DOCTOR OF PHILOSOPHY

Reimbursement decisions for new medicines: exploration of the preferences of decision-makers and the public

Linley, Warren G

Award date:
2013

Awarding institution:
Bangor University

Reimbursement decisions for new medicines: exploration of the preferences of decision-makers and the public

PhD Thesis

Warren G. Linley

Centre for Health Economics and Medicines Evaluation,
Institute of Medical and Social Care Research,
Bangor University

2013

Contents

Page

Tables

Appendix Tables

Thesis Summary

Aims: This thesis explores the preferences of different stakeholders, particularly national-level decision-makers and the public, to determine the relevance of current and proposed criteria and processes for medicines reimbursement in the UK National Health Service.

Methods and Results: Based on a systematic literature review of health-related conjoint analyses (Chapter 2), process-related aspects of health care are demonstrated to be important to stakeholders, albeit less so than health outcomes. Using revealed preference methods (Chapter 3) and a stated preference discrete choice experiment (Chapter 4), members of a medicines appraisal committee are prepared to trade-off cost effectiveness and health gains against other factors when making national decisions on new medicines. The first comprehensive empirical analysis of public preferences towards UK medicines prioritisation criteria (Chapter 5), demonstrates that several current criteria (e.g. the end-of-life premium, the special funding status for treatments of rare diseases, the Cancer Drugs Fund in England) do not reflect public preferences for resource allocation, but support is evident for the proposed criteria for rewarding new medicines with higher prices under the value-based pricing system commencing in 2014. From a comparative review of its reports, there is a degree of alignment between the views of the Citizens Council of the National Institute for Health and Clinical Excellence (NICE) and the wider public (Chapter 6), but evidence that it has influenced NICE decision-making processes is lacking.

Conclusions: Using a range of methods, this thesis confirms that current medicines reimbursement processes are inadequate, and moves towards value-based pricing of medicines are supported. Non-health, process-related aspects of health care should be explicitly considered in decision-making. The involvement of patients and public as stakeholders in medicines decision-making at all levels is supported. Efforts to demonstrate fair decision-making processes do not obviate the need for incorporation of relevant social value judgements in decision-making.

Acknowledgements

Many thanks are owed to many people who have made this work possible. First of all I wish to thank my supervisor, mentor, and friend, Prof. Dyfrig Hughes, whose commitment and enthusiasm for his subject is inspiring, and whose patience is second to none. Thank you also to Prof. Rhiannon Tudor-Edwards, my secondary supervisor, who has been there to call upon when needed. I am grateful to all colleagues and fellow students at the Centre for Health Economics and Medicines Evaluation, who have helped with various aspects of piloting of studies and have always been keen to provided suggestions and advice when requested, particularly Miss Emily Fargher. And of course, funding for this research was provided by an Institute of Medical and Social Care Research PhD studentship, for which I am very grateful.

Two of the Chapters of this thesis have explored decision-making by the All Wales Medicines Strategy Group (AWMSG), which would not have been possible without the co-operation of its Chair, Prof. Philip Routledge, and its secretariat, the All Wales Therapeutics and Toxicology Centre (AWTTC). I would like to acknowledge the assistance of Mrs Karen Samuels, Mrs Ruth Lang, and the administration team at AWTTC, who supported my efforts to engage AWMSG members and pharmaceutical companies in these studies. Special thanks are owed to the AWTTC assessment team, and the New Medicines Group and AWMSG appraisal committee members who kindly gave their time to complete my surveys.

I wish to acknowledge the informative discussions with Dr Stephen Lane, Mr Joshua Pink, and Dr Catrin Tudor-Smith, and the anonymous referees who provided constructive comments on three publications arising from this thesis so far. And many thanks to Prof. Bob Woods and the thesis examiners, Prof. Michael Drummond and Dr Nefyn Williams.

I have been incredibly fortunate to have had the love, support and encouragement of my parents in my every effort, academic or otherwise, throughout life. It is with much sadness that my mother passed away in October 2012. Having heard me talk of this work for almost four years, she would have loved to see this thesis reach this point, and I would dearly have loved to see her pretend to read it! Mum, I miss you and think of you every day.

Lastly, I thank my beautiful children, Rosie and Harry, and most of all, my beautiful wife, Rachel, who has lived, breathed and suffered this work every bit as much as I have. Your eternal support, encouragement and faith are truly wonderful, and to the extent that it is possible to dedicate a PhD thesis to anyone, this is for you.

Thesis Structure and Publications

The Institute of Medical and Social Care Research at Bangor University subscribes to a publication-based PhD model, which requires the preparation of four or more publishable manuscripts. This approach has a number of advantages, including early exposure of the research student to the processes and requirements of peer-reviewed publication.

Chapter 1 provides a general overview and the background necessary to the thesis. Each of Chapters 2 to 6 is presented as a stand-alone manuscript, which inevitably involves a degree of overlap and repetition of common themes. Chapter 7 then provides an integrated discussion of the findings, outlines the strengths and limitations of the approaches used, suggests areas for future research and concludes the thesis.

At the time of writing, Chapters 2 and 6 are being prepared for submission for publication, and Chapters 3, 4 and 5 have been published as:

- Linley WG, Hughes DA. Cost-effectiveness, clinical factors and policy influences on reimbursement decisions of the All Wales Medicines Strategy Group. *Pharmacoeconomics* 2012; 30 (9): 779-94.

- Linley WG, Hughes DA. Decision-makers' preferences for approving new medicines in Wales: a discrete choice experiment and exploration of external validity. *Pharmacoeconomics* 2013. doi: 10.1007/s40273-013-0030-0. [Epub ahead of print].

- Linley WG, Hughes DA. Societal views on NICE, Cancer Drugs Fund and Value-based Pricing criteria for prioritising medicines: a cross sectional survey of 4118 adults in Great Britain. *Health Economics* 2012. doi: 10.1002/hec.2872. [Epub ahead of print].

Chapter 1
Thesis introduction and background

1.1 Thesis overview

'...economics is notoriously a dismal science and cannot be otherwise when applied to medicine. Those who think that medicine is above economic laws are destined to have their illusion rudely shattered, for a limited and planned economy obviously cannot support a health service expanding with the acceleration which medicine now exhibits...'

