as being very good value for money. There are other effective treatment options available.**

Option P **Patients are not severely affected by their condition. With treatment the average patient has a small improvement in their health. The treatment is regarded as being fairly good value for money. There are no other effective treatment options available.**

(55)

Option O	1
Option P	2
Other (PLEASE WRITE IN AND CODE '3')	3
Don't Know/Unable to choose	4
None	5
Refused	6

(55)

Q10. SHOWCARD J **If a health decision-maker was given these two choices, which option would you support?** READ OUT. ROTATE. SINGLE CODE ONLY

Option Q **Patients are not severely affected by their condition. With treatment the average patient has a very small improvement in their health. The treatment is regarded as being fairly good value for money. There are no other effective treatment options available.**

Option R **Patients are severely affected by their condition. With treatment the average patient has a large improvement in their health. The treatment is regarded as being fairly poor value for money. There are other effective treatment options available.**

(56)

Option Q	1
Option R	2
Other (PLEASE WRITE IN AND CODE '3')	3
Don't Know/Unable to choose	4
None	5
Refused	6

(56)

Q11. SHOWCARD K **If a health decision-maker was given these two choices, which option would you support?** READ OUT. ROTATE. SINGLE CODE ONLY

Option S **Patients are severely affected by their condition. With treatment the average patient has a small improvement in their health. The treatment is regarded as being very poor value for money. There are other effective treatment options available.**

Option T **Patients are not severely affected by their condition. With treatment the average patient has a very small improvement in their health. The treatment is regarded as being very good value for money. There are no other effective treatment options available.**

(57)

Option S	1
Option T	2
Other (PLEASE WRITE IN AND CODE '3')	3
Don't Know/Unable to choose	4
None	5
Refused	6

(57)

Q12. SHOWCARD L **I would now like to talk to you a little about your answers. Overall, how difficult did you find it <u>understanding</u> these questions?** SINGLE CODE ONLY

	(58)
Very difficult	1
Fairly difficult	2
Not very difficult	3
Not at all difficult	4
Don't know	5

(58)

Q13. SHOWCARD L AGAIN **And overall, how difficult did you find it <u>answering</u> these questions (involving the choices)?** SINGLE CODE ONLY

	(59)	
Very difficult	1	ASK Q14
Fairly difficult	2	
Not very difficult	3	
Not at all difficult	4	GO TO Q15a OR Q15b
Don't know	5	

(59)

ASK ALL WHO SAY "VERY DIFFICULT" OR "FAIRLY DIFFICULT" AT Q13 (CODES 1 AND 2). OTHERS GO TO Q15A OR Q15B.

Q14. **And why do you say that?** PROBE FULLY AND WRITE IN.

ANY ANSWER (WRITE IN AND CODE '1')	1
None/no answer	X
Don't know	Y

(60)

ASK ALL

We would now like you to consider a similar but slightly different type of question. It involves setting priorities, and again it involves a difficult choice. However, please note that there is no right or wrong answers and that we are not trying to catch you out.

<u>NOTE TO INTERVIEWERS</u>:
RESPONDENTS SHOULD ONLY BE ASKED <u>EITHER</u> Q15a OR Q15b, AND <u>NOT</u> BOTH QUESTIONS. PLEASE ROTATE Q15 AFTER EACH INTERVIEW THAT YOU COMPLETE SO THAT YOU HAVE EQUAL NUMBERS OF PEOPLE HAVING ANSWERED Q15a AND Q15b. FOR EXAMPLE, IN ONE SAMPLE POINT, FOUR PEOPLE SHOULD BE ASKED Q15a AND ANOTHER FOUR ASKED Q15b.

Q15.a SHOWCARD M **Imagine an illness – illness A - that gives severe health problems, and an illness - illness B – that gives moderate health problems. Treatment will help patients with illness A a little, while it will help patients with illness B considerably. The cost of treatment is the same in both cases. An increase in funding is available but we are unable to treat all of the patients in both patient groups. Which of the three different options on this card come closest to your own view? Just read out the number.**

❑ ROTATE Q15a AND Q15b. TICK IF Q15a IS ASKED ☑. SINGLE CODE ONLY

		(61)
1)	Most of the increase should be allocated to treatment for illness B	1
2)	Most of the increase should be allocated to treatment for illness A	2
3)	The increase should be divided evenly between the two groups	3
	Other (PLEASE WRITE IN AND CODE '4')	4
	Don't Know	5
	None	6

(61)

Q15.b SHOWCARD N **Imagine an illness – illness A – where the patient group is disadvantaged, for example, from a low income family, and an illness - illness B – where the patients are from a more advantaged group. Treatment will help patients with illness A a little, while it will help patients with illness B considerably. The cost of treatment is the same in both cases. An increase in funding is available but we are unable to treat all of the patients in both patient groups. Which of the three different options on this card come closest to your own view? Just read out the number.**

❑ ROTATE Q15a AND Q15b. TICK IF Q15b IS ASKED ☑. SINGLE CODE ONLY

		(62)
1)	Most of the increase should be allocated to treatment for illness B	1
2)	Most of the increase should be allocated to treatment for illness A	2
3)	The increase should be divided evenly between the two groups	3
	Other (PLEASE WRITE IN AND CODE '4')	4
	Don't Know	5
	None	6

(62)

Q16. **Finally, do you wish to add anything else on this subject?**
PROBE FULLY AND WRITE IN.

ANY ANSWER (WRITE IN AND CODE '1')	1
None/no answer	X
Don't know	Y (63)

MORI

THIS MUST BE THE LAST PAGE OF THE QUESTIONNAIRE AND MUST BE SINGLE SIDED

MORI/J25567 1-5 Questionnaire No ☐☐ Serial No OUO (6-9) CARD * 10

Priority Health Setting in Southampton Survey
Main Stage - Questionnaire Version 1 (01/09/05)

Sample Point Number: ☐☐☐ (11) (12) (13)

Sample point name:

Interviewer Number: ☐☐☐☐/☐ (14) (15) (16) (17) (18) (14-18)

Interviewer Name:.....

Name/Initial/Title: Mr/Mrs/Ms/Miss

Address:

Full Postcode ☐☐☐☐☐☐☐ (19) (20) (21) (22) (23) (24) (25) (19-25)

QTEL1 **Do you have a fixed line telephone at home which you use for incoming and outgoing voice calls?**

(26)

Yes	1
WRITE IN Full tel. No	
No	2
Refused	3 GO TO QTEL2
Ex-directory	4 (26)

QTEL2 ASK IF NO FIXED LINE/REFUSED/EX-DIRECTORY (CODES 2-4). OTHERS CLOSE
Can I just check, do you have a mobile phone? IF YES ASK: **Can I take the number please?**

(27)

Yes	1
WRITE IN Full tel. No	
No	2
Refused	3 (27)

IT IS VERY IMPORTANT THAT YOU FILL IN THE SAMPLE POINT NUMBER, QUESTIONNAIRE NUMBER AND INTERVIEWER NUMBER ON BOTH THE FRONT AND BACK PAGE OF THE QUESTIONNAIRE.

MORI

SHOWCARDS FOR DCE:

QUESTIONNAIRE VERSION 1

SHOWCARD A

SEVERITY OF THE HEALTH CONDITION

Whether patients are severely affected by their condition

When considering severity we have judged that on at least one of the following areas patients have severe problems: (i) self-care *(e.g. unable to wash or dress themselves)*, (ii) unable to perform usual activities *(e.g. work, study, housework, family or leisure activities)*, (iii) extreme pain or discomfort, (iv) extreme anxiety or depression.

IMPROVEMENT IN HEALTH

Considering the average health improvement from treatment

Improvement in health refers to the benefits that the patient feels following treatment e.g. improvements in their mobility, improvements in their ability to perform usual activities, reduced pain, reduced anxiety.

In this questionnaire treatments offer one of the following:

- large improvement in health
- moderate improvement in health
- small improvement in health
- very small improvement in health

VALUE FOR MONEY

Whether or not the treatment offers 'value for money'

'Value for money' is a common consideration within the NHS. Value for money is about the efficient use of resources (e.g. doctor's time, hospital beds, healthcare funds).

In this questionnaire we use the following categories:

- very good value for money
- fairly good value for money
- fairly poor value for money
- very poor value for money

WHETHER OTHER EFFECTIVE TREATMENTS ARE AVAILABLE

Yes available or No not available

We assume that all patients will have usual and best supportive care available within the NHS (e.g. GP services, specialist outpatient appointments, and best supportive nursing care). Where referring to treatment options we are thinking about drugs, surgery, or specially developed services (such as physiotherapy, support services, occupational therapy, specialist education, preventative therapies, etc).

SHOWCARD B

Health condition: Long-term health condition

	OPTION A	OPTION B
Severity of patients	Not severely affected	Severely affected
Improvement in health	Large improvement in health	Small improvement in health
Does the treatment offer Value for money	Fairly good value for money	Very good value for money
Other effective treatment options available	No	Yes

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

MORI\J25567 – For use with questionnaire version 1

SHOWCARD C

Health condition: Long-term health condition

	OPTION C	OPTION D
Severity of patients	Severely affected	Not severely affected
Improvement in health	Large improvement in health	Small improvement in health
Does the treatment offer Value for money	Very good value for money	Very poor value for money
Other effective treatment options available	No	Yes

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

SHOWCARD D

Health condition: Long-term health condition

	OPTION E	OPTION F
Severity of patients	Not severely affected	Severely affected
Improvement in health	Large improvement in health	Moderate improvement in health
Does the treatment offer Value for money	Fairly poor value for money	Very poor value for money
Other effective treatment options available	No	Yes

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

SHOWCARD E

Health condition: Long-term health condition

	OPTION G	OPTION H
Severity of patients	Not severely affected	Severely affected
Improvement in health	Moderate improvement in health	Small improvement in health
Does the treatment offer Value for money	Fairly poor value for money	Very poor value for money
Other effective treatment options available	Yes	No

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

SHOWCARD F

Health condition: Long-term health condition

	OPTION I	OPTION J
Severity of patients	Severely affected	Not severely affected
Improvement in health	Large improvement in health	Moderate improvement in health
Does the treatment offer Value for money	Very good value for money	Fairly good value for money
Other effective treatment options available	No	Yes

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

SHOWCARD G

Health condition: Long-term health condition

	OPTION K	OPTION L
Severity of patients	Severely affected	Not severely affected
Improvement in health	Very small improvement in health	Large improvement in health
Does the treatment offer Value for money	Very poor value for money	Very good value for money
Other effective treatment options available	No	Yes

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

SHOWCARD H

Health condition: Long-term health condition

	OPTION M	OPTION N
Severity of patients	Not severely affected	Severely affected
Improvement in health	Small improvement in health	Very small improvement in health
Does the treatment offer Value for money	Fairly good value for money	Fairly poor value for money
Other effective treatment options available	Yes	No

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

SHOWCARD I

Health condition: Long-term health condition

	OPTION O	OPTION P
Severity of patients	Severely affected	Not severely affected
Improvement in health	Moderate improvement in health	Small improvement in health
Does the treatment offer Value for money	Very good value for money	Fairly good value for money
Other effective treatment options available	Yes	No

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

SHOWCARD J

Health condition: Long-term health condition

	OPTION Q	OPTION R
Severity of patients	Not severely affected	Severely affected
Improvement in health	Very small improvement in health	Large improvement in health
Does the treatment offer Value for money	Fairly good value for money	Fairly poor value for money
Other effective treatment options available	No	Yes

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

SHOWCARD K

Health condition: Long-term health condition

	OPTION S	OPTION T
Severity of patients	Severely affected	Not severely affected
Improvement in health	Small improvement in health	Very small improvement in health
Does the treatment offer Value for money	Very poor value for money	Very good value for money
Other effective treatment options available	Yes	No

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

DCE STUDY

MAIN QUESTIONNAIRE: VERSION 2

MORI/J25567 Questionnaire No: ☐☐ Serial No
1-5 OUO (6-9)
CARD 1 10

Priority Health Setting in Southampton
Main Stage - Questionnaire Version 2 (01/09/05)

Sample Point Number:
☐☐☐
(11) (12) (13)
Sample point name:

Gender

Male	1	
Female	2	(14)

WRITE IN & CODE EXACT AGE
Exact Age ☐☐ (15) (16) (15-16)

18-24	1	
25-34	2	
35-44	3	
45-54	4	
55-59	5	
60-64	6	
65+	7	(17)

QA. SHOWCARD O **Working Status of Respondent** SINGLE CODE ONLY

Working - Full-time (30+ hrs/wk)	1	
Working - Part-time (8-29 hrs/wk)	2	
Working (under 8 hrs/wk)	3	
Houseperson	4	
Retired	5	
Registered unemployed	6	
Unemployed but not registered	7	
Permanently sick/disabled	8	
On a training scheme	9	
Voluntary work	0	
Student	X	
Other	Y	(18)

Occupation of Chief Income Earner

Position/rank/grade

Industry/type of company

Quals/degree/apprenticeship

Number of staff responsible for

REMEMBER TO PROBE FULLY FOR PENSION AND CODE FROM ABOVE

Class

A	1	
B	2	
C1	3	
C2	4	
D	5	
E	6	(19)

Respondent is:

Chief Income Earner	1	
Not Chief Income Earner	2	(20)

QB. SHOWCARD P **Home Ownership** SINGLE CODE ONLY

Owned outright	1	
Buying on mortgage (privately)	2	
Buying on a mortgage from Council	3	
Rented from Council	4	
Rented from Housing Association	5	
Rented from private landlord	6	
Other (WRITE IN & CODE '7')	7	(21)

Interviewer Declaration

I confirm that I have carried out this interview face-to-face with the named person of the address attached and that I asked all the relevant questions fully and recorded the answers in conformance with the survey specification and within the MRS Code of Conduct and the Data Protection Act 1998.

Signature:.....

Interviewer Name (CAPS):.....

.....

Interviewer Number:
☐☐☐☐☐/☐
(22) (23) (24) (25) (26) (22-26)

Day of Interview 1 2 3 4 5 6 7
(Mon) (Thur) (Sun) (27)

Date of Interview: ☐☐/☐☐/05 (28-31)

Length of Interview: ☐☐ (minutes) (32-33)

QC. SHOWCARD Q **Household is:** SINGLE CODE ONLY

Single adult under 60	1	
Single adult 60 or over	2	
Two adults both under 60	3	
Two adults at least one 60 or over	4	
Three adults or more all 16 or over	5	
1-parent family with child/ren, at least one under 16	6	
2-parent family with child/ren at least one under 16	7	
Other (PLEASE WRITE IN & CODE '8')	8	
Not stated	9	(34)

QD. SHOWCARD R **In which would you place your total household income from all sources before tax and other deductions?** SINGLE CODE ONLY

Per Week	Per Year	(35)
Up to £86	Under £4,500	1
£87 - £125	£4,500-£6,499	2
£126 - £144	£6,500 - £7,499	3
£145 - £182	£7,500 - £9,499	4
£183 - £221	£9,500 - £11,499	5
£222 - £259	£11,500 - £13,499	6
£260 - £298	£13,500 - £15,499	7
£299 - £336	£15,500 - £17,499	8
£337 - £480	£17,500 - £24,999	9
£481 - £576	£25,000 - £29,999	0
£577 - £769	£30,000 - £39,999	X
£770 - £961	£40,000 - £49,999	Y
		(36)
£962 - £1,441	£50,000 - £74,999	1
£1,442-£1,922	£75,000 - £99,999	2
£1,923 +	£100,000+	3
	Refused	4 (35-36)

QE. SHOWCARD S **Which of the groups on this card do you consider you belong to?** SINGLE CODE ONLY

		(37)
WHITE:		
British	1	
Irish	2	
Any other white background (PLEASE WRITE IN)	3	
MIXED:		
White and Black Caribbean	4	
White and Black African	5	
White and Asian	6	
Any other mixed background (PLEASE WRITE IN)	7	
ASIAN OR ASIAN BRITISH:		
Indian	8	
Pakistani	9	
Bangladeshi	0	
Sri Lankan	X	
Any other Asian background (PLEASE WRITE IN)	Y	
		(38)
BLACK OR BLACK		
Caribbean	1	
African	2	
Any other black background (PLEASE WRITE IN)	3	
ARAB OR MIDDLE EASTERN:		
Arab	4	
Iranian	5	
Any other Arabic/Middle eastern background (PLEASE WRITE IN)	6	
CHINESE OR OTHER ETHNIC GROUP:		
Chinese	7	
Any other background (PLEASE WRITE IN)	8	
Refused	9	(37-38)

MORI

QF. SHOWCARD T **How is your health in general? Would you say it was...** SINGLE CODE ONLY

	(39)	
Very good	1	
Good	2	
Fair	3	
Bad	4	
Very bad	5	
Don't know	6	(39)

QG. **Do you or any members of this household have any longstanding illness, disability or infirmity? By long-standing we mean anything that has troubled you or a member of your household over a period of time, or that is likely to affect you/them over a period of time.** MULTICODE

	(40)	
Yes, respondent	1	
Yes, someone else in household	2	
No	3	(40)

ASK ALL WHO PERSONALLY HAVE OR SOMEONE ELSE IN THE HOUSEHOLD WHO HAVE A LONG-STANDING ILLNESS, DISABILITY OR INFIRMITY (CODES 1 OR 2 AT QG). OTHERS GO TO QI.

QH. **Does this illness, disability or infirmity limit [your and/or a member of your households] activities in any way?** MULTICODE

	(41)	
Yes, respondent	1	
Yes, someone else in household	2	
No	3	(41)

ASK ALL

QI. **Do you have private health insurance cover?** SINGLE CODE ONLY

	(42)	
Yes	1	
No	2	
Don't know	3	(42)

INTERVIEWER RECORD END TIME AFTER DEMOGRAPHICS ☐☐ : ☐☐ Hours Mins

INTERVIEWER RECORD START TIME ☐☐ : ☐☐ Hours Mins

INTRODUCTION/CONFIDENTIALITY

Good morning/afternoon/evening, I'm from MORI, the independent research organisation. We are doing a survey about some of the issues that are important in the decision-making process of the National Health Service (NHS), often referred to as priority-setting.

I would like to assure you that all the information we collect will be kept in the strictest confidence, and used for research purposes only. It will not be possible to identify any particular individual or address in the results. The interview may take about 25 minutes.... would you be willing to take part?

As I mentioned, this survey is about setting priorities in health. In the NHS decision-makers are asked to consider priority-setting issues and it often involves very difficult choices over the health care treatments and services that are to be provided within a limited budget. Decision-makers have to consider a range of health conditions such as heart disease, arthritis and mental illness. They also have to consider many different patient groups: some are severely affected by their health whilst some are not severely affected, some patients are able to get large health improvements while some are not able to get large health improvements.

This interview is hoping to gather views on how the public might wish to set priorities when faced with difficult choices. I would like you to put yourself in the position of a health care decision maker and to consider what it is like to be in their shoes when they have to make some of these difficult choices on how to best use its limited budget for the provision of health care services.

ASK ALL

Q1. SHOWCARD A **The four issues on this card represent some of the key issues decision makers have to consider when setting priorities. Please take a few minutes to read about these four issues and then I would like you to rank them in order of importance, from 1st (the most important) to 4th (the least important). You may think that some are equally important. Please note that there are no right and wrong answers here, I just want you to be familiar with the issues for the purposes of this interview.** READ OUT. SINGLE CODE FOR EACH

a) **Which do you think is the most important?**
b) **And which do you think is 2nd in order of importance?**
c) **And which do you think is 3rd in order of importance?**
d) **And which do you think is 4th in order of importance?**

INTERVIEWER NOTE – IT IS POSSIBLE TO HAVE MORE THAN ONE ISSUE RANKED EQUALLY. FOR EXAMPLE, IF WE HAD TWO EQUAL FIRST RANKINGS (AND A THIRD AND A FOURTH BUT NO SECOND), YOU WOULD CODE THIS AS 1, 1, 3, 4. PLEASE NOTE THAT YOU WOULD NOT CODE '2' BECAUSE IN THIS EXAMPLE 2ND MOST IMPORTANT HAS NOT BEEN MENTIONED.

		Severity of health condition	Improvement in health	Value for money	Other treatments available	Don't know	
a)	Most important	1	1	1	1	1	(43)
b)	2nd most important	2	2	2	2	2	(44)
c)	3rd most important	3	3	3	3	3	(45)
d)	4th most important	4	4	4	4	4	(46)

NOTE TO INTERVIEWER – PRESENT RESPONDENT WITH SHOWCARD B.
I am now going to ask you to consider some of the choices that a decision maker may have to face such as the ones on this card, where there are two treatment options and there is only funding to support one of them.

As you can see we have described the choices in a very general way, so that we can consider your views across a wide range of treatments and patient groups. However, to help you when considering the choices I would like you to think of patient groups with longer term health conditions, for example, where ill-health may affect patients' daily lives. They may have problems with mobility, carrying out their usual activities, or may have pain or anxiety, or a combination of these problems.

Q2. SHOWCARD B **A health care decision maker is faced with difficult choices on how to allocate its budget. Imagine a choice where there are two options for the use of available funds. Given that only one of the options can receive funding, which option would you support?** READ OUT. ROTATE. SINGLE CODE ONLY

Option A **Patients are not severely affected by their condition. With treatment for this condition the average patient has a large improvement in their health. The treatment is regarded as being fairly good value for money. There are no other effective treatment options available for this condition.**

Option B **Patients are severely affected by their condition. With treatment for this condition the average patient has a small improvement in their health. The treatment is regarded as very good value for money. There are other effective treatment options available for this condition.**

(47)

Option A	1
Option B	2
Other (PLEASE WRITE IN AND CODE '3')	3
Don't Know/Unable to choose	4
None	5
Refused	6

(47)

Q3.a SHOWCARD C **Again... a health care decision maker is faced with difficult choices on how to allocate its budget. Imagine a choice where there are two options for the use of available funds. Given that only one of the options can receive funding, which option would you support?** READ OUT. ROTATE. SINGLE CODE ONLY

Option C **Patients are severely affected by their condition. With treatment the average patient has a large improvement in their health. The treatment is regarded as very good value for money. There are no other effective treatment options available.**

Option D **Patients are not severely affected by their condition. With treatment the average patient has a small improvement in their health. The treatment is regarded as being very poor value for money. There are other effective treatment options available.**

(48)

Option C	1	GO TO Q4
Option D	2	ASK Q3.b
Other (PLEASE WRITE IN AND CODE '3')	3	GO TO Q4
Don't Know/Unable to choose	4	
None	5	
Refused	6	

(48)

ASK ALL WHO PICK OPTION D (CODE 2) AT Q3a. OTHERS PLEASE GO TO Q4.

Q3.b **Can you explain why you chose option D?**
PROBE FULLY AND WRITE IN.
ANY ANSWER (WRITE IN AND CODE '1') 1

None/no answer X

Don't know Y (49)

ASK ALL

Q4. SHOWCARD D **If a health decision-maker was given these two choices, which option would you support?** READ OUT. ROTATE. SINGLE CODE ONLY

Option E **Patients are not severely affected by their condition. With treatment the average patient has a moderate improvement in their health. The treatment is regarded as being very poor value for money. There are no other effective treatment options.**

Option F **Patients are severely affected by their condition. With treatment the average patient has a small improvement in their health. The treatment is regarded as being very good value for money. There are other effective treatment options available.**

	(50)
Option E	1
Option F	2
Other (PLEASE WRITE IN AND CODE '3')	3
Don't Know/Unable to choose	4
None	5
Refused	6

(50)

Q5. SHOWCARD E **If a health decision-maker was given these two choices, which option would you support?** READ OUT. ROTATE. SINGLE CODE ONLY

Option G **Patients are severely affected by their condition. With treatment the average patient has a very small improvement in their health. The treatment is regarded as being fairly poor value for money. There are other effective treatment options available.**

Option H **Patients are not severely affected by their condition. With treatment the average patient has a large improvement in their health. The treatment is regarded as being very poor value for money. There are no other effective treatment options available.**

	(51)
Option G	1
Option H	2
Other (PLEASE WRITE IN AND CODE '3')	3
Don't Know/Unable to choose	4
None	5
Refused	6

(51)

Q6. SHOWCARD F **If a health decision-maker was given these two choices, which option would you support?** READ OUT. ROTATE. SINGLE CODE ONLY

Option I **Patients are not severely affected by their condition. With treatment the average patient has a small improvement in their health. The treatment is regarded as being very good value for money. There are no other effective treatment options available.**

Option J **Patients are severely affected by their condition. With treatment the average patient has a very small improvement in their health. The treatment is regarded as being fairly good value for money. There are other effective treatment options available.**

	(52)
Option I	1
Option J	2
Other (PLEASE WRITE IN AND CODE '3')	3
Don't Know/Unable to choose	4
None	5
Refused	6

(52)

Q7. SHOWCARD G **If a health decision-maker was given these two choices, which option would you support?** READ OUT. ROTATE. SINGLE CODE ONLY

Option K **Patients are not severely affected by their condition. With treatment the average patient has a large improvement in their health. The treatment is regarded as being very poor value for money. There are other effective treatment options available.**

Option L **Patients are severely affected by their condition. With treatment the average patient has a moderate improvement in their health. The treatment is regarded as being very good value for money. There are no other effective treatment options available.**

	(53)
Option K	1
Option L	2
Other (PLEASE WRITE IN AND CODE '3')	3
Don't Know/Unable to choose	4
None	5
Refused	6

(53)

Q8. SHOWCARD H **If a health decision-maker was given these two choices, which option would you support?**

Is it helpful for me to continue reading out the options on the showcard? IF YES, PLEASE CONTINUE TO READ OUT (AND TO ROTATE). IF NO, PRESENT THE SHOWCARD AND LET THE RESPONDENT READ IT THEMSELVES. SINGLE CODE ONLY.

Option M **Patients are severely affected by their condition. With treatment the average patient has a small improvement in their health. The treatment is regarded as being fairly poor value for money. There are no other effective treatment options available.**

Option N **Patients are not severely affected by their condition. With treatment the average patient has a very small improvement in their health. The treatment is regarded as being very poor value for money. There are other effective treatment options available.**

	(54)
Option M	1
Option N	2
Other (PLEASE WRITE IN AND CODE '3')	3
Don't Know/Unable to choose	4
None	5
Refused	6

(54)

Q9. SHOWCARD I **If a health decision-maker was given these two choices, which option would you support?** READ OUT. ROTATE. SINGLE CODE ONLY

Option O **Patients are severely affected by their condition. With treatment the average patient has a large improvement in their health. The treatment is regarded as being fairly good value for money. There are other effective treatment options available.**

Option P **Patients are not severely affected by their condition. With treatment the average patient has a moderate improvement in their health. The treatment is regarded as being fairly poor value for money. There are no other effective treatment options available.**

	(55)
Option O	1
Option P	2
Other (PLEASE WRITE IN AND CODE '3')	3
Don't Know/Unable to choose	4
None	5
Refused	6

(55)

Q10. SHOWCARD J **If a health decision-maker was given these two choices, which option would you support?** READ OUT. ROTATE. SINGLE CODE ONLY

Option Q **Patients are not severely affected by their condition. With treatment the average patient has a very small improvement in their health. The treatment is regarded as being very good value for money. There are other effective treatment options available.**

Option R **Patients are severely affected by their condition. With treatment the average patient has a large improvement in their health. The treatment is regarded as being fairly good value for money. There are no other effective treatment options available.**

	(56)
Option Q	1
Option R	2
Other (PLEASE WRITE IN AND CODE '3')	3
Don't Know/Unable to choose	4
None	5
Refused	6

(56)

Q11. SHOWCARD K **If a health decision-maker was given these two choices, which option would you support?** READ OUT. ROTATE. SINGLE CODE ONLY

Option S **Patients are severely affected by their condition. With treatment the average patient has a moderate improvement in their health. The treatment is regarded as being fairly good value for money. There are no other effective treatment options available.**

Option T **Patients are not severely affected by their condition. With treatment the average patient has a small improvement in their health. The treatment is regarded as being fairly poor value for money. There are other effective treatment options available.**

	(57)
Option S	1
Option T	2
Other (PLEASE WRITE IN AND CODE '3')	3
Don't Know/Unable to choose	4
None	5
Refused	6

(57)

Q12. SHOWCARD L **I would now like to talk to you a little about your answers. Overall, how difficult did you find it <u>understanding</u> these questions?** SINGLE CODE ONLY

	(58)
Very difficult	1
Fairly difficult	2
Not very difficult	3
Not at all difficult	4
Don't know	5

(58)

Q13. SHOWCARD L AGAIN **And overall, how difficult did you find it <u>answering</u> these questions (involving the choices)?** SINGLE CODE ONLY

	(59)	
Very difficult	1	ASK Q14
Fairly difficult	2	
Not very difficult	3	
Not at all difficult	4	GO TO Q15c OR Q15d
Don't know	5	

(59)

ASK ALL WHO SAY "VERY DIFFICULT" OR "FAIRLY DIFFICULT" AT Q13 (CODES 1 AND 2). OTHERS GO TO Q15c OR Q15d.

Q14. **And why do you say that?** PROBE FULLY AND WRITE IN.

ANY ANSWER (WRITE IN AND CODE '1')	1
None/no answer	X
Don't know	Y

(60)

MORI

ASK ALL

We would now like you to consider a similar but slightly different type of question. It involves setting priorities, and again it involves a difficult choice. However, please note that there is no right or wrong answers and that we are not trying to catch you out.

<u>NOTE TO INTERVIEWERS</u>:
RESPONDENTS SHOULD ONLY BE ASKED <u>EITHER</u> Q15c OR Q15d, AND <u>NOT</u> BOTH QUESTIONS. PLEASE ROTATE Q15 AFTER EACH INTERVIEW THAT YOU COMPLETE SO THAT YOU HAVE EQUAL NUMBERS OF PEOPLE HAVING ANSWERED Q15c AND Q15d. FOR EXAMPLE, IN ONE SAMPLE POINT, FOUR PEOPLE SHOULD BE ASKED Q15c AND ANOTHER FOUR ASKED Q15d.

Q15.c SHOWCARD M **Imagine an illness – illness A - that gives severe health problems, and an illness - illness B – that gives moderate health problems. Treatment will help patients with illness A a little, while it will help patients with illness B considerably. The cost of treatment is the same in both cases. An increase in funding is available but we are unable to treat all of the patients in both patient groups. Which of the four different options on this card come closest to your own view? Just read out the number.**

❑ ROTATE Q15c AND Q15d. TICK IF Q15c IS ASKED ☑. SINGLE CODE ONLY

		(61)
1)	Most of the increase should be allocated to treatment for illness B	1
2)	Most of the increase should be allocated to treatment for illness A	2
3)	The increase should be divided evenly between the two groups	3
4)	I am not able to make a decision and would prefer that the choice be made by others	4
	Other (PLEASE WRITE IN AND CODE '5')	5
	Don't Know	6
	None	7

(61)

Q15.d SHOWCARD N **Imagine an illness – illness A – where the patient group is disadvantaged, for example, from a low income family, and an illness - illness B – where the patients are from a more advantaged group. Treatment will help patients with illness A a little, while it will help patients with illness B considerably. The cost of treatment is the same in both cases. An increase in funding is available but we are unable to treat all of the patients in both patient groups. Which of the four different options on this card come closest to your own view? Just read out the number.**

❑ ROTATE Q15c AND Q15d. TICK IF Q15d IS ASKED ☑. SINGLE CODE ONLY

		(62)
1)	Most of the increase should be allocated to treatment for illness B	1
2)	Most of the increase should be allocated to treatment for illness A	2
3)	The increase should be divided evenly between the two groups	3
4)	I am not able to make a decision and would prefer that the choice be made by others	4
	Other (PLEASE WRITE IN AND CODE '5')	5
	Don't Know	6
	None	7

(62)

MORI

Q16. **Finally, do you wish to add anything else on this subject?**
PROBE FULLY AND WRITE IN.

ANY ANSWER (WRITE IN AND CODE '1')	1	
None/no answer	X	
Don't know	Y	(63)

THIS MUST BE THE LAST PAGE OF THE QUESTIONNAIRE AND MUST BE SINGLE SIDED

MORI/J25567 Questionnaire No ☐☐ Serial No
1-5 OUO (6-9)
CARD * 10

Priority Health Setting in Southampton Survey
Main Stage - Questionnaire Version 2 (01/09/05)

Sample Point Number: ☐☐☐ (11) (12) (13)
Sample point name:

Interviewer Number: ☐☐☐☐/☐ (14) (15) (16) (17) (18) (14-18)
Interviewer Name:.....

Name/Initial/Title: Mr/Mrs/Ms/Miss

Address:

Full Postcode ☐☐☐☐☐☐☐ (19) (20) (21) (22) (23) (24) (25) (19-25)

QTEL1 **Do you have a fixed line telephone at home which you use for incoming and outgoing voice calls?**

(26)

Yes	1	
WRITE IN Full tel. No		
No	2	
Refused	3	GO TO QTEL2
Ex-directory	4	

(26)

QTEL2 ASK IF NO FIXED LINE/REFUSED/EX-DIRECTORY (CODES 2-4). OTHERS CLOSE
Can I just check, do you have a mobile phone? IF YES ASK: **Can I take the number please?**

(27)

Yes	1	
WRITE IN Full tel. No		
No	2	
Refused	3	

(27)

IT IS VERY IMPORTANT THAT YOU FILL IN THE SAMPLE POINT NUMBER, QUESTIONNAIRE NUMBER AND INTERVIEWER NUMBER ON BOTH THE FRONT AND BACK PAGE OF THE QUESTIONNAIRE.

SHOWCARDS FOR DCE:

QUESTIONNAIRE VERSION 2

SHOWCARD A

SEVERITY OF THE HEALTH CONDITION

Whether patients are severely affected by their condition

When considering severity we have judged that on at least one of the following areas patients have severe problems: (i) self-care *(e.g. unable to wash or dress themselves)*, (ii) unable to perform usual activities *(e.g. work, study, housework, family or leisure activities)*, (iii) extreme pain or discomfort, (iv) extreme anxiety or depression.

IMPROVEMENT IN HEALTH

Considering the average health improvement from treatment

Improvement in health refers to the benefits that the patient feels following treatment e.g. improvements in their mobility, improvements in their ability to perform usual activities, reduced pain, reduced anxiety.

In this questionnaire treatments offer one of the following:

- large improvement in health
- moderate improvement in health
- small improvement in health
- very small improvement in health

VALUE FOR MONEY

Whether or not the treatment offers 'value for money'

'Value for money' is a common consideration within the NHS. Value for money is about the efficient use of resources (e.g. doctor's time, hospital beds, healthcare funds).

In this questionnaire we use the following categories:

- very good value for money
- fairly good value for money
- fairly poor value for money
- very poor value for money

WHETHER OTHER EFFECTIVE TREATMENTS ARE AVAILABLE

Yes available or No not available

We assume that all patients will have usual and best supportive care available within the NHS (e.g. GP services, specialist outpatient appointments, and best supportive nursing care). Where referring to treatment options we are thinking about drugs, surgery, or specially developed services (such as physiotherapy, support services, occupational therapy, specialist education, preventative therapies, etc).