— Ffrancgon Roberts, Physician, Addenbrooke's Hospital, 1948

In the UK we benefit from a publically-funded National Health Service (NHS), which provides health services that are largely free at the point of delivery. The provision of health care, and use of health care services, involves complex decision-making on many levels by a wide range of stakeholders, including: patients and their carers, clinicians, budget holders, policy-makers, health technology industries, and the tax-paying public. An understanding of the priorities and trade-offs made by these stakeholders is essential to ensure that difficult decisions around the reimbursement and use of health interventions are aligned with the preferences of those who are affected by those decisions, and the broader objectives of the NHS to provide comprehensive care for all and the sustainable use of finite, tax-funded resources.

This thesis explores the preferences of different stakeholders involved in health care decision-making, particularly national-level decision-makers and the public, to determine the relevance of current and proposed criteria and processes for medicines reimbursement in the UK NHS.

1.2 The UK National Health Service history and finance

The UK National Health Service (NHS) was established in 1948 to address the inefficiencies in the disparate health services, and the inequities in access, that existed beforehand (Whitehead 1994). The founding principles have remained the same throughout its 65-year history: the NHS should provide comprehensive, high quality, needs-based care for all, free at the point of need, irrespective of ability to pay (Willink 1944; Dept Health 2012a). However, as the NHS is funded almost exclusively by general taxation, there are further obligations to provide the best value for tax-payers' money, and the sustainable use of finite resources (Dept Health 2012a). There are inevitable tensions between these competing obligations (Weale 1998). Issues of affordability were recognised and acknowledged before the NHS was established (Bevan 1948), and

rationing, both implicit and explicit, has been inevitable ever since (Klein and Maybin 2012).

There have been numerous reforms and reorganisations of the NHS over time, driven by the need to improve efficiencies. One of the most profound of these was the introduction in 1991 of a split between those responsible for the purchase and provision of care, which in effect created an internal market for health services. Prior to this, both the planning and delivery of services had been directed centrally by the Department of Health via regional health authorities, but the introduction of the purchaser-provider split shifted finances and commissioning responsibility to a local level. Coupled with the creation of self-governing hospital Trusts, the aims of these moves were to incentivise better management of costs and introduce competitive pressure on hospitals to deliver better quality and more efficient services (Brereton and Vasoodaven, 2010;). However, evidence that these aims were achieved appears mixed; increased management and transaction costs, and evidence of inequities of access as a result of the internal market have prompted reviewers to conclude that the increased costs of introducing competition have not been accompanied by the intended benefits. Nonetheless, the split between those responsible for commissioning care and those responsible for providing care has been retained in some form ever since, and has fostered an acute awareness of costs, efficiency and accountability across the entire service (Brereton and Vasoodaven, 2010).

NHS expenditure has increased each year in real terms by an average of 4% (Appleby 2013). In 2010/11, total UK government spending on the NHS was around £131 billion, equivalent to 9% of gross domestic product (Hawe et al., 2011). By far the largest expenditure is on staff costs but medicines, as the most common clinical intervention, are the next biggest single expenditure (National Prescribing Centre 2008), accounting for around 12% of the budget in 2011 (Hawe et al., 2011).

1.3 Medicines and the NHS

1.3.1 Medicines budgets

Since the reforms of 1991, around 80% of the NHS budget has been controlled by local NHS organisations. Prescribing budgets come within the unified budget allocations for local NHS organisations (Hawe et al., 2011); therefore, any overspend on medicines must be met by a reduction in other parts of the budget. It is little wonder then, that there is

much attention focused on the use of medicines by local and national NHS decision-makers, and steps taken to actively control prescribing budgets.

Around two-thirds of the NHS expenditure on medicines results from primary care prescribing (NHS Information Centre 2011). It is apparent from available data that the volume of medicines prescribed (per capita) has increased year on year in both primary care and hospital settings. NHS expenditure on hospital prescribing has increased, but NHS expenditure on primary care medicines has remained fairly constant, and as a proportion of overall NHS expenditure has decreased in the UK since 2003 (Hawe et al., 2011; Baillie et al., 2011).

The reasons for the relative decrease in primary care prescribing costs are complex, but are linked to the efforts to control prescribing budgets, including switching from branded medicines to their generic equivalents, and what some perceive to be a resistance to the use of high cost new medicines (OFT 2007; Baillie et al., 2011).

In 2010, an international review, commissioned by the UK Government, reported the UK to rank 8th out of 14 developed countries for the uptake of a range of medicines for the treatment of diseases causing significant morbidity or mortality (Richards 2010). There was much variation evident within and between countries. Overall, the UK ranked relatively high for three categories (thrombolytics for acute myocardial infarction, lung surfactants for respiratory distress syndrome and use of statins) and relatively low for seven categories (both newer and older cancer drugs, medicines for dementia, hepatitis C, multiple sclerosis, rheumatoid arthritis, and second-generation antipsychotics). Reasons identified for the variation between countries included the impact of health technology assessment (HTA) processes, the effectiveness of service planning, and prevailing clinical culture. The report concluded that countries that spend the most on health do not always have the highest levels of usage and low spenders can be high users of medicines (Richards 2010).

1.3.2 Medicines regulation and pricing

Before a new medicine is available for widespread use in humans the manufacturer must first obtain a marketing authorisation from the regulatory authorities (e.g. UK Medicines and Healthcare products Regulatory Agency [MHRA] or the European Medicines Agency [EMEA]). This requires the manufacturer to demonstrate the efficacy, safety and quality of

the medicine (EEC Directive 65/65/EEC1965; Medicines Act 1968). Demonstration of efficacy, and to an extent safety, is usually achieved through clinical trials programmes that compare the new medicine against either placebo or against an active comparator treatment in the patient population of interest.

Bringing a new medicine to the market is a costly and commercially risky endeavour; the most recent estimates place the average research and development costs at US$1.5 billion (Mestre-Ferrandiz et al., 2012) per successful launch. However, such sums have been refuted as gross overestimations propagated by the pharmaceutical industry (Relman and Angell, 2002; Light and Warburton, 2011). Whatever the actual cost, the time required for development and then approval of new medicines by regulatory authorities can certainly range widely from 5-10 years (Mestre-Ferrandiz et al., 2012), and can significantly erode the duration of patent exclusivity within which manufacturers are able to generate profits.

Needless to say, branded medicines that are within patent protection are generally priced at a premium. Once the patent has expired, other companies are able to manufacture and market generic (non-proprietary) versions of the same medicine, which introduces competition and usually results in a significant reduction in their prices. However, pharmaceutical manufacturers have been known to adopt strategies – known as 'evergreening' strategies – that, in effect, extend the market exclusivity of their medicines whilst offering little further therapeutic benefit. Examples include reformulation of the medicine, adaptation of dosage forms and isolation of active metabolites (House of Commons Health Committee 2005; Hughes and Ferner, 2010; Hitchings et al., 2012).

Alternatively, once a new class of medicine has been marketed, different manufacturers may produce alternative branded medicines within the same class, known as 'me too' medicines. These often offer little advantage over the original, but being branded they attract similar premium prices until their patents expire. There has been an increase in the number of 'me too' drugs being licensed over the last 10-15 years, and there are suggestions that this approach to pharmaceutical development is inherently less commercially risky, and so more profitable, than developing medicines that are truly innovative (House of Commons Health Committee 2005; Dept Health 2010a).

Arrangements for the pricing of branded medicines have, for the last half century, been based on the Prescription Pricing Regulation Scheme (PPRS). This is a voluntary agreement between the pharmaceutical industry and the UK government, which allows

pharmaceutical manufacturers to set the price of their new medicines at any level so long as the profits made from NHS sales across their entire portfolio of medicines do not exceed a specified amount (ABPI 2009). This, combined with the strategies for extending market exclusivity above, mean that the usual mechanisms of price controls based on market forces tend not to apply. Branded medicines account for over 70% of the costs but only 35% of the volume of medicines used in the NHS (Dept Health 2012b).

A review by the Office of Fair Trading (OFT) in 2007 concluded there were shortcomings in this pricing scheme; the prices the NHS has been paying for new medicines under PPRS do not reflect the therapeutic value they deliver, and the scheme offers little incentive for companies to invest in development of truly innovative medicines or those that address areas of unmet need (OFT 2007). This is a view that has been echoed by recent analyses of past marketing authorisations (Light and Lexchin, 2012). It is therefore proposed that a new pricing system should be introduced that better reflects the therapeutic value that medicines deliver (OFT 2007; Dept Health 2010a) (see section 1.5.3 for further details).

At this point it is worth noting that the licensing and pricing of medicines, under current pricing mechanisms,, reveal little about how effective they are or whether they represents a good use of limited health care resources compared with other treatments that are already in use in practice. HTA bodies have developed systems to evaluate the clinical and cost effectiveness of medicines relative to alternative treatments; however, before discussing these in detail (see section 1.5), it is useful to consider the wide range of stakeholder relationships involved in health care decision-making.

1.4 Stakeholder relationships in health care decision-making

Health care is argued to differ from other commodities in a number of ways, and consequently, health care markets differ from the classical economic models of competitive free markets (Arrow 1963; Culyer 1989). Competitive free markets are efficient where consumers are well-informed and can judge quality and prices to make rational choices to purchase commodities, the supply of which is limited only by demand. However, in health care markets such as exist in the UK NHS, prices are not determined by usual market forces (see section 1.3.2), and as services are provided to patients free at the point of need, demand is high and supply is limited. Patients do not possess the specialist knowledge required to make rational choices and so rely on the information and

advice of health care professionals, who are regulated with the aim of ensuring quality and patient protection (Culyer 1989). Collectively, the market for health care in the UK is highly dependent on government intervention.

1.4.1 Agency relationships

The majority of the population in the UK is registered with a general practitioner (GP) and, excluding emergency care, these primary care clinicians are the initial point of access for patients, and act as gatekeepers, to prescription medicines and other health care resources or services. The relationship between patients and their clinicians is often referred to as one of principal and agent, in which the clinician, who is presumed to be better informed about health and health care, acts on behalf of the patient in making decisions about their most appropriate care and treatment (Mooney and Ryan 1993; Scott and Vick 1999). Perfect agency would imply the agent makes decisions that fully reflect the principal's preferences and needs, unaffected by their own; however, this would require clinicians to be fully informed of their patients' preferences. Given their different remits, levels of knowledge, skills and experience, and different incentives, perfect agency is unlikely to exist (Culyer 1989; McGuire A et al., 1997).

Of course, health care decision-making is not confined to the patient-clinician relationship. Decisions on the availability of medicines to treat current and future patients, for example, are made on their behalf by a wide range of stakeholders and decision-makers: regulatory authorities (discussed in 1.3.2) grant permission for pharmaceutical companies to sell their medicines based on assessment of risks and benefits; HTA bodies (e.g. National Institute for Health and Clinical Excellence [NICE] in the UK – discussed in 1.5) make reimbursement decisions that determine whether the NHS should buy these medicines based on their clinical and cost effectiveness; Government and politicians issue national directives and targets that prioritise some disease areas, patient groups and/or treatments over others; and in the absence of national guidance, decision-makers with responsibility for local population health make decisions on which treatments should be included on their formularies and funded from local budgets.

These decision-makers therefore act as agents on behalf of patients, the wider public, and clinicians, and with good reason. It would impossible for patients or their clinicians to undertake robust assessments of all aspects of all possible treatments to determine their net clinical or economic value (Claxton et al., 2009). However, the view that 'decision-maker knows best' is increasingly challenged.

1.4.2 Stakeholder preferences

Greater efforts directed towards shared decision-making between patients and clinicians are suggested to improve patient acceptance and adherence to treatment, improve health outcomes and significantly reduce health service costs (Dept Health, 2010b; Coulter and Collins, 2011). In addition, greater efforts have been made to include lay people and patients among regulatory (Breckenridge, 2011) and NICE decision-making committees (Rawlins 2005), and local formulary decision-making groups (NICE 2012a).