SHOWCARD B

Health condition: Long-term health condition

	OPTION A	OPTION B
Severity of patients	Not severely affected	Severely affected
Improvement in health	Large improvement in health	Small improvement in health
Does the treatment offer Value for money	Fairly good value for money	Very good value for money
Other effective treatment options available	No	Yes

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

MORI\J25567 – For use with questionnaire version 2

SHOWCARD C

Health condition: Long-term health condition

	OPTION C	OPTION D
Severity of patients	Severely affected	Not severely affected
Improvement in health	Large improvement in health	Small improvement in health
Does the treatment offer Value for money	Very good value for money	Very poor value for money
Other effective treatment options available	No	Yes

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

SHOWCARD D

Health condition: Long-term health condition

	OPTION E	OPTION F
Severity of patients	Not severely affected	Severely affected
Improvement in health	Moderate improvement in health	Small improvement in health
Does the treatment offer Value for money	Very poor value for money	Very good value for money
Other effective treatment options available	No	Yes

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

SHOWCARD E

Health condition: Long-term health condition

	OPTION G	OPTION H
Severity of patients	Severely affected	Not severely affected
Improvement in health	Very small improvement in health	Large improvement in health
Does the treatment offer Value for money	Fairly poor value for money	Very poor value for money
Other effective treatment options available	Yes	No

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

SHOWCARD F

Health condition: Long-term health condition

	OPTION I	OPTION J
Severity of patients	Not severely affected	Severely affected
Improvement in health	Small improvement in health	Very small improvement in health
Does the treatment offer Value for money	Very good value for money	Fairly good value for money
Other effective treatment options available	No	Yes

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

SHOWCARD G

Health condition: Long-term health condition

	OPTION K	OPTION L
Severity of patients	Not severely affected	Severely affected
Improvement in health	Large improvement in health	Moderate improvement in health
Does the treatment offer Value for money	Very poor value for money	Very good value for money
Other effective treatment options available	Yes	No

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

SHOWCARD H

Health condition: Long-term health condition

	OPTION M	OPTION N
Severity of patients	Severely affected	Not severely affected
Improvement in health	Small improvement in health	Very small improvement in health
Does the treatment offer Value for money	Fairly poor value for money	Very poor value for money
Other effective treatment options available	No	Yes

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

SHOWCARD I

Health condition: Long-term health condition

	OPTION O	OPTION P
Severity of patients	Severely affected	Not severely affected
Improvement in health	Large improvement in health	Moderate improvement in health
Does the treatment offer Value for money	Fairly good value for money	Fairly poor value for money
Other effective treatment options available	Yes	No

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

SHOWCARD J

Health condition: Long-term health condition

	OPTION Q	OPTION R
Severity of patients	Not severely affected	Severely affected
Improvement in health	Very small improvement in health	Large improvement in health
Does the treatment offer Value for money	Very good value for money	Fairly good value for money
Other effective treatment options available	Yes	No

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

SHOWCARD K

Health condition: Long-term health condition

	OPTION S	OPTION T
Severity of patients	Severely affected	Not severely affected
Improvement in health	Moderate improvement in health	Small improvement in health
Does the treatment offer Value for money	Fairly good value for money	Fairly poor value for money
Other effective treatment options available	No	Yes

You are reminded that only one of these options can be provided from the limited funds available.

A decision has to be made by the health care provider.

APPENDIX 5

APPENDIX 5 (Appendix to Chapter 7): Discrete choice experiment – experimental design

Basic design for discrete choice study (main effects plan)

Scenario	A	B	C	D	E
1	0	0	0	0	0
2	0	1	1	1	1
3	0	2	2	2	2
4	0	3	3	3	3
5	1	0	1	2	3
6	1	1	0	3	2
7	1	2	3	0	1
8	1	3	2	1	0
9	2	0	2	3	1
10	2	1	3	2	0
11	2	2	0	1	3
12	2	3	1	0	2
13	3	0	3	1	2
14	3	1	2	0	3
15	3	2	1	3	0
16	3	3	0	2	1

- Source: www.research.att.com/~njas/oadir/ 0a.16.5.4.2 design plan, 94.4% efficient (Sloan, 2003)
- Need only use columns A to C for the study here (4^2 x 2^2)

Adapted design for discrete choice study (main effects plan) 4 attributes: 2 x 2-level, 2 x 4-level

Scenario	A	B	C1	C2
1	0	0	0	0
2	0	1	0	1
3	0	2	1	0
4	0	3	1	1
5	1	0	0	1
6	1	1	0	0
7	1	2	1	1
8	1	3	1	0
9	2	0	1	0
10	2	1	1	1
11	2	2	0	0
12	2	3	0	1
13	3	0	1	1
14	3	1	1	0
15	3	2	0	1
16	3	3	0	0

- Basic design adapted to create 2 2-level attributes C1, C2, from column C
- Column C: C1 created where 0 or 1 = 0, 2 or 3 = 1; C2 created where 0 or 2 = 0, 1 or 3 = 1

Foldover design (Louviere *et al* 2000)

- Pair choices by creating foldover copy
- 2 level attributes foldover 0=1, 1=0
- 4-level attributes foldover 0=1, 1=2, 2=3, 3=0

	Adapted Design			
Choice/pair	A	B	C1	C2
1	0	0	0	0
2	0	1	0	1
3	0	2	1	0
4	0	3	1	1
5	1	0	0	1
6	1	1	0	0
7	1	2	1	1
8	1	3	1	0
9	2	0	1	0
10	2	1	1	1
11	2	2	0	0
12	2	3	0	1
13	3	0	1	1
14	3	1	1	0
15	3	2	0	1
16	3	3	0	0

FOLDOVER			
A	B	C1	C2
1	1	1	1
1	2	1	0
1	3	0	1
1	0	0	0
2	1	1	0
2	2	1	1
2	3	0	0
2	0	0	1
3	1	0	1
3	2	0	0
3	3	1	1
3	0	1	0
0	1	0	0
0	2	0	1
0	3	1	0
0	0	1	1

- There is zero correlation between attributes (columns) (i.e. col AB, AC, AD, BC). Both sides of the foldover checked for correlation. For balance in design, we see that the attribute levels appear the same number of times in each column.

2 blocks of 8 questions:

Additional column from the experimental design was used to split the 16 pairs into 2 blocks of 8 pairs (2 x 8).

Random ordering for each block of 8 questions:

When blocked into 2 x 8, using experimental methods to provide greatest balance, the ordering of each of the 8 pairs (2 x 8 blocked sets of questions) was determined via a random draw.

APPENDIX 6

APPENDIX 6 (APENDIX to Chapter 7)

REPORTING DCE ANALYSIS USING CONDITIONAL LOGIT MODEL

ANALYSIS IN STATA 8.1

MAIN EFFECTS (MODEL A)

```
. clogit  response asc sev othertx himp1 himp2 himp3 vfm1 vfm2 vfm3, group( dce
> group_id)

Iteration 0:  log likelihood = -1255.0672
Iteration 1:  log likelihood = -1121.1702
Iteration 2:  log likelihood = -1116.8388
Iteration 3:  log likelihood = -1116.8077
Iteration 4:  log likelihood = -1116.8077

Conditional (fixed-effects) logistic regression   Number of obs   =       4054
                                                  LR chi2(9)      =     576.40
                                                  Prob > chi2     =     0.0000
Log likelihood = -1116.8077                       Pseudo R2       =     0.2051

------------------------------------------------------------------------------
    response |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         asc |   .2688554   .0527779     5.09   0.000     .1654125    .3722983
         sev |   .5392658   .0533445    10.11   0.000     .4347125    .6438192
     othertx |   .0781391    .052769     1.48   0.139    -.0252862    .1815644
       himp1 |   1.377303   .0917833    15.01   0.000      1.197411    1.557195
       himp2 |   1.042834   .1050436     9.93   0.000       .836952    1.248715
       himp3 |   .4916643   .0878891     5.59   0.000     .3194048    .6639238
        vfm1 |   1.165533   .0892049    13.07   0.000     .9906949    1.340372
        vfm2 |   1.012153   .1058264     9.56   0.000     .8047376    1.219569
        vfm3 |   .2889197   .0872898     3.31   0.001     .1178348    .4600045
------------------------------------------------------------------------------
```

MAIN EFFECTS WITH INTERACTION (MODEL B)

. clogit response asc sev othertx himp1 himp2 himp3 vfm1 vfm2 vfm3 other_hstat, group(dcegroup_id)

note: 32 groups (32 obs) dropped due to all positive or
all negative outcomes.
Iteration 0: log likelihood = -1234.361
Iteration 1: log likelihood = -1101.5381
Iteration 2: log likelihood = -1097.1351
Iteration 3: log likelihood = -1097.1035
Iteration 4: log likelihood = -1097.1035

Conditional (fixed-effects) logistic regression Number of obs = 3990
LR chi2(10) = 571.45
Prob > chi2 = 0.0000
Log likelihood = -1097.1035 Pseudo R^2 = 0.2066

response	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
asc	.2730598	.0531972	5.13	0.000	.1687952	.3773245
sev	.5313652	.0537938	9.88	0.000	.4259313	.636799
othertx	.1243038	.0557855	2.23	0.026	.0149663	.2336412
himp1	1.375578	.0925976	14.86	0.000	1.19409	1.557066
himp2	1.04433	.1059989	9.85	0.000	.8365759	1.252084
himp3	.4958704	.0888163	5.58	0.000	.3217936	.6699473
vfm1	1.174411	.0900122	13.05	0.000	.9979907	1.350832
vfm2	1.031414	.1067475	9.66	0.000	.8221931	1.240636
vfm3	.3149636	.0883064	3.57	0.000	.1418863	.48804
other_hstat	-.4288868	.1829253	-2.34	0.019	-.7874138	-.0703598

REPORTING DCE ANALYSIS USING PROBIT MODEL (FIXED AND RANDOM EFFECT PROBIT)

1. RANDOM EFFECT PROBIT – MAIN EFFECTS WITH INTERACTION

```
. iis  unique_id
. tis  dcegroup_id

. xtprobit  response sev othertx himp1 himp2 himp3 vfm1 vfm2 vfm3  other_hstat,
> i( unique_id)
Fitting comparison model:
Iteration 0:   log likelihood =   -2787.83
Iteration 1:   log likelihood = -2416.5284
Iteration 2:   log likelihood = -2411.6115
Iteration 3:   log likelihood = -2411.6082
Fitting full model:
rho =  0.0     log likelihood = -2411.6082
rho =  0.1     log likelihood = -2498.0051
Iteration 0:   log likelihood = -2411.6092
Iteration 1:   log likelihood = -2411.6092
Iteration 2:   log likelihood = -2411.6092
Iteration 3:   log likelihood = -2411.6092

Random-effects probit regression              Number of obs      =      4022
Group variable (i): unique_id                 Number of groups   =       259

Random effects u_i ~ Gaussian                 Obs per group: min =         6
                                                             avg =      15.5
                                                             max =        16
                                              Wald chi2(9)       =    667.38
Log likelihood  = -2411.6092                  Prob > chi2        =    0.0000
------------------------------------------------------------------------------
    response |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sev |    .614728   .0421834    14.57   0.000     .5320501     .697406
      othertx |   .1180588   .0430663     2.74   0.006     .0336504    .2024671
       himp1 |   1.093842    .061287    17.85   0.000     .9737212    1.213962
       himp2 |   .6263999   .0595515    10.52   0.000      .509681    .7431188
       himp3 |   .5747901   .0587575     9.78   0.000     .4596275    .6899528
        vfm1 |   .8631932   .0602749    14.32   0.000     .7450565     .98133
        vfm2 |   .5996563    .05955    10.07   0.000     .4829404    .7163722
        vfm3 |    .360648    .058489     6.17   0.000     .2460118    .4752843
  other_hstat |  -.2460521   .1038988    -2.37   0.018      -.44969   -.0424142
       _cons |  -1.372314   .0664904   -20.64   0.000    -1.502633   -1.241995
-------------+----------------------------------------------------------------
     /lnsig2u |       -14   32.85559                     -78.39578    50.39578
-------------+----------------------------------------------------------------
     sigma_u |   .0009119   .0149802                     9.47e-18    8.78e+10
         rho |   8.32e-07   .0000273                     8.98e-35           1
------------------------------------------------------------------------------
Likelihood-ratio test of rho=0: chibar2(01) =     0.00 Prob >= chibar2 = 1.000
```

2. FIXED EFFECT PROBIT – MAIN EFFECTS WITH INTERACTION

. probit response sev othertx himp1 himp2 himp3 vfm1 vfm2 vfm3 other_hstat

Iteration 0: log likelihood = -2787.83
Iteration 1: log likelihood = -2416.5284
Iteration 2: log likelihood = -2411.6115
Iteration 3: log likelihood = -2411.6082

Probit estimates — Number of obs = 4022
LR chi2(9) = 752.44
Prob > chi2 = 0.0000
Log likelihood = -2411.6082 — Pseudo R^2 = 0.1350

response	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
sev	.614728	.0421834	14.57	0.000	.5320501	.6974059
othertx	.1180588	.0430662	2.74	0.006	.0336505	.2024671
himp1	1.093842	.061287	17.85	0.000	.9737213	1.213962
himp2	.6263999	.0595515	10.52	0.000	.509681	.7431188
himp3	.5747901	.0587575	9.78	0.000	.4596275	.6899527
vfm1	.8631932	.0602749	14.32	0.000	.7450566	.9813299
vfm2	.5996563	.05955	10.07	0.000	.4829405	.7163722
vfm3	.360648	.0584889	6.17	0.000	.2460118	.4752843
other_hstat	-.2460521	.1038986	-2.37	0.018	-.4496896	-.0424146
_cons	-1.372314	.0664904	-20.64	0.000	-1.502632	-1.241995

REPORTING DATA ANALYSIS IN PROBIT MODEL WITH EFFECTS CODING (as an alternative to dummy variables)

1. RANDOM EFFECT PROBIT – MAIN EFFECTS WITH INTERACTION (effects coding)

```
. iis unique_id
. tis dcegroup_id

. xtprobit response sevb otherb himp1b himp2b himp3b vfm1b vfm2b vfm3b, i( unique_id)

Fitting comparison model:
Iteration 0:  log likelihood = -2810.0187
Iteration 1:  log likelihood = -2437.9593
Iteration 2:  log likelihood = -2433.1226
Iteration 3:  log likelihood = -2433.1195
Iteration 4:  log likelihood = -2433.1195

Fitting full model:
rho =  0.0    log likelihood = -2433.1194
rho =  0.1    log likelihood = -2521.1809
Iteration 0:  log likelihood = -2433.1204
Iteration 1:  log likelihood = -2433.1204

Random-effects probit regression              Number of obs      =      4054
Group variable (i): unique_id                 Number of groups   =       259
Random effects u_i ~ Gaussian                 Obs per group: min =         6
                                                             avg =      15.7
                                                             max =        16
                                              Wald chi2(8)       =    669.49
Log likelihood  = -2433.1204                  Prob > chi2        =    0.0000
------------------------------------------------------------------------------
    response |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        sevb |   .3089501   .0209992    14.71   0.000     .2677924    .3501077
      otherb |   .0453293   .0209108     2.17   0.030     .0043449    .0863137
     himp1b |   .5196386   .0372829    13.94   0.000     .4465654    .5927118
     himp2b |    .051397   .0359838     1.43   0.153    -.0191299     .121924
     himp3b |   .0008159   .0353305     0.02   0.982    -.0684307    .0700624
      vfm1b |   .4095048   .0369391    11.09   0.000     .3371055    .4819041
      vfm2b |    .141051   .0363467     3.88   0.000     .0698128    .2122893
      vfm3b |   .1004632   .0355699    -2.82   0.005    -.1701788   -.0307475
      _cons |   .0101742    .020944     0.49   0.627    -.0308753    .0512237
-------------+----------------------------------------------------------------
   /lnsig2u |        -14   32.46628                     -77.63273    49.63273
-------------+----------------------------------------------------------------
    sigma_u |   .0009119   .0148027                      1.39e-17    5.99e+10
        rho |   8.32e-07   .000027                       1.93e-34           1
------------------------------------------------------------------------------
Likelihood-ratio test of rho=0: chibar2(01) =     0.00 Prob >= chibar2 = 1.000
```

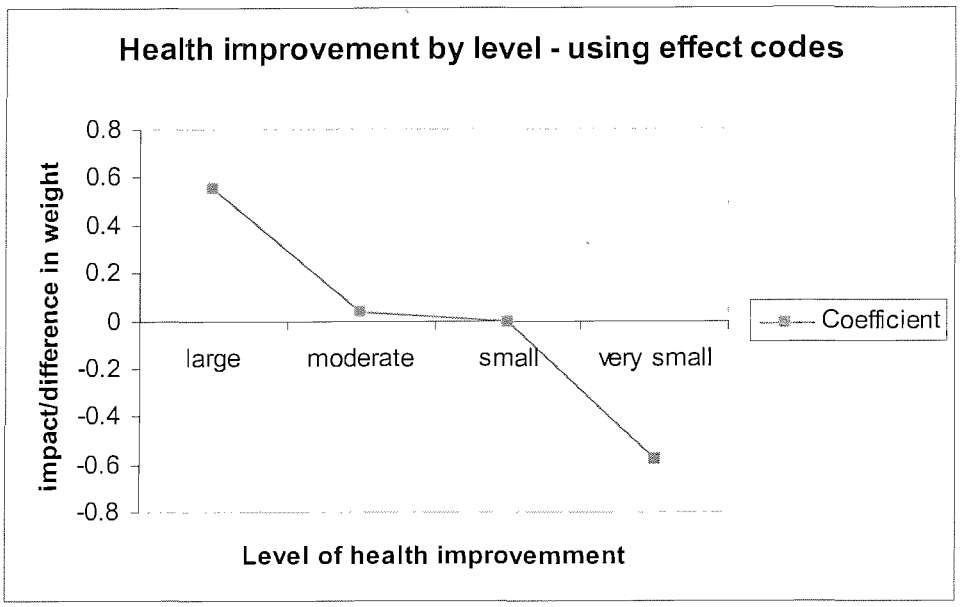
2. FIXED EFFECT PROBIT – MAIN EFFECTS WITH INTERACTION (with effects codes)

. probit response sevb otherb himp1b himp2b himp3b vfm1b vfm2b vfm3b

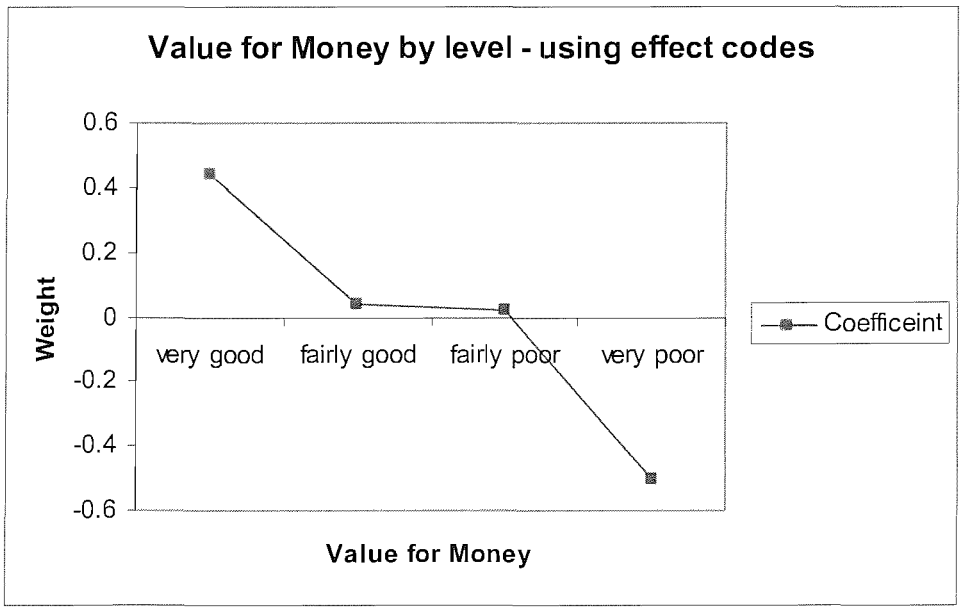
Iteration 0: log likelihood = -2810.0187
Iteration 1: log likelihood = -2437.9593
Iteration 2: log likelihood = -2433.1226
Iteration 3: log likelihood = -2433.1195
Iteration 4: log likelihood = -2433.1195

Probit estimates — Number of obs = 4054
LR chi2(8) = 753.80
Prob > chi2 = 0.0000
Log likelihood = -2433.1195 — Pseudo R^2 = 0.1341

response	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
sevb	.3089501	.0209992	14.71	0.000	.2677924	.3501077
otherb	.0453293	.0209108	2.17	0.030	.0043449	.0863137
himp1b	.5196386	.0372829	13.94	0.000	.4465654	.5927118
himp2b	.051397	.0359838	1.43	0.153	-.0191299	.121924
himp3b	.0008159	.0353305	0.02	0.982	-.0684307	.0700624
vfm1b	.4095048	.0369391	11.09	0.000	.3371055	.4819041
vfm2b	.141051	.0363467	3.88	0.000	.0698128	.2122893
vfm3b	-.1004632	.0355699	-2.82	0.005	-.1701788	-.0307475
_cons	.0101742	.0209439	0.49	0.627	-.0308752	.0512235



	Health improvement	Coefficient
large	1	0.519639
moderate	2	0.051397
small	3	0.000816
very small	4	-0.57185



	Value for money	Coefficeint
very good	1	0.409505
fairly good	2	0.141051
fairly poor	3	-0.10046
very poor	4	-0.45009

APPENDIX 7

APPENDIX 7 (Appendix to Chapter 8)

A presentation of the sample of health technologies appraised by NICE, selected for use as examples in Chapter 8, and presented in Table 21 of Chapter 8.

The summary provides outline detail on the key issues informing the NICE technology appraisal process, as relevant to the consideration of decision making presented in Table 21 (Chapter 8).

The vast majority of the information presented in the short outlines below is drawn from the published NICE guidance (reports available online at www.nice.org.uk). Further additional information has been drawn from the published independent assessment reports associated with the appraisal (available online at www.hta.ac.uk), this has mainly been necessary for the information on health improvement, and further detail around cost effectiveness estimates.

1 GUIDANCE ON THE USE OF IMATINAB FOR CHRONIC MYELOID LEUKAEMIA

1. NICE Guidance TA70 (October 2003)

Chroinc myeloid leukaemia (CML)
CML is one of the most common types of leukaemia. In CML, the bone marrow produces an excessive number of abnormal stem cells. The abnormal cells eventually suppress the production of normal white blood cells, which act to protect the body against infection. In 95% of people with CML there is a chromosomal abnormality which influences cellular processes such as proliferation, differentiation and survival. CML usually has three identifiable phases: the chronic phase (typically lasting 3-5 years), the accelerated phase (seen in around 2/3rds of cases), and the blast crisis (lasting 3-6 months and leading to death).

Technology
Imatinib (Glivec) is the first in a new class of cancer drugs, the signal transduction inhibitors, rationally designed to competitively inhibit BCR-ABL tyrosine kinase activity. By blocking specific signals in cells expressing BCR-ABL protein, imatinab reduces the uncontrolled proliferation of white blood cells that is a characteristic feature of chronic myeloid leukaemia (CML).

NICE Guidance/Recommendation
NICE recommend imatinib as first-line treatment for people with Philadelphia-chromosone-positive CML in the chronic phase. Imatinib is recommended as an option for the treatment of people with Philadelphia-chromosone-positive CML who initially present in the accelerated phase or with blast crisis.

Alternative Treatments (Yes)
The guidance from NICE regarded interferon alfa (IFN-α) as the most appropriate comparator (alternative treatment). Other potential options were stem cell transplant (but availability highly limited), and conventional chemotherapy with hydrosyurea (HU).

Cost-effectiveness
In their guidance NICE discuss estimates of cost-effectivness compared to IFN-α and HU. They discuss estimates of cost-effectiveness available in the published literature, from the manufacturer, and from the independent assessment team. NICE suggest that the estimate of imatinab vs. IFN-α, at £26,000 per QALY (from the assessment team) is a reasonable estimate and state on this basis that it represents a cost-effective treatment option. They state that when compared to HU it is not a cost-effective treatment option, but accept that IFN-α is the most appropriate comparator.

Severity: How CML affects patients with the condition
Patients can be regarded as being severely affected by this condition.
The chronic phase (initial phase) of CML is usually relatively stable and benign, but accelerated phase and blast crisis represent a rapid development of disease leading to death.

Health Improvement
In the cost per QALY estimates cited above (£26,000 per QALY) the mean QALY gain for imatinab vs. IFN-α is 1.99 QALYs per person treated, over the time horizon of the model used (20 year time horizon). This equates to a mean improvement per year of circa. 0.10 per person.

2 GUIDANCE ON THE USE OF DROTRECOGIN ALFA (ACTIVATED) FOR SEVERE SEPSIS

1. NICE Guidance TA084 (September 2004)

Severe Sepsis
Sepsis is a clinical response to infection, and it is termed severe when it is associated with organ failure, tissue hypoperfusion or hypotension.

Technology
Drotrecogin alfa (activated) is a recombinant form of human activated protein C. It is a treatment for patients with severe sepsis, licensed for use in patients with severe sepsis with multiple organ failure when added to best standard care.

NICE Guidance/Recommendation
NICE recommend drotrecogin alfa (activated) for use in adult patients who have severe sepsis that has resulted in multiple organ failure, and who are being provided with optimum intensive care. They state that it should only be initiated and supervised by a specialist with intensive care skills and experience in the care of patients with sepsis.

Alternative Treatments (No)
The alternative treatment scenario is regarded as best supportive treatment involving bothe treatment of the underlying infection, primarily with antibiotics, and supportive treatment (which may include fluids, steroids, vasopressors and ventilatory and renal support). No direct alternative treatment is available

Cost-effectiveness
In their guidance NICE discuss estimates of cost-effectivness compared to best supportive care. They discuss estimates of cost-effectiveness available in the published literature, from the manufacturer, and from the independent assessment team. The UK analyses presented to NICE indicated a cost per QALY 0f less than £11,000. NICE state that overall they were persuaded the intervention was a cost-effective option for patients with severe sepsis whose risk of death was increased because of multiple organ failure.

Severity: How severe sepsis affects patients with the condition
This is a serious and life threatening condition, requiring intensive care, and resulting in a mean mortality rate of 30% to 50%. Patients can be regarded as being severely affected by this condition.

Health Improvement
In the cost per QALY estimates cited the assessment group estimated a gain of 1.4 life years per treated patient, a mean incremental QALY improvement of 0.810 per treated person is reported (Green *et al*, 2005).

3 GUIDANCE ON THE USE OF RILUZOLE (RILUTEK) FOR THE TREATMENT OF MOTOR NEURONE DISEASE

1. NICE Guidance TA020 (January 2001)

Motor Neurone Disease (MND)
MND is characterised by degeneration of motor neurones of the brain, brain stem or spinal cord. MND usually starts insidiously with symptoms and signs including stumbling, foot drop, weakened grip, slurred speech, cramp, muscle wasting, twitching and tiredness. Other symptoms include muscle stiffness, paralysis, incoordination and impaired speech, swallowing and breathing. Most patients die from ventilatory failure, resulting from progressive weakness and wasting of limb, respiratory and bulbar muscles within approximately 3 years of the onset of symptoms.

Technology
Rilozole is licensed for use to extend life or the time to mechanical ventilation for persons with amyotrophic lateral sclerosis (ALS), the most common form of MND.

NICE Guidance/Recommendation
Riluzole is recommended for the treatment of individuals with the ALS form of MND. NICE state that it should be initiated by a neurological specialist with expertise in the management of MND.

Alternative Treatments (No)
At the time of the appraisal (and at current time) riluzole is the only drug licensed for treatment of MND (ALS). The only alternative is supportive and palliative care.

Cost-effectiveness
In their guidance NICE discuss estimates of cost-effectiveness compared to supportive care. They discuss estimates of cost-effectiveness available from the manufacturer (£18-£29,000 per QALY) and from the independent assessment team (£34-£43,500 per QALY). NICE considered that the net increase in cost for the NHS was reasonable when set against the benefits, assessed as extended months of acceptable (to patients) quality of life.

Severity: How MND affects patients with the condition
Patients can be regarded as being severely affected by this condition (as above).

Health Improvement
The NICE guidance states that the suggested health gain is a median of 2 to 4 months of tracheostomy free survival. The cost per QALY estimates cited in the published assessment report (Stewart *et al*, 2001) are a base case of £58,000 per QALY, when the mean QALY gain per patient is 0.09, and a range of between £25,000 per QALY (using a mean QALY gain of 0.21) to a scenario where riluzole is dominated by supportive care (with no QALY gain). Where NICE guidance refers to estimates of cost per QALY from the assessment team, it is not clear how to cross-reference these with the published report. There is an appendix with updated further CEA analyses, but data on mean QALY gains is not presented. For the current thesis, the above published range of mean QALY benefits is expected to capture the health improvements offered by riluzole.

4 GUIDANCE ON THE USE OF CONTINUOUS SUBCUTANEOUS INSULIN INFUSION FOR DIABETES

1. NICE Guidance TA057 (February 2003)

Diabetes

Diabetes is a chronic metabolic disorder caused by defects in insulin secretion and action. There are two major types of diabetes, type 1 and type 2. In type 1 diabetes the pancreas makes little or no insulin and people depend on daily insulin injections to survive. Type 2 diabetes results from failure of insulin production to overcome insulin resistance, it is a progressive disease in which insulin production declines as the disease progresses. An impaired insulin effect results in increased levels of glucose in the blood which can, if prolonged, cause microvascular and macrovascular damage in the body. Common complications of diabetes include visual impairment, kidney failure, angina, myocardial infarction, stroke, foot ulceration and erectile dysfunction.

Technology

Continuous subcutaneous insulin infusion (CSII) devices are external pumps comprising a programmable pump and insulin storage reservoir to which the patient is continuously connected. Insulin is administered to the patient via a needle or cannula inserted under the skin. The pump delivers insulin continuously at a constant or variable basal rate with an additional boost dose delivered at meal times.

NICE Guidance/Recommendation

NICE recommend CSII as an option for people with type 1 diabetes provided that (i) multiple-dose insulin therapy has failed, and (ii) those receiving treatment have the commitment and competence to use the therapy effectively. NICE state that CSII therapy should be initiated only be a trained specialist team, and that those beginning CSII therapy should be provided with specific training in its use.
NICE do not recommend CSII for people with type 2 diabetes who require insulin therapy. NICE state that there was insufficient evidence to draw any conclusions from studies comparing CSII with MDI in type 2 diabetes.

Alternative Treatments (Yes)

For type 1 diabetes the comparator was stated as multiple-dose insulin therapy (MDI).

Cost-effectiveness

In their guidance NICE discuss estimates of cost-effectiveness compared to MDI They discuss estimates of cost-effectiveness available from the manufacturer (£8,400 per QALY), and from cost analysis available from the independent assessment team, as well as suggestions from the assessment team that CSII could result in a very high cost per QALY (circa. £500,000 per QALY) when using the manufacturer model with more pessimistic assumptions NICE had no reliable information on estimates of cost-effectiveness from the formal clinical evidence. NICE suggest that an estimate of the increase in utility may be of the order of 0.035 or more, and this with an estimated net cost of £1,100 per year led the committee to recommend CSII in type 1 diabetes (on this basis, assuming such a gain each year [and that is uncertain] it can be estimated at circa. £31,500 per QALY, although the assessment group suggest a much higher estimate).

Health Improvement

In the discussions around cost-effectiveness the mean QALY improvement is reported as circa. 0.035 or more (this is a hypothesised health gain, and no analysis is presented to support this estimate).

Severity: How type 1 diabetes affects patients with the condition
Whilst longer term complications can be very serious, and the ongoing need for insulin is a burden for patients and carers, people with diabetes are able to lead a full and normal life where disease is managed and controlled effectively. Therefore, it would seem reasonable to regard patients with type 1 diabetes as not being severely affected by their condition (i.e. no severe impairment on mobility, usual care, activities of daily living, etc.

Health Improvement
As above, NICE suggest that an estimate of the increase in utility may be of the order of 0.035 or more. The assessment group report, suggests that it is difficult to estimate to health gain, in terms of QALYS, with benefits being of a fairly intangible nature (i.e. lifestyle convenience), with patients unlikely to suggest QALY gains via accepting a risk of death [standard gamble technique] or forgoing future life-expectancy [time trade-off technique]. There is no analysis to support the NICE estimate of health improvement.

5 GUIDANCE ON THE USE OF DRUGS FOR ALZHEIMER'S DISEASE (DONEPEZIL, GALANTAMINE, RIVASTIGMINE, MEMANTINE)

1. NICE Guidance TA111 (2006)

Alzheimer's Disease

Alzheimer's disease (AD) is the most common form of dementia. It is a degenerative cerebral disease with characteristic neuropathological and neurochemical features. It is a chronic progressive mental disorder that adversely affects higher cortical functions including memory, thinking and orientation. Disease progression is characterised by deterioration of cognition (thinking, conceiving, reasoning) and functional ability (activities of daily living) and a disturbance in behaviour and mood. People with AD lose the ability to carry out routine daily activities like dressing, toileting, travelling and handling money, and as a result many of them require a high level of care. People with mild dementia are sometimes able to cope without assistance, but as the disease progresses, all eventually require the aid of carers, and about half need residential care.

Technology

Donepezil, galantamine and rivastigmine are cholinesterase inhibitors, licensed (in the UK) for the symptomatic treatment of people with mild to moderately severe AD. These drugs were recommended for use by NICE in an initial guidance issued in 2001, this appraisal reviews that earlier guidance. Memantine is an NMDA-receptor antagonist licensed for the treatment of moderate to severe AD (more severely affected patient group).

NICE Guidance/Recommendation

The NICE final appraisal determination is subject to appeal at the time of writing. It states that the three cholinesterase inhibitors donepezil, galantamine and rivastigmine are recommended as options in the management of people with AD of moderate severity only (MMSE score of between 10-20), under the guidance of specialists in the care of people with dementia. Those prescribed these drugs are required to be reviewed every 6-months. These drugs are not recommended for people with mild AD (MMSE greater than 20). Memantine is not recommended as a treatment options for people with AD, except as part of well designed clinical studies.

Alternative Treatments (None)

These drugs are the only treatments available for those with AD. The alternative is best supportive care. All drugs have been compared to placebo. Memantine can be used together with cholinesterase inhibitors in those patients with moderately-severe AD.

Cost-effectiveness

In their guidance NICE discuss estimates of cost-effectiveness compared to placebo (best supportive care). They discuss estimates of cost-effectiveness available in the published literature, from the manufacturer, and from the independent assessment team. For the cholinesterase inhibitors NICE state that the economic model presented by the assessment group was the most appropriate to use to assess cost-effectiveness. The assessment group presented cost per QALY estimates for cholinesterase inhibitors that ranged from £70,000 to £97,000 per QALY. NICE took other factors into consideration and presented a wide range of additional analyses, and sub-group analyses. Finding that cost per QALY estimates for those people with moderate AD (MMSE 10-20) could be reduce to between £30,000 to £40,000, and NICE suggested it was possible that the cost-effectiveness of cholinesterase inhibitors in the moderate AD group could be between, £23,000 to £35,000 (depending on the choice of cholinesterase inhibitor, and the inclusion of carer benefits in the adjusted analyses). The NICE guidance notes that the cost-effectiveness of these drugs could be less favourable than the findings in these additional analyses. However, in consideration of the evidence the NICE appraisal committee concluded that the evidence suggested that for

people with moderate AD (MMSE 10-20) treatment with cholinesterase inhibitors was cost effective. The cost-per-QALY estimates for mild AD were between £56,000 and £72,000. For memantine the NICE appraisal committee considered the evidence on clinical effectiveness to be insufficient. On the basis of this evidence it concluded that memantine could not reasonably be considered a cost-effective therapy for moderately severe to severe AD (indicating a cost-per-QALY of between £70,000 to £90,000).