Having established that perfect agency is unlikely to exist (Culyer 1989; McGuire A et al., 1997) it would be useful to understand how preferences vary between different types of stakeholders when making health care decisions. Identification of common areas of agreement or disagreement between different stakeholders could potentially provide useful insights for clinicians and also higher-level decision-makers such as HTA or regulatory bodies.

1.5 Health technology assessment and the economics of health care decision-making

1.5.1 UK HTA bodies

The shift of budgets and commissioning responsibility to a local level brought with it issues of fragmented decision-making and geographical variation in access to expensive new medicines, in what became known as 'post code prescribing'. In response, the National Institute for Health and Clinical Excellence (NICE) was established in 1999 to provide national recommendations on the use of medicines and other health technologies based on evidence of their clinical and cost effectiveness. Combined with a statutory funding direction for England and Wales, to make funding available for medicines within three months of positive NICE guidance being issued, the aim was to reduce variation in the availability and quality of care (Dept Health 1998).

The approach to decision-making taken by NICE is discussed below (section 1.5.4) but the process of technology appraisal is important to understand here, as this sets the context of other HTA bodies in the UK.

Initially NICE focused on conducting multiple technology appraisals (MTA), i.e. the simultaneous appraisal of several medicines with marketing authorisations for the same indications. Individual manufacturers submitted clinical and economic evidence in support of their own products, and NICE also commissioned academic centres to prepare independent reviews of the available clinical literature and construct independent economic models. These collective sources of evidence were then considered in the appraisal process. However, this approach is time consuming – recent retrospective analyses suggest a median time from final scope to completion of 74 weeks, which is 14 weeks over NICE's target (Casson et al., 2013) – and local NHS organisations were accused of delaying the use of new medicines until NICE guidance had been issued, in what became known as 'NICE blight' (Haycox 2008). In response to criticisms of delays, in 2005 NICE introduced an additional single technology appraisal (STA) process, in which only one medicine is appraised for a single indication and only an independent review of the manufacturer's submitted evidence is undertaken [NICE 2009a]. This has significantly reduced the time taken for NICE guidance to be issued, to a median of 48 weeks, although this is still 5 weeks over NICE's target (Casson et al., 2013).

NICE's STA process is similar to the HTA processes adopted separately by Scotland and Wales. NICE technology appraisal guidance is not compulsory in Scotland, and the Scottish Medicines Consortium (SMC) was established in 2002 to undertake appraisals of medicines and issue guidance to local NHS bodies. In contrast to NICE, which only appraises medicines referred to it by the Department of Health, the SMC appraises all new medicines and new licensed indications. Comparisons of NICE's STAs and SMC's appraisals suggest similar recommendation rates (90% and 80%, respectively), but SMC issues guidance much closer to market authorisation of medicines than does NICE (median 7.35 months versus 16.05 months) (Ford, et al., 2012).

In Wales, although NICE technology appraisal guidance is compulsory, the All Wales Medicines Strategy Group (AWMSG) began appraising high cost medicines in 2002, and in 2010 expanded this to include all medicines and new licensed indications where NICE guidance was not imminent; however, should NICE guidance on the same medicines become available, this supersedes the earlier recommendations of AWMSG.
Despite the initial aims of NICE in 1999, and the statutory funding direction, recent audits have shown regional variation exists in the uptake of NICE and AWMSG-recommended medicines in England (NHS Information Centre 2012) and Wales (Baillie et al., 2011). These findings have prompted the Chief Executive of the NHS in England to instruct local NHS organisations to publish their medicines formularies and make compliance with

NICE technology appraisals a contractual requirement under the NHS Operating Framework (Dept Health 2012c).

1.5.2 Preferences of HTA decision-makers

There is an increasing body of evidence on the factors that influence HTA bodies' reimbursement decisions for new medicines, including the decisions of NICE in the UK. Based on logistic regression modelling of past recommendations, the incremental cost effectiveness ratio (ICER – discussed in section 1.5.4.2), the uncertainty surrounding the ICER, the number of patients affected by the disease, the availability of alternative therapies, the number of supporting randomised controlled trials (RCTs), and the presence of a submission from a patient organisation, have been reported as significant determinants of NICE recommendations (Devlin & Parkin, 2004; Dakin et al., 2006).

Alternative approaches to explore decision-making include stated preference techniques, such as discrete choice experiments (DCEs). These have a particular appeal within the health economics arena being grounded as they are on random utility/consumer theory, which underpins welfare economics theory (see 1.5.4.1 below), and demand theory (Lancaster 1966; Ryan et al., 2008). They permit estimation of the impact, and willingness of decision-makers to trade-off, various characteristics (attributes) making up hypothetical choice alternatives. However, if DCE-based stated preferences are to reflect real-life decision-making, it is imperative that attributes and levels employed to describe choice alternatives have face validity, and that choice tasks represent as far as possible the true nature of decision-problems (Lancsar and Louviere, 2008). Few DCEs have been conducted in HTA committee members to date (Tappenden et al., 2007; Koopmanschap et al., 2010; Whitty et al., 2011) and none of these, nor many other health-related DCEs to date, have compared stated decision-making behaviours against actual behaviours (De Bekker-Grob et al., 2010).

1.5.3 Future HTA processes and Value-based Pricing in the UK

Currently, AWMSG, NICE, and SMC make recommendations on the use of medicines based among other things on their clinical and cost-effectiveness, the latter of which is determined in part by the price of the medicine agreed with the Department of Health under the PPRS scheme (see 1.3.2). In response to the OFT report on pharmaceutical pricing (OFT 2007), a new system of value-based pricing is to be implemented in the UK from 2014, which aims to ensure that the price of branded medicines better reflects the

therapeutic value and wider benefits they deliver (Dept Health 2010a). Details of how the scheme will be operated are still lacking, but the outcome of appraisal will be an acceptable price rather than a recommendation for use (Hughes 2011). Nonetheless, the methods of HTA and the economic theory underpinning healthcare decision-making remain largely the same.