Severity: How Alzheimer's disease affects patients with the condition
Patients can be regarded as being severely affected by this condition. When people have mild AD the condition itself may not dictate a health state description that is of a severe nature, but the rapid onset of disease will lead to severe impairment.

Health Improvement
The NICE guidance states that the NICE appraisal committee considered the cost-effectiveness model submitted by the assessment group to be the most appropriate. This model estimated a mean QALY gain of between 0.032 to 0.035. Additional analyses by NICE (using this model) estimated a mean QALY gain up to 0.06, and NICE noted that it was not persuaded that these average QALY gains could be reasonably increased any further.

6 GUIDANCE ON THE USE OF BETA INTERFERON AND GLATIRAMER ACETATE FOR THE TREATMENT OF MULTIPLE SCELROSIS

1. NICE Guidance TA032 (2002)

Multiple Sclerosis

Multiple sclerosis is a disabling neurological disease. It is characterised by repeated episodes of inflammation of the nervous tissue in the brain and spinal cord, resulting in the removal of the insulating myelin sheath covering the nerves. Multiple areas of scar tissue (sclerosis) form along the nerve fibres, slowing or blocking transmission of signals to and from the brain and spinal cord, so that movements such as movement and sensation may be lost. There are several forms of MS. The vast majority of people (80%-90%) start with relapsing remitting MS (RRMS). In this form of the disease, recurrent attacks of loss of neurological function, termed relapses, are separated by periods of complete or incomplete recovery, described as remissions. After about 10-years (without treatment), about half of people with MS begin a continuous downward progression, which may also include relapses. This form of MS is known as secondary progressive (SPMS).

Technology

These treatments are presented as disease-modifying treatments, targeted at reducing the frequency and/or severity of relapses and/or slowing the course of disease.
The range of beta interferons available work by reducing the inflammatory process the characterises and MS relapse, but the mode of action is uncertain. Beta interferons are licensed for treatment of both RRMS and SPMS.
Glatiramer acetate is licensed for treatment of RRMS. It works by reducing the inflammation around nerves. Such inflammation usually precedes an MS relapse.

NICE Guidance/Recommendation

On the balance of their clinical and cost-effectiveness neither beta interferon nor glatiramer acetate is recommended for the treatment of multiple sclerosis.

Alternative Treatments (None)

Beta interferon and glatiramer are the only disease-modifying treatment options available.

Cost-effectiveness

In their guidance NICE discuss estimates of cost-effectiveness compared to best supportive care (placebo). They discuss estimates of cost-effectiveness initially presented by manufacturers and from the independent assessment team. These initial estimates ranged from £10,000 to £3 million pounds per QALY. NICE commissioned additional analyses. When this additional modelling considered analysis over 5- and 10-year periods, the cost-per-QALY ranged from £380,000 to £780,000 for 5-years, and from £190,000 to £425,000 for the 10-year analysis. Whilst using a longer term time horizon leads to assumptions becoming increasingly unreliable, NICE considered a 20-year time horizon with estimates of £40,000 to £90,000 per QALY. On the balance of costs and benefits NICE found that the beta interferons and glatiramer acetate were not cost effective.

Severity: How multiple sclerosis affects patients with the condition

MS has an adverse and often highly debilitating impact on the quality of life of people with MS and their families. Even in its early stages MS restricts activity and undermines confidence, with patients often unable to continue with employment, and unable to take part in usual activities. For the purposes of this thesis it is clear that patients can be regarded as being severely affected by this condition.

Health Improvement

In the cost per QALY estimates cited ranging £40,000 to £90,000 per QALY, the mean QALY improvement is reported at between 0.32 to 1.02 QALYs. These estimates vary around the wide and varied analyses presented. For example, where 5-year analyses are presented the mean QALY gain is reported at between 0.04 to 0.06 QALYs.

7 GUIDANCE ON THE USE OF TRASTUZUMAB FOR THE ADJUVANT THERAPY OF EARLY-STAGE HER2-POSITIVE BREAST CANCER

1. NICE Guidance TA107 (2006)

Early Stage Breast Cancer (Stages I to III)

Breast cancer is a complex heterogeneous disease characterised by uncontrolled growth and spread of abnormal cells. The stage of a breast cancer is a measure of how far it has progressed. So, a cancer that is small and confined to the breast is at an early stage, whereas one that has spread to many different parts of the body is at an advanced stage. Stage 1 - the earliest stage – is when the cancer is only in the breast and does not affect the skin overlying the breast. Stage 2 - the next stage – is when the cancer has spread to the lymph nodes under the arm. Stage 3 - this is when the cancer involves the skin of the breast (whether or not it has spread to the lymph nodes under the arm). Stage 4 is when the cancer has spread to other parts of the body. Breast cancer is the leading cause of death amongst women aged 35-54 years in the UK. The aetiology of breast cancer is unclear, although it is likely that hormonal and genetic factors play a role. Figures suggest that approximately 50% or women presenting with early stage breast cancer (stages I to III) will eventually progress to develop advanced disease.

Technology

Trastuzumab (Herceptin; Roche Products) is a recombinant humanised IgG1 monoclonal antibody directed against the human epidermal growth factor receptor 2 (HER2). It is licensed for the treatment of patients with early-stage HER-2 positive breast cancer, following surgery, chemotherapy, and radiotherapy (if applicable). It is licensed as an intravenous infusion, given at intervals of either 1 or 3-weeks.

NICE Guidance/Recommendation

Trastuzumab, given at 3-week intervals for 1-year or until disease recurrence, is recommended as a treatment option for women with early-stage HER2-positive breast cancer following surgery, chemotherapy and radiotherapy (if applicable).

Alternative Treatments (Yes)

This is the first of the class of drugs available for the licensed indication, but other treatment options are available. In their guidance NICE discuss estimates of cost-effectiveness based on the comparison of trastuzumab following surgery and standard adjuvant therapy to surgery and standard adjuvant therapy alone. Standard adjuvant therapy was defined as chempotherapy without trastuzumab.

Cost-effectiveness

In their guidance NICE discuss estimates of cost-effectiveness with trastuzumab following surgery and standard adjuvant therapy compared to surgery and standard adjuvant therapy alone. Standard adjuvant therapy was defined as chemotherapy without trastuzumab. NICE discuss estimates of cost-effectiveness available from the manufacturer, and from the independent assessment team (Evidence Review Group). The manufacturer submitted estimates of the cost per QALY ranging from £2,387 to £8,689. The Evidence Review Group presented an estimate of £18,000 per QALY, for the 3-weekly regimen. Other scenarios presented by the review group ranged £16,000 to £33,000 per QALY. NICE concluded that the cost per QALY estimates from the manufacturer were likely to be underestimates, and that the review groups estimate of £18,000 per QALY was more likely to reflect the cost effectiveness of trastuzumab (although there was uncertainty in analysis).

Severity: How breast cancer affects patients with the condition

As above, given the nature of disease, the impact of treatments, and the high mortality rates, patients can be regarded as being severely affected by this condition.

Health Improvement

Clinical evidence reports the benefit from treatment as a reduction in the relative risk of recurrence at 1 year follow-up. The RCT providing the primary source of clinical evidence, reported 87.1% of patients in the control arm and 92.5% of patients in the trastuzumab arm free from disease at 1-year follow-up, equating to a 46% reduction in the relative risk of recurrence. There was also a 24% relative reduction in the risk of mortality for patients taking trastuzumab. In the cost per QALY estimates presented by the manufacturer the mean QALY gain is in excess of 2 QALYs (e.g. for ICER of £5,678, mean QALY gain of 2.43 QALYs). The cost effectiveness analysis from the evidence review group does not present mean incremental costs and benefits separately.

8 GUIDIANCE ON THE USE OF BORTEZOMIB IN THE TREATMENT OF MULTIPLE MYELOMA

1. NICE Guidance TA (*in progress*) (2007)

Multiple Myeloma
Multiple myeloma is a relapsing, recurring haematological cancer arising from malignant, antibody producing plasma cells in the bone marrow. It is a debilitating incurable disease with a poor prognosis (survival at 1-year approx 60%, at 5-years approx. 25%). Morbidity associated with multiple myeloma is significant, typically symptoms are painful, distressing and disabling (including fractures of the long bones and vertebral collapse, renal failure and anaemia).

Technology
Bortezomib is an anti-cancer drug that belongs to a novel class of drugs know as proteasome inhibitors. It is licensed (in UK) as monotherapy for the treatment of progressive multiple myeloma in patients who have received at least on prior therapy and who have undergone, or are unsuitable for bone marrow transplantation.

NICE Guidance/Recommendation
Preliminary recommendations: Bortezomib monotherapy in its licensed indication is not recommended for the treatment of patients with multiple myeloma except for use in well-designed clinical studies.

Alternative Treatments (Yes)
The primary clinical evidence available (APEX RCT) compares bortezomib to high dose dexamethasone. Whilst there may be some debate over the relative effectiveness of alternative therapies, there are a number of other treatment options available (including thalidomide, anti-cancer chemotherapy, repeat stem cell transplant).

Cost-effectiveness
In their guidance NICE discuss estimates of cost-effectiveness compared to HDD. They discuss estimates of cost-effectiveness available from the manufacturer, and comments on cost-effectiveness from the independent assessment team (evidence review group). The manufacturer submitted an estimate of £31,000 per life-year-gained, equating this to £38,000 per QALY. The review group raised a number of concerns over the cost effectiveness estimates presented by the manufacturer, including concerns over the possible overestimation of survival benefits. However, based on the data presented by the manufacturer, the review group suggested a range of £27,000 to £45,000 per QALY. NICE guidance states that it considers the cost per life year estimate presented by the manufacturer to be an underestimate, and that the estimated £38,000 per QALY was an underestimate, and concluded that bortezomib had not been shown to be cost effective compared with current practice in the NHS.

Severity: How multiple myeloma affects patients with the condition
Patients can be regarded as being severely affected by this condition.

Health Improvement
In the cost per life year estimates cited (£30,750 per life-year-gained) the mean difference in overall survival was 9.9 months, which equates to 0.90 years, and a mean QALY gain of 0.73 (assuming a simple health state value of 0.80 applied to mean survival). This mean QALY estimate is the basis for the estimated £38,000 per QALY that NICE felt was an underestimate of the true cost-effectiveness of bortezomib.

9 GUIDANCE ON THE USE OF PHOTODYNAMIC THERAPY FOR AGE-RELATED MACULAR DEGENERATION

1. NICE Guidance TA068 (2003)

Age Related Macular Degeneration (ARMD)

ARMD is characterised by irreversible damage to the central part of the retina (the macula) resulting in progressive loss of central vision. Peripheral vision is not affected, so individuals retain some useful vision. The condition has two forms, dry and wet. The wet form is characterised by development of new blood vessels beneath the retina, known as choroidal neovascularisation (CNV). CNV can be subdivided into classic and occult forms according to its appearance. The classic form is associated with more rapid progression than the occult form. In people with wet form ARMD, the newly formed blood vessels may leak fluid and blood, leading to scar formation and permanent damage to the macula. Individuals lose visual acuity and contrast sensitivity. They may also experience distortion of vision. In 70% of cases eyes with CNV will have severe loss of vision within 2-years of diagnosis.

Technology

The aim of photodynamic therapy (PDT) is to destroy CNV lesions without damaging the overlying retina, thereby slowing or halting the progression of vision loss. The treatment involves the infusion of a light sensitive agent, followed by light activation of the drug. At the time of the appraisal only verteporfin, a benzoporphyrin derivative, is available for this indication, but other agents were said to be in development.

NICE Guidance/Recommendation

PDT is recommended for the treatment of wet ARMD for individuals who have a confirmed diagnosis of 'classic with no occult' CNV, and best corrected visual acuity 6/60 or better. PDT should be carried out only by retinal specialists with expertise in the use of this technology.
PDT is not recommended for the treatment of people with predominantly classic subfoveal CNV associated with wet ARMD, except as part of ongoing or new clinical studies.

Alternative Treatments (None)

At the time of the appraisal there were no alternative treatment options. For most patients with ARMD management consists of best supportive care.

Cost-effectiveness

In their guidance NICE discuss estimates of cost-effectiveness compared to best supportive care (placebo). They discuss estimates of cost-effectiveness available in the published literature, from the manufacturer, and from the independent assessment team. Published cost effectiveness data (USA) report a cost per QALY between £61,000 and £122,000, at 2 years. The manufacturer submission presents an estimate of approx. £70.500 per QALY, and the assessment group present a revised estimate of £80,000 per QALY at 2-years.
NICE report additional sub-group analyses undertaken for the appraisal process, with estimates of between £10,000 and £57,000 (at 2-years) for PDT in classic with no occult CNV. Further analyses by NICE, based on a scenario assumption of change in visual acuity, produced a cost per QALY of £26,000 (at 2-years) for this subgroup. The NICE guidance states that this latter estimate was viewed as the most reasonable estimate of cost effectiveness for the subgroup (PDT in classic with no occult CNV).
Additional modelling generated a cost per QALY of £55,000 for all people with predominantly classic CNV. At 5-years the cost-effectiveness ratios were reduced to around £8,500 for classic with no occult CNV, and £34,000 for the predominantly classic group. This additional analysis suggested that for those with predominantly classic CNV with any element of occult CNV the estimated cost per QALY was around £164,000 at 2-years and £120,000 at 5-years.
On the basis of this evidence the NICE appraisal committee concluded that PDT was likely to

be cost effective for people with classic with no occult CNV, but not for those with predominantly classic with some occult CNV.

Severity: How ARMD affects patients with the condition
Central vision loss particularly impairs the perception of fine visual detail and colours. Activities such as reading, recognising faces and driving are affected. Ability to work may be impaired, and significant loss of independence may occur. Deteriorating vision has an impact on emotional well-being, and individuals are likely to suffer depression and anxiety due to loss of vision and reduction in independence.
For the purposes of this thesis patients can clearly be regarded as being severely affected by this condition.

Health Improvement
In initial cost effectiveness estimates presented by the assessment group the mean QALY gain was of the order of 0.03 QALYs. Disaggregated data on QALYs for subsequent/additional analyses are not presented. Nor are data presented for the subgroups used in the final appraisal determination. Considering the commentary presented for sensitivity analyses in the assessment report, the current candidate does not expect the mean QALY improvement to differ substantially from that presented above i.e. not likely to be over a mean of 0.05 QALYs (given the estimate of mean costs in the base case at around £5,000), and judged (by the candidate) to be very unlikely to be over a mean QALY gain of 0.10 QALYs.

10 GUIDANCE ON THE USE OF ANAKINRA FOR RHEUMATOID ARTHRITIS

1. NICE Guidance TA072 (2003)

Rheumatoid arthritis
Rheumatoid arthritis (RA) is a chronic, progressive destructive and disabling condition that is associated with considerable morbidity and mortality, impacts severely on quality of life, and represents a considerable economic burden. RA affects all aspects of life, from education and employment through to family and social lives.

Technology
Interleukin-1 (IL-1) has been identified as one of the factors responsible for the damaging inflammatory processes that occur in RA. Anakinra is a recombinant, non-glycosylated form of human IL-1-receptor antagonist that inhibits the activity of IL-1, seeking to protect both cartilage and bone. Anakinra is licensed for used in combination with methotrexate in patients who have had an inadequate response to methotrexate alone.

NICE Guidance/Recommendation
Based on the balance of its clinical benefits and cost effectiveness anakinra is not recommended for the treatment of rheumatoid arthritis, except in the context of controlled long-term clinical study.

Alternative Treatments
A range of treatment options are available for RA. Management of RA involves physical therapy, surgical intervention and drug treatment, all running in parallel. Conventional drug therapy includes various combinations of non-steroidal anti-inflammatory drugs, analgesics, corticosteroids and disease-modifying anti-rheumatic drugs (DMARDs). The tumour necrosis factor a (TNFα) inhibitors etanercept and infliximab are available (at time of appraisal) for people with continuing, clinically active RA that has not responded adequately to at least two DMARDs, including methotrexate.
The key trial, informing the clinical evidence for anakinra in the appraisal, compared anakinra pus methotrexate to methotrexate plus placebo.

Cost-effectiveness *(CG Judgment: Very Poor)*
In their guidance NICE discuss estimates of cost-effectiveness compared to They discuss estimates of cost-effectiveness available from the manufacturer, and from the independent assessment team. The manufacturers model produced estimates of between £20,510 and £21,752 per QALY, however NICE suggest that there were a number of concerns over the industry analyses. The assessment team presented estimates of cost per QALY based on a range of scenarios, with estimates from £67,400 to over £500,000 per QALY. NICE considered additional analyses, using effectiveness data that it felt most appropriate, and this showed a cost per QALY in excess of £69,000.
The committee concluded that, although there was evidence of the clinical effectiveness of anakinra in the short term, the extent of the benefit was not sufficient to justify its cost.

Severity: How rheumatoid arthritis affects patients with the condition
As noted above, patients can be regarded as being severely affected by this condition.

Health Improvement
In the cost per QALY estimates presented by the assessment group (Clark *et al*, 2004) the mean QALY gains were between 0.016 and 0.109. Data on health gain is not presented in disaggregated form for the other cost-effectiveness estimates discussed in the NICE guidance.