1.5.4 The economics of health care decision-making

In publicly funded health services, resources always have been, and always will be, limited. Difficult decisions therefore have to be made around how best to use those resources . Economic analyses may inform those decisions by comparing the benefits and costs of competing uses of limited resources to determine their most efficient allocation; however, as demonstrated by the founding principles of the NHS, that exist today, (Dept Health 2012a) (see section 1.1), issues of fairness in the distribution of health care resources are also of paramount importance. Value judgements need to be made on the desirability of alternative allocations of health care resources and conclusions of value for money.

1.5.4.1 Normative economics: welfarist and non-welfarist approaches

The branch of economics concerned with the application of value judgements is termed normative economics, of which there are broadly two schools: welfare and extra-welfare economics. There is a broad contemporary literature discussing each approach (Culyer 1989; Sassi et al., 2001; Brouwer et al., 2008), and what follows is a brief overview that is sufficient to provide the background to this thesis.

The theory underpinning welfare economics is essentially that of random utility/consumer theory, which is assumed in any competitive free market. Individuals are assumed to be the best judges of their own wellbeing, or utility, and they act in rational ways to maximise their own utility in the choices they make. The utility that individuals derive from health care is determined only by the outcomes of their consumption of health care, such that the process of health care has no bearing on their utility. In terms of decision-making, the aim is to maximise social welfare, which is the product of the utilities of individual members of society (Culyer 1989).

It is implicit within this welfarist framework, then, that an increase in social welfare arising from a change in the current distribution of resources is better than none. Any change in

that distribution that can increase the utility of one member of society without reducing the utility of another is an improvement, known as a Pareto improvement. Taking this further, a *potential* Pareto improvement in social welfare can occur if a redistribution of resources leads to an increase in utility for some individuals and a decrease in utility for others, so long as those whose utility increases could potentially compensate those whose utility decreases and still remain better off than before (Kaldor 1939; Hicks 1939). The notion of compensation implies some willingness to pay for those increases in utility, and it is the collective willingness to pay that provides a measure of the benefit or value of a given use of health resources. This willingness to pay may therefore be used as a method of economic evaluation, known as cost-benefit analysis (CBA) (Drummond et al., 1997).

There are a number of limitations to this approach. Health care systems do not represent the competitive free markets to which standard consumer theory may be applied (Arrow 1963; Culyer 1989) (see section 1.4). In the context of limited resources, there are unlikely to be many policies that would lead to a redistribution of resources for the benefit of some individuals without impacting adversely on others (Coast, 2004) and, in reality, compensation for those adversely affected is unlikely to be forthcoming, which introduces potential equity issues around who should gain and who should lose from any given redistribution of resources.

The use of utility to reflect wellbeing is challenged because each individuals' valuation of their utility would be affected by their own characteristics, such as need (Sen 1985; Culyer 1989). It is therefore likely that preferences will differ between individuals, which raises problems in the extent to which social welfare, as framed by welfare economics, could ever truly exist: mathematically, Arrow's impossibility theorem demonstrates that, under such circumstances of conflicting preferences or utility functions, complete and consistent social choices are impossible unless some entity is permitted to dominate, or dictate, choices (Arrow 1950).

Other problems also exist in relation to the CBA framework of economic evaluation that is derived from the welfarist model. CBA requires assignment of monetary values to length and quality of life, which may be an uncomfortable exercise for some, and an individual's willingness to pay will be influenced by their ability to pay, so that policies informed by economic evaluations using willingness to pay may favour the wealthy (Coast 2004).

Therefore, despite its grounding in economic theory, the welfarist model appears to have limited relevance to the allocation of health care resources (Culyer 1989; Tsuchiya and

Williams, 2001). This has led to the development of an alternative, extra-welfarist approach to economic evaluation. This permits additional considerations to utility to be taken into account in an individual's assessment of wellbeing, namely need, which is met with the use of health care resources, the outcomes of which are some level of health (Culyer 1989). However, in practice, rather than health being used as additional considerations to utility, it has replaced utility as the maximand.

As attaching some monetary value to benefits is problematic (as noted for CBA), the costs of health resources are left separate from the benefits in extra-welfarist economic evaluations, giving rise to measures of costs per unit of effect (e.g. cost per life year saved). These types of economic evaluations are known as cost-effectiveness analyses (CEA), where the measure of benefits is left in natural units (such as life years saved). A problem with this approach, though, is that different health programmes may have different natural units of benefit, which makes comparisons across several health programmes difficult. Furthermore, there may be more than one outcome of health programmes that is important. Therefore, to try to address these types of problems, an extension of CEA has been developed, known as cost utility analysis (CUA) (Weinstein and Stason, 1977; Drummond et al.,1997).

In CUA, multiple outcomes experienced over time may be incorporated into the unit of effectiveness, and each time-based state is quality adjusted by a set of values or weights (utilities) that reflect the relative desirability (in terms of health-related quality of life) of time spent in that state. Therefore, elements of both length and quality of life are incorporated within a single index called a quality-adjusted life year (QALY). As a common unit of benefit, an advantage of QALYs is that they enable the broad comparison of different health programmes (Weinstein and Stason, 1977; Drummond 1989); however, this requires some assumptions that may be viewed as limitations of the QALY approach (Mooney 1989; Rawles 1989).

The extent to which a QALY is composed of length or quality of life is considered irrelevant, so individuals are assumed to be risk neutral for length of life. In addition, QALYs are also aggregated across all individuals, such that one QALY gained in one group is considered exactly the same as one QALY gained in another group, which of itself offers little room to consider the different levels of health of patients before treatment or any equity-related issues. Questions are also raised about whose values should be used to derive the utility weights to quality adjust health states (Dolan 1999).

Nonetheless, QALYs are the preferred measure of health benefits for national decision-makers in the UK (NICE 2008a; AWMSG 2011a; SMC 2012).

1.5.4.2 Decision rules and cost effectiveness thresholds

Irrespective of whether CEA or CUA is used, the decision-rules to determine cost effectiveness are more complex than those required for CBA. For the latter, as both costs and benefits are expressed in monetary values, if the sum of the benefits is greater than the sum of the costs, then the programme is efficient (cost effective). However, for CEA/CUA, a direct comparison between the value of effects and costs is not made, and trade-offs therefore need to be made between the different costs and different benefits of alternative uses of health resources (Drummond, et al., 1997).

Appendix to Chapter 1 provides a detailed explanation of the method of considering the costs and benefits of alternative use of health resources. In brief, this is achieved by estimating the incremental costs (C) to achieve an additional unit of health benefit (E, however measured) with a new health technology (A) relative to an alternative use of those health resources (B, i.e. a relevant comparator treatment). This is termed the incremental cost effectiveness ratio (ICER) and is given by:

$$ICER = \frac{C_A - C_B}{E_A - E_B}$$

Depending on whether the ICER for the new health intervention is less than or greater than a threshold value (λ), representing the decision-maker's maximum willingness to pay for an additional unit of effectiveness, then the new intervention would be deemed to be a cost effective use of health resources, or not.

There has been much debate in the literature about what the threshold value is, or should be, in the UK (Rawlins and Culyer, 2004; McCabe et al., 2008). In 2004, NICE's guide to the methods of technology assessment stated a fixed ICER threshold was not used, but suggested that a technology with a most plausible ICER below £20,000 per QALY gained was unlikely to be rejected on grounds of cost ineffectiveness and the case of recommending a technology with an ICER above £30,000 per QALY would need to be increasingly strong (NICE 2004). A similar view was also stated by the Chair and Vice Chair of NICE the same year, who also acknowledged there was no empirical basis for this threshold range (Rawlins and Culyer, 2004). Empirical studies have been undertaken

in attempts to determine the social value of a QALY, but have encountered numerous methodological challenges that render estimates unstable (Baker et al. 2010; Dolan et al., 2008). Analyses of past NICE decisions suggested a threshold value in excess of £40,000 per QALY (Devlin and Parkin, 2004), and the most recent study suggests a conservative threshold of less than £18,317 per QALY would be appropriate for reflecting the opportunity cost of the benefits foregone in recommending a new technology, based on local NHS programme budgeting data from England (Claxton et al., 2013).

But whatever the threshold value assumed, the decision-rules of CEA/CUA under the extra-welfarist approach, and of CBA under the welfarist approach, implicitly assume that the aim is to maximise health gain within the constraints of limited resources. Not only is this potentially inconsistent with the wider stated aims and objectives of the NHS (Dept Health 2012a), but it is also at odds with the empirical evidence of the public's views on the distribution of health care resources (Dolan et al., 2005; Bobinac et al., 2012).

In summary, the theories underlying economic evaluation of health care, be they welfarist or extra-welfarist, have many limitations in the context of the wider objectives and obligations of health care services. As noted by Culyer (1989), neither approach can yield final answers. Economic evaluation provides a tool to aid decision-making but cannot be the determinant of allocation decisions. It is therefore necessary to consider other factors, beyond economic efficiency, that are relevant to health care decision-making.

1.5.5 Ethics and equity in health care decision-making

1.5.5.1 Distributive justice and accountability for reasonableness

The approaches to economic evaluation described above reflect a utilitarian philosophy, which considers that society should seek to maximise overall utility in its approach to the distribution of health and health care. In contrast to this, a qualified egalitarian philosophy would consider that society should distribute resources to ensure each member receives a fair share of the opportunities available, which requires equality among individuals but permits inequalities that contribute significantly to the benefit of the least advantaged (Beauchamp and Childress, 2001). In the context of the overall aims and objectives of the NHS, to provide comprehensive, needs based care for all, whilst providing the most effective, sustainable use of finite resources (Dept Health 2012a), neither philosophy alone is adequate.

In the absence of consensus on the moral principles or philosophies that should govern priority setting, it is argued that fairness and legitimacy should be achieved through the processes involved (Daniels 2000). The dominant ethical priority setting framework to have emerged, and to which NICE aims to subscribe (Rawlins 2005, NICE 2008b), is the '*accountability for reasonableness*' framework, which requires that four conditions are met for fair and legitimate decision-making: decisions and their rationales should be transparent and made public; the rationales should be reasonable and relevant; there should be opportunities for challenge and revisions; and, there should be regulation to ensure the first three conditions are met (Daniels and Sabin, 1997). The second of these conditions, ensuring that the rationales for decisions are relevant requires that the reasons, evidence, and rationales used in priority-setting are those that "fair minded" people would accept as being relevant to meeting population health needs fairly in the context of resource constraints, which therefore requires involvement of the people affected by those decisions (Daniels and Sabin, 1997; Rawlins 2005).

1.5.5.2 Public involvement in HTA decision-making

There has been increased involvement of patients and the public in health care decision-making (Abelson et al. 2003; Whitty 2013), and NICE has gone further than most HTA bodies to involve public views in its decision-making processes (Buxton and Chambers 2011; Shah et al. 2011a). In 2002, NICE established a Citizen's Council, an adaptation of the Citizens jury model, to incorporate public views and inform on the social values that should be applied in its decision-making (Rawlins 2009). This is a standing committee composed of 30 lay volunteer members of the public who meet periodically for three days at a time to deliberate on topics around which NICE requires a public view. Fifteen Citizens Council meeting reports have been published to date, which are intended to contribute to NICE's Social Value Judgements guide to development of guidance (NICE 2008b).

However, the extent to which the views of a small group of self-selected volunteers can reflect the diverse views of the general population is open to challenge (Ryan, et al., 2001; Abelson et al., 2003; Buxton and Chambers 2011). An early evaluation of NICE's Citizens Council was unable to comment on its ability to influence NICE decisions (Davis et al., 2005), and assessments of other public engagement exercises similarly indicate there has been little focus on their effectiveness, or on comparisons of different methods of engagement (Lenaghan 1999; Abelson et al., 2003; Mitton et al., 2009).