APPENDIX 8

APPENDIX 8 (Appendix to Chapter 8)

NICE HEALTH TECHNOLGY APPRAISAL GUIDANCE (CURRENT), At June 2006 (source: www.nice.org.uk; and Raftery (2006))

NICE Guidance	Technology Appraisal Title	NICE recommendation (should the technology be available within NHS?)
TA001	Wisdom teeth - removal	No (due to lack of evidence or poor cost-effectiveness)
TA002	Hip disease - replacement prostheses	Yes (minor restrictions)
TA010	Asthma (childern under 5) - Inhaler devices	Yes (major restrictions)
TA014	Hepatitis C - alpha interferon and ribavarin	Yes (major restrictions)
TA017	Colorectal cancer - laparoscopic surgery	No (due to lack of evidence or poor cost-effectiveness)
TA019	Alzheimer's disease - donepezil, rivastigmine and galantamine	Yes (minor restrictions)
TA020	Motor neurone disease - riluzole	Yes
TA022	Obestity - orlistat	Yes (major restrictions)
TA023	Brain cancer - temozolomide	Yes (major restrictions)
TA024	Wound care - debriding agents	Yes
TA025	Pancreatic cancer - gemcitabine	Yes (major restrictions)
TA027	Osterarthritis and Rheumatoid arthritis - cox II inhibitors	Yes (major restrictions)
TA029	Leukaemia (lymphocytic) - fludarabine	Yes (minor restrictions)
TA030	Breast cancer - taxanes (review)	Paclitaxel/docetaxel 1st line: Yes (minor restrictions) Docetaxel, 2nd line: Yes (major restrictions)
TA031	Obesity - sibutramine	Yes (major restrictions)
TA032	Multiple sclerosis - beta interferon and glatiramer acetate	No (due to lack of evidence or poor cost-effectiveness)
TA034	Breast cancer (adv) - trastuzumab	Yes
TA035	Arthritis (juvenile idiopathic) - etanercept	Yes
TA036	Rheumatoid arthritis - etanercept and infliximab	Yes (minor restrictions)
TA037	Lymphoma (follicular non-Hodgkin's) - rituximab	No (due to lack of evidence or poor cost-effectiveness)
TA038	Asthma (older children) - inhaler devices	Yes (minor restrictions)
TA039	Smoking cessation - bupropion and nicotine replacement therapy	Yes (minor restrictions)
TA040	Crohn's disease - infliximab	Yes (minor restrictions)
TA041	Pregnancy - routine anti-D prophylaxis for rhesus negative women	Yes
TA042	Growth hormone deficiency (children) - human growth hormone	Yes (major restrictions)
TA043	Schizophrenia - atypical antipsychotics	Yes
TA044	Hip disease - metal on metal hip resurfacing	Yes (minor restrictions)
TA046	Obesity (morbid) - surgery	Yes (major restrictions)
TA047	Acute coronary syndromes - glycoprotein IIb/IIIa inhibitors (review)	Yes (major restrictions)
TA048	Renal failure - home versus hospital haemodialysis	Yes
TA049	Central venous catheters - ultrasound locating devices	2-D imaging: Yes (minor restrictions) Audio guided Doppler: No
TA052	Myocardial infarction - thrombolysis	Yes
TA053	Diabetes (types 1 and 2) - long acting insulin analogues	Type 1: Yes Type 2: Yes (major restrictions)
TA054	Breast cancer - vinorelbine	First line: No, 2nd line: Yes (major restrictions)
TA055	Ovarian cancer - paclitaxel (review)	Yes (minor restrictions)

TA056	Stress incontinence -tension-free vaginal tape	Yes (minor restrictions)
TA057	Diabetes - Insulin pumps (continuous subcutaneous insulin infusion - CSII)	Type 1: Yes (major restrictions)
TA058	Flu treatment - zanamivir (review), amantadine and oseltamivir	Yes (minor restrictions) in 'at risk adults', No in healthy adults
TA059	Electroconvulsive therapy (ECT)	Schizophrenia: No Severe depressive, catatonia, manic episode: Yes (major restrictions)
TA060	Diabetes (Types 1 & 2) - Patient education models	Yes
TA061	Colorectal cancer - capecitabine and tegafur uracil	Yes (minor restrictions)
TA062	Breast cancer - capecitabine	Yes (capecitabine + docetaxel), Yes * (capecitabine mono, 2nd line)
TA063	Diabetes (type 2) - glitazones (review)	Yes (major restrictions)
TA064	Growth hormone deficiency (adults) - human growth hormone	Yes (major restrictions)
TA065	Non-hodgkin's lymphoma - rituximab	Yes (major restrictions)
TA066	Bipolar disorder - new drugs	Yes
TA067	Flu prevention - amantadine and oseltamivir	Oseltamivir, post exposure: Yes (major restrictions) Prophylaxis (both), and amantadine post exposure: No
TA068	Macular degeneration (age related) - photodynamic therapy	No
TA069	Cervical cancer - cervical screening (review)	Yes
TA070	Leukaemia (chronic myeloid) - imatinib	Yes
TA071	Ischaemic heart disease - coronary artery stents	Bare-metal stent: Yes Drug eluting stent: Yes (major restrictions)
TA072	Rheumatoid arthritis - anakinra	No
TA073	Angina and myocardial infarction - myocardial perfusion scintigraphy	Yes (major restrictions)
TA074	Trauma - fluid replacement therapy	Yes (major restrictions)
TA075	Hepatitis C - pegylated interferons, ribavarin and alfa interferon	Yes (major restrictions)
TA076	Epilepsy (adults) - newer drugs	Yes (major restrictions)
TA077	Insomnia - newer hypnotic drugs	Yes (major restrictions)
TA078	Menstrual bleeding - fluid-filled thermal balloon and microwave endometrial ablation	Yes (minor restrictions)
TA079	Epilepsy (children) - newer drugs	Yes (minor restrictions)
TA080	Acute coronary syndromes - clopidogrel	Yes (major restrictions)
TA081	Atopic eczema - Topical corticosteroids (frequency of use)	Yes (minor restrictions)
TA082	Atopic eczema - pimecrolimus and tacrolimus	Tacrolimus: Yes (minor restrictions) Pimecrolimus: Yes (major restrictions)
TA083	Hernia - Laparascopic surgery (review)	Yes (minor restrictions)
TA084	Severe sepsis - Drotrecogin alfa (activated)	Yes (major restrictions)
TA085	Renal transplantation (adults) - immuno-suppressive regimens	Yes / Yes (major restrictions); (across various drugs)
TA086	Gastro-intestinal stromal tumours (GIST) - Imatinib	Yes (major restrictions)
TA087	Osteoporosis - Secondary prevention	Yes (major restrictions)
TA088	Bradycardia - dual chamber pacemakers	Yes (major restrictions)
TA089	Cartilage injury - ACI (review)	No
TA090	Vascular disease - clopidogrel and dipryidamole	Yes (major restrictions)
TA091	Ovarian cancer (advanced) - paclitaxel, pegylated liposomal doxorubicin hydrochloride and topotecan (review)	Yes (major restrictions)
TA092	Tooth decay - HealOzone	No
TA093	Colorectal cancer (advanced) - irinotecan, oxaliplatin and raltitrexed (review)	Irinotecan, oxaliplatin: Yes (minor restrictions) Raltitrexed: No
TA094	Cardiovascular disease - statins	Yes (in adults with clinical evidence of CVD)
TA095	Arrhythmia - ICDs (review)	Yes (major restrictions)
TA096	Hepatitis B (chronic) - adefovir dipivoxil and pegylated interferon alpha-2a	Yes (minor restrictions)
TA097	Depression and anxiety - computerised cognitive behavioural therapy (CCBT)	Yes (major restrictions)
TA098	Attention deficit hyperactivity disorder (ADHD) - methylphenidate, atomoxetine and dexamfetamine (review)	Yes
TA099	Renal transplant - immunosuppresive regimens for children and adolescents	Bbasiliximab, daclizumab, tacrolimus, mycophenolate mofetil: Yes (minor restrictions) Mycophenolate sodium, sirolimus: No
TA100	Colon cancer (adjuvant) - capecitabine and oxaliplatin	Yes (minor restrictions)
TA101	Prostate cancer (hormone-refractory) - docetaxel	Yes
TA102	Conduct disorder in children - parent training/education programmes	Group-based: Yes Individual training: Yes (major restrictions)
TA103	Psoriasis - efalizumab and etanercept	Yes (minor restrictions)
TA104	Psoriatic arthritis - etanercept and infliximab	Yes (minor restrictions)
TA105	Colorectal cancer - laparoscopic surgery (review)	Yes
TA106	Hepatitis C - peginterferon alfa and ribavirin	Yes (minor restrictions)
TA107	Breast cancer (early) - trastuzumab	Yes (HER2 Positive)

BIBLIOGRAPHY

Abellan-Perpinan JM, Pinto-Prades JL. Health state after treatment: a reason for discrimination? *Health Economics* 1999 Dec;8(8):701-707.

Acheson D. *Independent inquiry into inequalities in health.* London: Stationery Office; 1998.

Adamowicz W. Stated preference methods: random utility models, econometric analysis and welfare measurement. Paper presented at 'Advancing the methodology of discrete choice experiments in health economics workshop'. Oxford, University of Oxford, 23-24 Setptember; 2003.

Andersson F, Lyttkens CH. Preferences for equity in health behind a veil of ignorance. *Health Economics* 1999 Aug;8(5):369-78.

Arneson R. Equality and equality of opportunity for welfare. *Philosophical Studies* 1989;56.

Bailey JW. *Utilitarianism, Institutions, and Justice.* New York: Oxford University Press; 1997.

Baltussen R, Stolk E, Chisholm D, Aikins M. Towards a multi-criteria approach for priority setting: an application to Ghana. *Health Economics* 2006;15(7):689-96.

Baron J, Ubel PA. Revising a priority list based on cost-effectiveness: the role of the prominence effect and distorted utility judgments. *Medical Decision Making* 2001; 21(4):278-87.

Barrett A, Roques T, Small M, Sith RD. How much will herceptin really cost? *British Medical Journal* 2006;333:1118-20.

Bateman I, Carson R, Day B, Hanemann M, HAnley N, Hert T, *et al. Evaluating valuation with stated preference techniques: a manual.* Cheltenham: Edward Elgar; 2002.

Battisla RN, Hodge MJ. Setting priorities and selecting topics for clinical practice guidelines. *Canadian Medical Association Journal* 1995;153(9):1233-7.

Bayoumi AM, Hoch J. What's fair? The attributes of equitable health care resource allocation decisions. Abstract of 27th Annual Meeting of the Society of Medical Decision Making (October 21-24). 2005.

Beach LR. *The Psychology of Decision Making.* London: Sage; 1997.

Beach MC, Asch DA, Jepson C, Hershey JC, Mohr T, McMorrow S, *et al.* Public response to cost-quality tradeoffs in clinical decisions. *Medical Decision Making* 2003;23(5):369-78.

Belkar R, Fiebig DG. Use of discrete choice experiments in health economics: An update of the literature. Working Paper. School of Economics, University of New South Wales, Australia; 2004.

Bergson A. A reformulation of certain aspects of welfare theory. *Quarterly Journal of Economics* 1938;68:233-52.

Birch S, Gafni A. On being NICE in the UK: guidelines for technology appraisal for the NHS in England and Wales. *Health Economics* 2002;11:185-91.

Bjork S, Rosen P. Setting Health-Care Priorities in Sweden - The Politicians Point-of-View. *Health Policy* 1993;26(2):141-54.

Bleichrodt H, Doctor J, Stolk E. A nonparametric elicitation of the equity-efficiency trade-off in cost-utility analysis. *Journal of Health Economics* 2005;24(4):655-78.

BMA. *Healthcare Funding Review*. London: British Medical Association; 2001.

Bowling A. Health care rationing: the public's debate. *British Medical Journal* 1996; 16;312(7032):670-4.

Bowling A, Mariotto A, Evans O. Are older people willing to give up their place in the queue for cardiac surgery to a younger person? *Age & Ageing* 2002; 31(3):187-92.

Bradley N, Sweeney K, Waterfield M. The health of their nation: how would citizens develop England's health strategy? *British Journal of General Practice* 1999; 49(447):801-5.

Brock D. Priority to the worst off in health care resource prioritisation. In: Battin M, Rhodes R, Silvers A, editors. *Health Care and Social Justice*. Oxford University Press; New York. 2001.

Broome J. *Weighing Goods*. Oxford: Blackwell; 1991.

Brouwer WBF, Koopmanschap MA, Rutten FFH. Productivity costs measurement through quality of life? A response to the recommendation of the Washington Panel. *Health Economics* 1997; 6(3):253-259.

Browning CJ, Thomas SA. Community values and preferences in transplantation organ allocation decisions. *Social Science and Medicine* 2001;52(6):853-861.

Brown JR. *The Laboratory of the mind: Thought Experiments in the Natural Sciences.* Routledge (London); 1993.

Bryan S, Buxton M, Sheldon R, Grant A. Magnetic resonance imaging for the investigation of knee injuries: an investigation of preferences. *Health Economics* 1998;7(7):595-603.

Bryan S, Gold L, Sheldon R, Buxton M. Preference measurement using conjoint methods: an empirical investigation of reliability. *Health Economics* 2000;9(5):386-95.

Bryan S, Roberts T, Heginbotham C, McCallum A. QALY-maximisation and public preferences: results from a general population survey. Health *Economics* 2002; 11(8):679-93.

Bryan S, Dolan P. Discrete choice experiments in health economics. For better or for worse? *Health Economics* 2004;3:199-202.

Bryan S, Williams I, McIver S. Seeing the NICE side of cost-effectiveness analysis: A qualitative investigation of the use of CEA in NICE technology appraisals. *Health Economics* 2006;16(2):179-93.

Burgess L, Street D. Optimal designs for 2k choice experiments. *Communications in Statistics - Theory and Methods* 2003;32:2185-206.

Burgess L, Street D, Louviere JL. Quick and easy choice sets: Constructing optimal and nearly optimal stated choice experiments. *International Journal of Research in Marketing* 2005;22:459-70.

Burls A, Bradley P. Public Health Ethics. In: *Medical Ethics Today – the practice and philosophy.* British Medical Association. London: British Medical Association 2003.

Busschbach JJ, Hessing DJ, de Charro FT. The utility of health at different stages in life: a quantitative approach. *Social Science and Medicine* 1993;37(2):153-158.

Campbell T. *Justice. Issues in Political Theory.* (2nd Edition). Palgrave Macmillan; 2001.

Charny MC, Lewis PA, Farrow SC. Choosing who shall not be treated in the NHS. *Social Science and Medicine* 1989;28(12).

Choudhry N, Slaughter P, Sykora K, Naylor CD. Distributional dilemmas in health policy: large benefits for a few or smaller benefits for many? *Journal of Health Services Research and Policy* 1997;2(4):212-6.

Coast J. Is economic evaluation in touch with society's health values? *British Medical Journal* 2004;329:1233-6.

Cookson R, Dolan P. Public views on health care rationing: a group discussion study. *Health Policy* 1999;49(1-2):63-74.

Cropper ML, Aydede SK, Portney PR. Preferences for life saving programs: How the public discounts time and age. *Journal of Risk and Uncertainty* 1994;8:243-65.

Culyer AJ. The normative economics of health care finance and provision. *Oxford Review of Economic Policy* 1989;5(1):34-58.

Culyer AJ. Commodities, characteristics of commodities, characteristics of people, utilities, and quality of life. In: Baldwin S, Godfrey C, Propper C, editors. *Quality of Life: Perspectives and Policies*. London: Routledge; 1990.

Culyer AJ, Wagstaff A. Equity and equality in health and health care. *Journal of Health Economics* 1993; 12(4):431-57.

Culyer AJ. Equity - Some theory and its policy implications. *Journal of Medical Ethics* 2001; 27(4):275-83.

Culyer AJ. *The Dictionary of Health Economics*. Cheltenham: Edward Elgar; 2005.

Dakin HA, Devlin NJ, Odeyemi IAO. "Yes,"No" or "Yes, but"? Multinomial modelling of NICE decision-making. *Health Policy* 2006;77:352-67.

Daniels N, Sabin J. The ethics of accountability in managed care reform. *Health Affairs* 1998;17:50-64.

Daniels N. Accountability for reasonableness. *British Medical Journal* 2000;321:1300-1.

Department of Health. *Inequalities in Health: The Black Report.* London: Department of Health; 1980.

Department of Health. *Patients Charter.* London: The Department of Health; 1992.

Department of Health (NHS Management Executive). *Local Voices: The views of local people in purchasing for health.* London: Department of Health; 1992.

Department of Health. *The New NHS: Modern, Dependable*. London: Department of Health (UK); 1997.

Department of Health. *The NHS Plan*. London: Department of Health; 2000.

Department of Health. *Involving patients and the public in healthcare: A discussion document*. London: Department of Health; 2001.

Department of Health. *Tackling health inequalities: A programme for action*. London: Department of Health (UK); 2003.

Department of Health. *Directions to Primary Care Trusts and NHS trusts in England concerning arrangements for the funding of technology appraisal guidance from the National Institute for Clinical Excellence*. www.dh.gov.uk; 2004.

Devlin N, Parkin D. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. *Health Economics* 2004;13:437-52.

Dicker A, Armstrong D. Patients' views of priority setting in health care: An interview survey in one practice. *British Medical Journal* 1995; 311(7013):1137-9.

Dixon A, Le Grand J, Henderson J, Murray R, Poteliakhoff E. *Is the NHS equitable? A review of the evidence*. London: London School of Economics; 2003. Health and Social Care Discussion Paper No.11.

Dolan P. The measurement of individual utility and social welfare. *Journal of Health Economics* 1998;17(1):39-52.

Dolan P, Green C. Using the person trade-off approach to examine differences between individual and social values. *Health Economics* 1998;7(4):307-12.

Dolan P, Cookson R, Ferguson B. Effect of discussion and deliberation on the public's views of priority setting in health care: focus group study. *British Medical Journal* 1999;318(7188):916-9.

Dolan P, Cookson R. A qualitative study of the extent to which health gain matters when choosing between groups of patients. *Health Policy* 2000;51(1):19-30.

Dolan P, Robinson A. The measurement of preferences over the distribution of benefits: The importance of the reference point. *European Economic Review* 2001;45(9):1697-709.

Dolan P, Shaw R. A note on the relative importance that people attach to different factors when setting priorities in health care. *Health Expectations* 2003;6(1):53-9.

Dolan P, Shaw R. A note on a discussion group study of public preferences regarding priorities in the allocation of donor kidneys. *Health Policy* 2004;68(1):31-6.

Dolan P, Shaw R, Tsuchiya A, Williams A. QALY maximisation and people's preferences: A methodological review of the literature. *Health Economics* 2005;14(2):197-208.

Dolan P, Tsuchiya A. Health priorities and public preferences: the relative importance of past health experience and future health prospects. *Journal of Health Economics* 2005;24(4):703-14.

Donaldson C, Gerard K. *Economics of Health Care Financing. The Visible Hand*. Second Edition. Basingstoke: Palgrave Macmillan; 2005.

Dworkin R. What is equality? *Philosophy and Public Affairs* 1981;10:185-246.

Dworkin R. *Sovereign Virtue: The Theory and Practice of Equality*. 2000.

Edwards RT, Boland A, Wilkinson C, Cohen D, Williams J. Clinical and lay preferences for the explicit prioritisation of elective waiting lists: survey evidence from Wales. *Health Policy* 2003;63(3):229-37.

Emmelin MAC, Lindholm LA, Stenlund HC, Dahlgren LG. Pol-ethical considerations in public health. The views of Swedish health care politicians. *European Journal of Public Health* 1999; 9(2):124-30.

EuroQol Group. EuroQol - a new facility for the measurement of health related quality of life. *Health Policy* 1990;16:199-208.

Farrar S, Ryan M, Ross D, Ludbrook A. Using discrete choice modelling in priority setting: an application to clinical service developments. *Social Science and Medicine* 2000;50(1):63-75.

Fiebig DG, Louviere JL, Walsman DM. *Contemporary Issues in Modelling Discrete Choice Experimental Data in Health Economics*. Working Paper. School of Economics, University of New South Wales, Australia; 2005.

Fowler FJ, Jr., Berwick DM, Roman A, Massagli MP. Measuring public priorities for insurable health care. *Medical Care* 1994;32(6):625-39.

Furnham A, Simmons K, McClelland A. Decisions concerning the allocation of scarce medical resources. *Journal of Social Behavior & Personality* 2000;15(2):185-200.

Gafni A, Birch S. NICE methodological guidelines and decision making in the National Health Service in England and Wales. *Pharmacoeconomics* 2003;21(3):149-57.

Gerard K. Cost-utility in practice - A policy makers guide to the state-of-the-art. Health Policy 1992;21(3):249-79.

Gerard K. *Economic aspects of consumer involvement in health care benefit assessment*. Southampton: PhD Thesis, University of Southampton; 2005.

Goddard M, Smith P. Equity of access to health care services: theory and evidence from the UK. *Social Science and Medicine* 2001;53(9):1149-62.

Gravelle H, Rees R. *Microeconomics*. (3rd Edition). Pearson Education; 2004.

Green C, Brazier JE, Deverill M. Valuing health-related quality-of-life: A review of health state valuation techniques. *Pharmacoeconomics* 2000;17(2):151-65.

Green C. On the societal value of health care: what do we know about the person trade-off technique? *Health Economics* 2001;10(3):233-43.

Green MJ, Fong S, Mauger DT, Ubel PA. Rationing HIV medications: what do patients and the public think about allocation policies? *Journal of Acquired Immune Deficiency Syndrome* 2001;26(1):56-62.

Greene WH. *Econometric Analysis*. Fourth Edition. New Jersey: Prentice Hall; 2000.

Greene WH. *Econometric Analysis*. Fifth Edition. New Jersey: Prentice Hall; 2003.

Gyrd-Hansen D, Sogaard J. Analysing public preferences for cancer screening programmes. *Health Economics* 2001;10(7):617-34.

Gyrd-Hansen D, Slothuus U. The citizen's preferences for financing public health care: a Danish survey. *Int J Health Care Finance Econ* 2002;2(1):25-36.

Gyrd-Hansen D. Investigating the social value of health changes. *Journal of Health Economics* 2004; 23(6):1101-16.

Hahn FH. On some difficulties of the utilitarian economist. In: Sen A, Williams B, editors. *Utilitariansim and Beyond*. Cambridge University Press; 1982.

Ham C. *Health Care Reform: Learning from international experience.* Open University Press; 1997.

Ham C, Robert G. *Reasonable rationing. International experience of priority setting in health care*. Maidenhead, UK: Open University Press; 2003.

Hamlin AP. *Ethics and Economics (2 Volumes) (International Library of Critical Writings in Economics)*. Cheltenham: Edward Elgar; 1996.

Hanemann WM. Welfare evaluations in contingent valulation experiments with discrete responses. *American Journal of Agricultural Economics* 1984;66:332-41.

Hanley N, Ryan M, Wright RE. Estimating the monetary value of health care: lessons from environmental economics. *Health Economics* 2003;12(1):3-16.

Harsanyi JC. *Essays in Ethics, Social Behaviour, and Scientific Explanation*. Dordrecht: Reidel; 1976.

Hauck K, Shaw R, Smith PC. Reducing avoidable inequalities in health: a new criterion for setting health care capitation payments. *Health Economics* 2002;11(8):667-77.

Hauck K, Smith PC, Goddard M. *The economics of priority setting for health care. A Literature review*. Washington, USA: Health Nutrition and Population (NHP) Discussion Paper, The World Bank; 2003.

Hausman D, McPherson MS. Taking ethics seriously: Economics and contemporary moral philosophy. *Journal of Economic Literature* 1993;XXXI:671-731.

Hensher DA, Rose JM, Greene WH. *Applied Choice Analysis: A Primer*. Cambridge University Press; 2005.

Hicks J. *Value and Capital*. Oxford: Clarendon Press; 1939.

Holmes AM. A method to elicit utilities for interpersonal comparisons. *Medical Decision Making* 1997;17(1):10-20.

Hope T. Rationing and life-saving treatments: should identifiable patients have higher priority? *Journal of Medical Ethics* 2001;27(3):179-85.

Hurley J. An overview of the normative conomics of the health sector. In: Culyer AJ, Newhouse JP, editors. *Handbook of Health Economics*. Elsevier; 2000.

Jan S, Mooney G, Ryan M, Bruggemann K, Alexander K. The use of conjoint analysis to elicit community preferences in public health research: a case study of hospital services in South Australia. *Australian and New Zealand Journal of Public Health* 2000;24(1):64-70.

Jenkings KN, Barber N. What constitutes evidence in hospital new drug decision making? *Social Science and Medicine* 2004;58:1757-66.

Johannesson M, Gerdtham U. A note on the estimation of the equity-efficiency trade-off for QALYs. *Journal of Health Economics* 1996;15(3):359-368.

Johannesson M, Johansson PO. The economics of ageing: on the attitude of Swedish people to the distribution of health care resources between the young and the old. *Health Policy* 1996;37(3):153-61.

Johannesson M, Johansson PO. Is the valuation of a QALY gained independent of age? Some empirical evidence. *Journal of Health Economics* 1997;16(5):589-599.

Johri M, Ubel PA. Setting organ allocation priorities: Should we care what the public cares about? *Liver Transplantation* 2003; 9(8):878-80.

Johri M, Damschroder LJ, Zikmund Fisher B, Ubel PA. The importance of age in allocating health care resources: does intervention-type matter? *Health Economics* 2005;14(7):669-78.

Kaldor N. Welfare propositions and interpersonal comparison of utility. *Economic Journal* 1939;49:549-52.

Kinnunen J, Lammintakanen J, Myllykangas M, Ryynanen O-P, Takala J. Health care priorities as a problem of local resource allocation. *Int J Health Plann Management* 1998; 13(3):216-29.

Kjaer T. A review of the discrete choice experiment - with emphasis on its application in health care. *Health Economic Papers* 2005;1:2005.

Kluge EH, Tomasson K. Health care resource allocation: complicating ethical factors at the macro-allocation level. *Health Care Anal* 2002;10(2):209-20.

Kneeshaw J. What does the public think about rationing?: a review of the evidence. In: New B, editor. *Rationing: Talking and Action in Health Care*. London: BMJ; 1999.

Kocur G, Alder T, Hyman W. *Guide to forecasting travel demand with direct utility assessment*. Washington DC: US Department of Transportation.

Konow J. Fair shares: Accountability and cognitive dissonance in allocation decisions. American Economic Review 2000;90(4):1072-1091.

Konow J. Fair and square: The four sides of distributive justice. *Journal of Economic Behaviour and Organisationl* 2001;46(2):137-164.

Konow J. Which is the fairest one of all? A positive analysis of theories of justice. *Journal of Economic Literature* 2003; XLI:1188-1239.

Kuder LB, Roeder PW. Attitudes Toward Age-Based Health-Care Rationing - A Qualitative Assessment. *Journal of Aging and Health* 1995;7(2):301-27.

Lancaster K. A new approach to consumer theory. *Journal of Political Economy* 1966;74:132-57.

Lancaster KJ. *Consumer Demand: A New Approach*. New York: Columbia University Press; 1971.

Lanscar E, Louviere J. Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences? *Health Economics* 2006;15(8):797-811.

Lees A, Scott N, Scott SN, MacDonald S, Campbell C. Deciding how NHS money is spent: a survey of general public and medical views. *Health Expectations* 2001; 5(2):47-54.

Lewis PA, Charny M. Which of two individuals do you treat when only their ages are different and you can't treat both? *Journal of Medical Ethics* 1989;15(1):28-32.

Lian OS. Norms for priority setting among health professionals: a view from Norway. *Sociology of Health & Illness* 2001;23(3):357-85.

Lindblad AK, Hartzema AG, Jansson L, Feltelius N. Patients' views of priority setting for new medicines. A qualitative study of patients with rheumatoid arthritis. *Scand J Rheumatol* 2002;31(6):324-9.

Lindholm L, Rosen M. On the measurement of the nation's equity adjusted health. *Health Economics* 1998;7(7):621-628.

Lindholm LA, Emmelin MA, Rosen ME. Health maximization rejected - The view of Swedish politicians. *European Journal of Public Health* 1997;7(4):405-10.

Little IMD. Ethics, *Economics and Politics. Principles of Public Policy*. Oxford: Oxford University Press; 2002.

Louviere JL. Conjoint analysis modelling of stated preferences. *Journal of Transport Economics and Policy* 1988;22(1):93-119.

Louviere.J., Hensher D, Swait J. *Stated Choice Methods: Analysis and Applications*. Cambridge: Cambridge University Press; 2000.

Luce RD, Suppes P. Preference Utility and Subjective Probability. In: Luce RD, Bush RR, Galanter E, editors. *Handbook of Mathematical Psychology*. 1965. p. 249-410.

Mariotto A, De LD, Buono MD, Favaretti C, Austin P, Naylor CD. Will elderly patients stand aside for younger patients in the queue for cardiac services? *Lancet* 1999:7;354(9177):467-70.

Mark T, Swait J. Using stated preference and revealed preference modeling to evaluate prescribing decisions. *Health Economics* 2004;13(6):563-73.

Martin D, Abelson J, Singer P. Participation in health care priority-setting through the eyes of the participants. *Journal of Health Services Research and Policy* 2002;7(4):222-9.

McFadden D. Conditional logit analysis of qualitative choice behaviour. In: Zarembka P, editor. *Frontiers in Economics*. New York: Academic Press; 1973.

McGuire A. Theoretical concepts in economic evaluation of health care. In: Drummond M, McGuire A, editors. *Economic Evaluation in Health Care. Merging Theory with Practice*. Oxford: Oxford University Press; 2001.

McIntosh E. Using discrete choice experiments within a cost-benefit analysis framework. Some considerations. *Pharmacoeconomics* 2006;24(9):855-68.

Mill JS. *Utilitarianism*. Penguin Classics (republished 1987); 1863.

Mirrlees JA. The economic uses of utilitarianism. In: Sen A, Williams B, editors. *Utilitarianism and Beyond*. Cambridge: Cambridge University Press; 1982.

Mooney G, Jan S, Wiseman V. Examining preferences for allocating health care gains. *Health Care Analysis* 1995;3(3):261-265.

Mooney G. "Communitarian claims" as an ethical basis for allocating health care resources. *Social Science and Medicine* 1998;47(9):1171-80.

Mooney G. Vertical equity in health care resource allocation. *Health Care Analysis* 2000;8(3):203-15.

Mullen MA, Kohut N, Sam M, Blendis L, Singer PA. *Access to adult liver transplantation in Canada: a survey and ethical analysis*. Canadian Medical Association Journal 1996;154(3):337-42.

Murray JL, Lopez AD. *The global burden of disease: A comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020*. Cambridge: Harvard University Press; 1996.

Murray JL, Lopez AD. *Alternative projections of mortality and disability by cause 1990-2020: Blobal Burden of Disease Study*. Lancet 1997;349:1498-504.

Myllykangas M, Ryynanen OP, Lammintakanen J, Isomaki VP, Kinnunen J, Halonen P. Clinical management and prioritization criteria. Finnish experiences. *J Health Organ Manag* 2003;17(5):338-48.

Nath SK. *A Perspective of Welfare Economics*. London: Macmillan; 1973.

National Institute for Clinical Excellence. *Framework Document*. London: NICE; 2000.

National Institute for Clinical Excellence. *Guidance on the use of riluzole (Rilutek) for the treatment of motor neurone disease*. London: National Institute for Clinical Excellence; 2001. Guidance No.20.

National Institute for Clinical Excellence. *Beta interferon and glatarimer acetate for the treatment of multiple sclerosis*. London: National Institute for Clinical Excellence; 2002. Guidance No.32.

National Institute for Clinical Excellence. *Report of the Citizens Council: Determining Clinical Need*. London: National Institute for Clinical Excellence (www.nice.org.uk); 2003.

National Institute for Clinical Excellence. *Anakinra for rheumatoid arthritis*. London: National Institute for Clinical Excellence; 2003. Technology Appraisal Guidance No.72.

National Institute for Clinical Excellence. *Guidance on the use of imatinab for chronic myeloid leukaemia*. London: National Institute for Health and Clinical Excellence; 2003. Technology Appraisal Guidance No: TA070.

National Institute for Clinical Excellence. *Guidance on the use of continuous subcutaneous insulin infusion for diabetes*. London: National Institute for Health and Clinical Excellence; 2003. Technology Appraisal Guidance: TA057.

National Institute for Clinical Excellence. *Guidance on the use of photodynamic therapy for age-related macular degeneration*. London: National Institute for Health and Clinical Excellence; 2003. Technology Appraisal Guidance: TA068.

National Institute for Clinical Excellence. *NICE Citizens Council Report on Age*. London: National Institute for Clinical Excellence (www.nice.org.uk); 2004.

National Institute for Clinical Excellence. *Guide to the Methods of Technology Appraisal*. London: National Institute for Clinical Excellence; 2004 (April).

National Institute for Clinical Excellence. *Drotrecogin alfa (activated) for severe sepsis*. London: National Institute for Health and Clinical Excellence; 2004. Technology Appraisal Guidance: TA084.

National Institute for Health and Clinical Excellence. *Social value judgements. Principles for the development of NICE guidance*. Consultation paper (April). London: National Institute for Health and Clinical Excellence. (April) 2005

National Institute for Health and Clinical Excellence. *Social value judgements. Principles for the development of NICE guidance*. London: National Institute for Health and Clinical Excellence (www.nice.org.uk). (8th December) 2005

National Institute for Health and Clinical Excellence. *Donepezil, galantamine, rivastigmine (review) and memantine for the treatment of Alzheimer's disease*. London: National Institute for Health and Clinical Excellence; 2006. Technology Appraisal Guidance: TA111.

National Institute for Health and Clinical Excellence. *Bortezomib monotherapy for relapsed multiple myeloma: Final Appraisal Determination*. London: National Institute for Health and Clinical Excellence (October); 2006.

National Institute for Health and Clinical Excellence. *Trastuzumab for the adjuvant treatment of early stage HER2-positive breast cancer*. London: National Institute for Health and Clinical Excellence; 2006. Technology Appraisal Guidance: TA107.

Neuberger J, Adams D, MacMaster P, Maidment A, Speed M. Assessing priorities for allocation of donor liver grafts: survey of public and clinicians. *British Medical Journal* 1998;317(7152):172-5.

Neumann PJ. *Using cost effectiveness analysis to improve health care*. Opportunities and barriers. Oxford: Oxford University Press; 2005.

New Zealand Ministry of Health. *New Zealand Health Strategy 2000*. New Zealand Ministry of Health (www.moh.govt.nz); 2000.

New B. Rationing. *Talk and Action in Health Care*. London: BMJ Publishing Group; 1997.

New B. The rationing agenda in the NHS. *British Medical Journal* 1996: 22;312(7046):-1601.

Newdick C. *Who should we treat? Rights, rationing, and resources in the NHS*. 2nd ed. Oxford, New York: Oxford University Press; 2005.

Ng Y-K. *Welfare economics: introduction and development of basic concepts*. Revised Edition. Basingstoke: Macmillan Press; 1983.

Nord E. Norwegian health care politicians do not wish to maximize health. *Journal of the Norwegian Medical Association* 1993;113:1371-1373 (Text in Norwegian).

Nord E. The trade-off between severity of illness and treatment effect in cost- value analysis of health care. *Health Policy* 1993;24(3):227-38.

Nord E. The relevance of health state after treatment in prioritizing between different patients. *Journal of Medical Ethics* 1993;19(1):37-42.