1.5.5.3 Criteria for prioritising medicines in the UK

Several medicines with plausible ICER estimates in excess of the £20,000 to £30,000 per QALY threshold range (discussed in section 1.5.4.2) have been approved by NICE for use via the NHS (e.g., sunitinib for advanced renal cancer, riluzole for motor neurone disease). Justification for this departure from the usual cost effectiveness threshold range includes the social value judgements of NICE's Citizen Council. Based predominantly on its views, six specific criteria besides clinical and cost effectiveness have been put forward by the Chair of NICE and Chairs of NICE Technology Appraisal Committees as reflecting societal preferences in the allocation of health resources (Rawlins et al., 2010): severity of the underlying illness; significant innovations; disadvantaged populations; children; end-of-life treatments; and cases where disease symptoms are not well reflected or addressed in clinical trials or health-related quality of life measures (stakeholder persuasion).

Both the All Wales Medicines Strategy Group (AWMSG) and the Scottish Medicine Consortium (SMC) also permit additional considerations in the appraisal of medicines for the treatment of rare diseases (AWMSG 2011b; SMC 2010). In England, the Cancer Drugs Fund (CDF) was introduced in 2011 to facilitate access specifically to cancer medicines that have received a negative opinion from NICE on the grounds they do not represent a good use of NHS resources, or which have not yet been appraised [Dept Health 2010c]. And under the VBP system to be implemented from 2014 (see section 1.5.3), it is proposed that medicines will be rewarded with higher prices if they: treat severe conditions; address unmet needs; are innovative; and have wider societal benefits [Dept Health 2010d].

Of these criteria, only severity of disease is shown consistently to be a valid criterion for prioritising health care resources in the empirical ethics literature (Dolan et al., 2005; Shah 2009). For the other criteria, the evidence is either mixed (Dolan et al., 2005; Sassi et al., 2001) or absent. It is apparent, therefore, that current prioritisation criteria used in pricing and reimbursement systems in the UK, and recent initiatives to address their perceived short-comings, lack robust supporting empirical evidence that they reflect societal preferences for the allocation of scarce health resources.

1.6 Thesis aims

In order to explore the preferences of different stakeholders and determine the relevance of current and proposed criteria and processes for medicines reimbursement in the UK NHS, seven research questions are addressed in the following Chapters:

Chapter 2:

Research question 1: *How and to what extent do key considerations in health care decision-making differ between decision-makers (agents) and those affected by their decisions (principals)? In particular, are there differences in preferences between principals and agents towards benefits and harms of health care interventions, and towards health-related and process-related aspects of health care?*

Methods: Systematic literature review of health-related conjoint analyses involving two or more stakeholders involved in principal-agent type relationships.

Unique contribution: In contrast to previous reviews, this systematic review has characterised the extent to which differences may exist in the preferences of a range of stakeholders, across a range of health care services and interventions. The importance of process-related considerations relative to health-related considerations can be characterised for the first time.

Chapter 3:

Research question 2: *What does previous guidance of the All Wales Medicines Strategy Group (AWMSG) reveal about the factors that influence reimbursement decisions for new medicines?*

Methods: Secondary data collection and analysis using logistic regression modelling to explore the revealed preferences of AWMSG for recommending the use of new medicines in NHS Wales.

Unique contribution: This is the first study to explore the revealed preferences of an alternative decision-making body to NICE in the UK. Importantly, this study contributes to the design and validity testing of the stated preference study described in Chapter 4.

Chapter 4:

Research question 3: *What are the stated preferences of AWMSG appraisal committee members towards specific new medicines adoption criteria?*

Research question 4: *Are the stated preferences of AWMSG appraisal committee members externally valid?*

Methods: Primary data collection and analysis, using stated preference, discrete choice methods.

Unique contribution: This is the first study to explore the stated preferences of an alternative decision-making body to NICE in the UK. This is one of only a few studies of the external validity of stated preferences in the health economics arena, and is the first exploration of the external validity of the stated preferences of any HTA body.

Chapter 5:

Research question 5: *Do the current prioritisation criteria used by NICE and the UK Government's assumptions used to justify the introduction of the Cancer Drugs Fund in England reflect the public's stated preferences for health care resource allocation?*

Research question 6: *Do the proposed criteria for rewarding new medicines with higher prices under the future value-based pricing system in the UK reflect the public's stated preferences for health care resource allocation?*

Methods: Primary data collection and analysis, using choice-based methods to elicit the preferences of a large sample of the public.

Unique contribution: This is the first comprehensive empirical analysis of public views on current and proposed prioritisation criteria that are central to UK policies on medicines reimbursement. This is the first empirical analysis of public preferences towards cancer treatments, and has been used to support the Welsh Government's rejection of the introduction of a Cancer Drugs Fund in Wales. It was also the first empirical analysis of UK public views towards the funding of treatments for rare diseases, and NICE's end-of-life policy.

Chapter 6:

Research question 7: *How effective is NICE's Citizens Council in reflecting the views of the public and influencing NICE's decision-making?*

Methods: Secondary data collection and comparative review of NICE documentation and public survey findings.

Unique contribution: This is the first review to explore the congruence of NICE Citizens Council's views with UK public views, and their implementation in NICE process guidelines. Few studies of public engagement exercises have explored effectiveness beyond participation.

Chapter 2

Systematic review of health-related conjoint analyses

2.1 Abstract

Background: The preferences of clinicians and patients for health care interventions are known to differ. Shared decision-making is high on the UK health policy agenda, and there are increasing expectations to incorporate the views of all relevant stakeholders into decision making processes at all levels.

Objectives: To characterise how and the extent to which key considerations in health care decision-making differ between decision-makers (agents) and those affected by their decisions (principals).

Data sources and study eligibility: The following databases were searched from their inception up to 24th July 2012 for published, English-language, health-related conjoint analyses conducted in two or more respondent types that may reasonably be viewed as being involved in a principal-agent type relationship: EMBASE, HMIC, AMED, MEDLINE, IPA, CINAHL, Biomedical Reference collection: Corporate, SSCI, A&HCI, CPCI-S, CPCI-SHH, EconLit and PyschINFO.