Nord E, Richardson J, arounas-Kirchmann K. Social evaluation of health care versus personal evaluation of health states. Evidence on the validity of four health-state scaling instruments using Norwegian and Australian surveys. *Int J Technol Assess Health Care* 1993;9(4):463-78.

Nord E. The person-trade-off approach to valuing health care programs. *Medical Decision Making* 1995; 15(3):201-8.

Nord E, Richardson J, Street A, Kuhse H, Singer P. Maximizing health benefits vs egalitarianism: An Australian survey of health issues. *Social Science and Medicine* 1995a; 41(10):1429-37.

Nord E, Richardson J, Street A, Kuhse H, Singer P. Who cares about cost - Does economic-analysis impose or reflect social values. *Health Policy* 1995;34(2):79-94.

Nord E, Street A, Richardson J, Kuhse H, Singer P. The significance of age and duration of effect in social evaluation of health care. *Health Care Analysis* 1996;4(2):103-11.

Nord E, Pinto JL, Richardson J, Menzel P, Ubel P. Incorporating societal concerns for fairness in numerical valuations of health programmes. *Health Economics* 1999;8(1):25-39.

Nord E. Concerns for the worse off: Fair innings versus severity. *Social Science and Medicine* 2005; 60(2):257-63.

Nozick R. *Anarchy, State and Utopia*. Oxford: Blackwell; 1974.

Oddsson K. Assessing attitude towards prioritizing in healthcare in Iceland. *Health Policy* 2003;66(2):135-46.

Office for National Statistics. *Disadvantaged Households. Results from the 2000 General Household Survey*. London: National Statistics; 2002.

Office for National Statistics. *Focus on Social Inequalities*. London: National Statistics; 2004.

Oliver A. Prioritizing health care: is "health" always an appropriate maximand? *Medical Decision Making* 2004;24(3):272-80.

Oliver A, Mossialos E, Robinson R. Health technology assessment and its influence on health-care priority setting. *Int J Technol Assess Health Care* 2004; 20(1):1-10.

Olsen JA. Persons vs years: two ways of eliciting implicit weights. *Health Economics* 1994;3(1):39-46.

Olsen JA. Theories of justice and their implications for priority setting in health care. *Journal of Health Economics* 1997;16(6):625-39.

Osberg L. The Equity Efficiency Trade-off in Retrospect. *Canadian Business Economics* 1995 (Spring).

Parfit D. Equality or priority. The Lindley Lecture. Department of Philosophy, University of Kansas; 1991.

Pearmain D, Swanson J, Kroes E, Bradley M. *Stated preference techniques. A guide to practice*. Richmond, Surrey: The Haag, Netherlands: Steer, Davis, Gleave and Hague Consulting Group; 1991.

Perneger TV, Martin DP, Bovier PA. Physicians' attitudes toward health care rationing. *Medical Decision Making* 2002;22(1):65-70.

Pickard S, Sheaf R. Primary care groups and NHS rationing: Implications of the Child B case. *Health Care Analysis* 1999;7(1):37-56.

Raftery J. NICE: faster access to modern treatments? Analysis of guidance on health technologies. *British Medical Journal* 2001;323:1300-3.

Raine R, Hutchings A, Black N. Is publicly funded health care really distributed according to need? The example of cardiac rehabilitation in the UK. *Health Policy* 2003;63(1):63-72.

Ratcliffe J. Public preferences for the allocation of donor liver grafts for transplantation. *Health Economics* 2000;9(2):137-48.

Rawlins MD, Culyer AJ. National Institute for Clinical Excellence and its value judgements. *British Medical Journal* 2004;329:224-6.

Rawls J. *A Theory of Justice*. New York: Oxford University Press; 1971.

Rawls J. *Political Liberalism*. Cambridge University Press; 1993.

Rawls J. *A Theory of Justice*. Revised Edition. New York: Oxford University Press; 1999.

Reinhardt UE. Reflections on the meaning of efficiency: can efficiency be separated from equity. *Yale Law Policy* Review 1992;10:302-315.

Reinhardt UE. Abstracting from distributional effects, this policy is inefficient. In: Barer MGT, Stoddart G, editors. *Health, Health Care and Health Economics: Perspectives on Distribution*. Toronto: John Wiley & Sons; 1998.

Roberts T, Bryan S, Heginbotham C, McCallum A. Public involvement in health care priority setting: an economic perspective. *Health Expectations* 1999;2(4):235-44.

Rodriguez-Miguez E, Pinto-Prades JL. Measuring the social importance of concentration or dispersion of individual health benefits. *Health Economics* 2002;11(1):43-53.

Rodriguez E, Pinto JL. The social value of health programmes: is age a relevant factor? *Health Economics* 2000;9(7):611-621.

Rosen P, Karlberg I. Opinions of Swedish citizens, health-care politicians, administrators and doctors on rationing and health-care financing. *Health Expectations* 2002;5(2):148-55.

Rosen S. Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy* 1974;(82):34-55.

Ryan M. *Using consumer preferences in health care decision making: the application of conjoint analysis*. London: Office of Health Economics; 1996.

Ryan M. Using conjoint analysis to take account of patient preferences and go beyond health outcomes: an application to in vitro fertilisation. *Social Science and Medicine* 1999;48(4):535-46.

Ryan M. Discrete choice experiments in health care. *British Medical Journal* 2004;328:360-1.

Ryan M, Gerard K. Using discrete choice experiments to value health care programmes: current practice and future research reflections. *Applied Health Economics and Health Policy* 2003;2(1):1-10.

Ryan M, Scott DA, Reeves C, Bate A, Russell EM. Eliciting public preferences for healthcare: a systematic review of techniques. *Health Technology Assessment* 2001;5(5).

Ryan M, Watson V, Amaya-Amaya,M. Methodological issues in monetary valuation of benefits in health care. *Experet Review in Pharmacoeconomics Outcomes Research* 2003;3(6):717-727.

Ryynanen OP, Myllykangas M, Kinnunen J, Takala J. Attitudes to health care prioritisation methods and criteria among nurses, doctors, politicians and the general public. *Social Science and Medicine* 1999;49(11):1529-39.

Sassi F, Archard L, Le GJ. Equity and the economic evaluation of healthcare. *Health Technology Assessment* 2001; 5(3).

Sassi F, Le GJ, Archard L. Equity versus efficiency: A dilemma for the NHS. *British Medical Journal* 2001;323(7316):762-3.

Sassi F. Setting priorities for the evaluation of health interventions: when theory does not meet practice. *Health Policy* 2003;63(2):141-54.

Schwappach DL. Resource allocation, social values and the QALY: a review of the debate and empirical evidence. *Health Expectations* 2002;5(3):210-22.

Schwappach DL. The equivalence of numbers: the social value of avoiding health decline: an experimental Web-based study. *BMC Medical Information and Decision Making* 2002;2(1):3.

Schwappach DLB. Does it matter who you are or what you gain? An experimental study of preferences for resource allocation. Health Economics 2003;12(4):255-67.

Schwappach DL. Are preferences for equality a matter of perspective? *Medical Decision Making* 2005;25(4):449-59.

Schwappach DL, Strasmann TJ. "Quick and dirty numbers"? The reliability of a stated-preference technique for the measurement of preferences for resource allocation. *Journal of Health Economics* 2006;25(3):432-48.

Scott A, Watson MS, Ross S. Eliciting preferences of the community for out of hours care provided by general practitioners: a stated preference discrete choice experiment. *Social Science and Medicine* 2003;56:803-814.

Scott A, Ubach C, French F, Needham G. Examining the preferences of health care providers: and application to hospital consultants. In: *Using Discrete Choice Experiments to Value Health and Health Care*. Eds. Ryan M, Gerard K. 2006: Kluwer.

Sen AK. Personal utilities and public judgments: or What's wrong with welfare economics. *Economic Journal* 1979;89:537-58.

Sen AK. *Choice, Welfare and Measurement*. Oxford: Blackwell; 1982.

Sen A. *Commodities and Capabilities*. Amsterdam: North-Holland; 1985.

Sen A, Williams B. *Utilitarianism and Beyond*. Cambridge: Cambridge University Press; 1982.

Sen AK. *Inequality Reexamined*. NewYork: Oxford University Press; 1992.

Sen A. Why health equity? *Health Economics* 2002;11(8):659-66.

Shmueli A. Survival vs. quality of life: a study of the Israeli public priorities in medical care. *Social Science and Medicine* 1999;49(3):297-302.

Sheldon TA, Cullum N, Dawson D, Lankshear A, Lowson K, Watt I, West P, Wright D, Wright J. What's the evidence that NICE guidance has been implemented? Results from a national evaluation using time series analysis, audit of patients' notes, and interviews. *British Medical Journal* 2004;329:999.

Shmueli A. Horizontal equity in medical care: A study of the Israeli public's views. *Israeli Medical Association Journal* 2000; 2(10):746-52.

Sidgwick H. *Essays on Ethics and Method*. Oxford: Oxford University Press (Published update 2000); 1874.

Skitka LJ, Tetlock PE. Allocating scarce resources: a contingency model of ditributive justice. *Journal of Experimental Psychology* 1992;28:491-522.

Skjoldborg U, Gyrd-Hansen D. Conjoint analysis: the cost variable an Achilles heel? *Health Economics* 2003;12:479-91.

Sloan N. A Library of Orthogonal Arrays. http://www research att com/~njas/oadir/ (last accessed January 2006) 2003 Available from: URL: www.research.att.com/~njas/oadir

Smith PC, Ginnelley L, Sculpher M. *Health Policy and Economics: Opportunities and Challenges (State of Health)*. Open University Press; 2004.

Sorensen RA. *Thought Experiments*. Oxford University Press (Oxford); 1992.

Southampton City Council. *Southampton. The 2001 Census Results*. Southampton City Council (July), www.southampton.gov.uk; 2004.

Starfield B, Paganini JM. Is equity a scientific issue? *Journal of Epidemiology and Community Health* 2000;54(5):324-5.

STATA Corporation. *Stata Reference Manual, release 8*. College Station, TX: STATA Corporation; 2003.

Stolk EA, Pickee SJ, Ament AH, Busschbach JJ. Equity in health care prioritisation: an empirical inquiry into social value. *Health Policy* 2005;74(3):343-55.

Street D, Burgess L. *Quick and easy choice sets: using word processing packages to construct near-optimal stated choice experiments*. Sydney: University of Technology; 2003.

Street D, Burgess L. Optimal and near optimal pairs for the estimation of effects in the 2-level choice experiment. *Journal of Statistical Planning and Inference* 2004;118:185-99.

Stronks K, Strijbis AM, Wendte JF, Gunning-Schepers LJ. Who should decide? Qualitative analysis of panel data from public, patients, healthcare professionals, and insurers on priorities in health care. *British Medical Journal* 1997;315(7100):92-6.

Sugden R, Williams A. *The Principles of Practical Cost-Benefit Analysis*. Oxford: Oxford University Press; 1978.

Swait J, Adamowicz W. Choice environment, market complexity, and consumer behaviour: a theoretical and empirical approach for incorporating decision complexity into models of consumer choice. *Organisational Behaviour and Human Decision Processes* 2001;86(2):141-67.

Tappnden P, Brazier J, Ratcliffe J. *Does the National Institute for Health and Clinical Excellence take account of factors such as uncertainty and equity as well as incremental cost-effectiveness in commissioning health care servicers? A binary choice experiment.* Sheffield (UK): School of Health and Related Research, University of Sheffield; 2006. Discussion Paper 06/10 (HEDS).

The Stationery Office. *Health and Social Care Act*. London: The Stationery Office; 2002.

The Stationery Office. *National Health Service Reform and Health Care Professions Act*. London: The Stationery Office; 2002.

Tsuchiya A, Williams A. Welfare economics and economic evaluation. In: Drummond M, McGuire A, editors. *Economic Evaluation in Health Care. Merging Theory with Practice*. Oxford: Oxford University Press; 2001.

Tsuchiya A, Dolan P, Shaw R. Measuring people's preferences regarding ageism in health: some methodological issues and some fresh evidence. *Social Science and Medicine* 2003;57(4):-96, 2003.

Tsuchiya A, Miguel LS, Edlin R, Wailoo A, Dolan P. Procedural justice in public health care resource allocation. *Applied Health Economics and Health Policy* 2005;4(2):119-27.

Ubel PA, Loewenstein G. The Efficacy and Equity of Retransplantation - An Experimental Survey of Public-Attitudes. *Health Policy* 1995;34(2):145-51.

Ubel PA, Dekay ML, Baron J, Asch DA. Cost-effectiveness analysis in a setting of budget constraints - Is it equitable? *New England Journal of Medicine* 1996;334(18):1174-7.

Ubel PA, Loewenstein G. Distributing scarce livers: The moral reasoning of the general public. *Social Science and Medicine* 1996;42(7):1049-55.

Ubel PA, Loewenstein G. Public perceptions of the importance of prognosis in allocating transplantable livers to children. *Medical Decision Making* 1996;16(3):234-41.

Ubel PA, Spranca MD, Dekay ML, Hershey JC, Asch DA. Public preferences for prevention versus cure: What if an ounce of prevention is worth only an ounce of cure? *Medical Decision Making* 1998;18(2):141-8.

Ubel PA. How stable are people's preferences for giving priority to severely ill patients? *Social Science and Medicine* 1999;49(7):895-903.

Ubel PA, Richardson J, Prades JL. Life-saving treatments and disabilities. Are all QALYs created equal? *Int J Technol Assess Health Care* 1999;15(4):738-48.

Ubel PA, Baron J, Nash B, Asch DA. Are preferences for equity over efficiency in health care allocation "all or nothing"? *Medical Care* 2000;38(4):366-73.

Ubel PA, Jepson C, Baron J, Mohr T, McMorrow S, Asch DA. Allocation of transplantable organs: Do people want to punish patients for causing their illness? *Liver Transplantation* 2001;7(7):600-7.

Ubel PA, Baron J, Asch DA. Preference for equity as a framing effect. *Medical Decision Making* 2001;21(3):180-9.

UK Government. Statutory Instrument 1999 No. 220. The National Institute for Clinical Excellence (establishment and constitution) order 1999.

Varekamp I, Krol LJ, Danse JA. Age rationing for renal transplantation? The role of age in decisions regarding scarce life extending medical resources. *Social Science and Medicine* 1998;47(1):113-120.

Viney R, Lanscar E, Louviere J. Discrete choice experiments to measure preferences for health and health care. *Expert Review of Pharmacoeconomics Outcomes Research* 2002;2:319-26.

Viney R, Savage E, Louviere JL. Empirical investigation of experimental design properties of discrete choice experiments in health care. *Health Economics* 2005;14:349-62.

Wagstaff A. QALYs and the equity-efficiency trade-off. *Journal of Health Economics* 1991 May;10(1):21-41.

Wailoo A, Anand P. The nature of procedural preferences for health-care rationing decisions. *Social Science and Medicine* 2005;60(2):223-36.

Webster C. *The National Health Service-A Political History*. Oxford: Oxford University Press; 2002.

Weinstein M, Pliskin JS. Perspectives on Health-years Equivalents: HYEs what are the issues? *Medical Decision Making* 1996;16:205-6.

Weinstein MC, Siegel JE, Garber AM, Lipscomb J, Luce BR, Manning WG, Torrance GW. Productivity costs, time costs and health-related quality of life: a response to the Erasmus Group. *Health Economics* 1997;6(5):505-510.

Williams A. Ethics and efficiency in the provision of health care. In: Bell M, Mendus S, editors. *Philosophy and Medical Welfare*. Cambridge: Cambridge University Press; 1988.

Williams A. QALYs and ethics: a health economist's perspective. *Social Science and Medicine* 1996;43(12):1795-804.

Williams A. Intergenerational equity: an exploration of the 'fair innings' argument. *Health Economics* 1997;6(2):117-32.

Williams A. Economics, ethics and the public in health care policy. *International Social Science Journal* 1999;(161):297-312.

Williams I, Bryan S, McIver S. The use of cost effectiveness analysis in technology coverage decisions: evidence from the UK. *Society for Medical Decision Making, Annual Meeting* (San Francisco), (October) 2005.

Williams I, Bryan S. Understanding the limited impact of economic evaluation in health care resource allocation: A conceptual framework. *Health Policy* 2007;80(1):135-143.

Williams I, Bryan S, McIver S. How should cost-effectiveness analysis be used in health technology coverage decisions? Evidence from the National Institute for Health and Clinical Excellence approach. *Journal of Health Services Research and Policy* 2007;12(2):73-9.

Wilmot S, Ratcliffe J, Allen C. How well do members of the public deal with a distributive justice problem in health care? *Journal of Health Services Research and Policy* 2004;9(1):7-13.

Wiseman V, Mooney G, Berry G, Tang KC. Involving the general public in priority setting: Experiences from Australia. *Social Science and Medicine* 2003;56(5):1001-12.

Wiseman V. Comparing the preferences of health professionals and members of the public for setting health care priorities: experiences from Australia. *Applied Health Economics and Health Policy* 2005;4(2):129-37.

Woolhead GM, Donovan JL, Chard JA, Dieppe PA. Who should have priority for a knee joint replacement? *Rheumatology* 2002;41(4):390-4.

World Health Organisation. *World Health Report 1999 - Making a Difference*. Geneva: World Health Organisation; 1999.

World Health Organisation (WHO). *The World Health Report 2000: Health systems: improving performance*. Geneva: World Health Organisation (WHO); 2000.

World Health Organisation. *World Health Report 2002. Reducing Risks, Promoting Healthy Lives*. Geneva: World Health Organisation; 2002.

Young HP. *Equity. In Theory and Practice*. New Jersey & Chichester: Princeton University Press; 1994.

Zweibel NR, Cassel CK, Karrison T. Public attitudes about the use of chronological age as a criterion for allocating health care resources. *Gerontologist* 1993;33(1):74-80.