Methods: A search strategy was designed with reference to previous systematic reviews of conjoint analysis-type studies. Statistically significant attributes in each eligible conjoint analysis were categorised according to: baseline patient/disease characteristics; health intervention effectiveness/benefits; health intervention safety/risks; non-health/process-related aspects of health care; and costs/economic considerations. Their implied rank order of importance to respondents was compared and contrasted to make inferences on the extent to which the views and preferences of principals and agents differ.

Results: From 34,555 database records, 45 studies were included, from which 95 different principal-agent comparisons were extracted. The majority involved patients versus clinicians. The rank order of importance of all attributes included in each comparison experiment differed in 75% of cases, was the same in 18%, and was unclear in the remainder. Patients and clinicians differed in the importance attached to all attribute groups in 58%-69% of experiments, with similar figures for other stakeholders. There were no statistically significant differences between patients and clinicians in their benefit-risk or their health outcomes-process trade-offs. Patients ranked 83% of benefit attributes and 58% of safety attributes among the top three most important attributes (NS, $p>0.1$), with similar figures for clinicians. Health-related attributes were ranked among the top three most important attributes significantly more than process-related attributes were by

both patients (68% vs. 42%, respectively; $p<0.05$) and clinicians (70% vs. 44%, respectively; $p<0.05$).

Limitations: Conjoint analyses require respondents to express preferences for hypothetical scenarios, and stated preferences may differ when faced with real choices. Categorising attributes is subjective. The simple rank ordering of importance does not inform on how much more important the first ranked attribute is relative to the second and subsequent ranked attributes. Sample size was small for subgroup analyses across different respondents and different attribute types.

Conclusions and implications: The preferences of principals and agents differ at all levels of health care decision-making, more often than not. Initiatives to encourage decision-makers at all levels to actively engage and incorporate the views of those affected by their decisions are supported. The importance of non-health, process-related aspects of health care is confirmed, lending support to proposals to consider the non-health benefits of medicines under the future value-based pricing system; however, the findings further suggest that health outcomes should be given more weight than process-related aspects in those considerations.

2.2 Introduction

The White Paper *Equity and Excellence: Liberating the NHS* put patient choice at the heart of the recent English NHS reforms (Dept Health 2010b). Greater efforts directed towards shared decision-making between patients and clinicians are suggested to improve patient acceptance and adherence to treatment, improve health outcomes and significantly reduce health service costs (Dept Health 2010b; Coulter and Collins, 2011) which is an attractive prospect in the context of the unprecedented efficiency savings demanded across the NHS in coming years (Dept Health, 2009; Institute for Fiscal Studies/Nuffield Trust 2012).

However, decision-making in health is not confined to the patient-doctor relationship. Decisions on the availability of medicines to treat current and future patients, for example, are made on their behalf by a wide range of stakeholders and decision-makers: regulatory agencies determine whether the benefit-risk profile of medicines is positive before granting and continuing marketing authorisations; health technology assessment (HTA) bodies, such as the National Institute for Health and Clinical Excellence (NICE), determine whether medicines represent a good use of limited resources when judged against various priority setting criteria; Government and politicians issue national directives and targets that prioritise some disease areas, patient groups and/or treatments over others; and in the absence of national guidance, local decision-makers with responsibility for local population health make decisions on which treatments should be included on their formularies and funded from local budgets. These decision-makers therefore act as agents on behalf of patients, the wider public, and clinicians (principals).

Notwithstanding, the view that 'decision-maker knows best' is increasingly challenged. Efforts have been made to include lay people and patients among regulatory decision-making committees (Breckenridge 2011), and formal benefit-risk assessment methods that can include public and patient preferences alongside expert preferences are being explored (Cross and Garrison, 2008; European Medicines Agency 2010). With respect to the processes of NICE, the '*accountability for reasonableness*' framework for ethical priority-setting (Daniels and Sabin, 1997) requires that decisions and their rationales should be transparent and relevant. This requires involvement of the people affected by those decisions (Daniels and Sabin, 1997), which implies involvement of current and future patients (Rawlins 2005). Involvement of patients and wider stakeholders at all levels of health care decision-making is therefore increasingly encouraged and sought.

Previous reviews of related empirical evidence have discussed isolated examples of studies demonstrating differences in patients' and clinicians' views towards treatment choices (Montgomery and Fahey, 2001; Mulley et al., 2012), but have not characterised the extent to which differences may exist in the preferences of a range of principals and agents, across a range of health care services and interventions.

The objective of this study was to explore the relative importance to principals and agents of common considerations in health care decision-making, via a systematic review of published health-related conjoint analyses. Conjoint analyses are stated preference experiments that are used increasingly in the health care arena to estimate the relative importance to respondents of salient attributes of products and services (Johnson 2006; De Bekker-Grob et al., 2010). The attributes of interest for this study include: patient or disease characteristics; the effectiveness of interventions (including impact on quality of life); the safety profile of interventions; and economic factors (costs or cost effectiveness). Non-health, process-related aspects of health care (e.g. convenience, waiting times, information provision) may also be important to patients. However, there have been suggestions that current approaches used in health technology assessment may not consider these adequately (OHE 2007; Shah et al., 2011b; Sussex et al., 2011), and failure to do so may result in incorrect recommendations (OHE 2007; Watson et al., 2009). Therefore, the importance of non-health, process-related attributes, relative to health-related attributes, was also explored.

2.3 Methods

Reporting of the methods, results and discussion is based around the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Moher et al., 2009) (see Table 2.1A, Appendix to Chapter 2). This systematic review has not been formally registered on a systematic review database.

2.3.1 Database search strategy

The following databases were searched from their inception up to 24th July 2012: EMBASE, HMIC, AMED, MEDLINE, IPA, CINAHL, Biomedical Reference collection: Corporate, SSCI, A&HCI, CPCI-S, CPCI-SHH, EconLit and PyschINFO. The search strategy was developed based on the terms used in previous systematic reviews of discrete choice conjoint analyses (Ryan and Gerard 2003; Guttman et al., 2009; De Bekker-Grob et al., 2010; OHE 2007), and was trialled to test if papers known to the authors were identified. Following refinement, the final search terms employed were:

"Conjoint" OR "Conjoint analysis" OR "Conjoint measurement" OR "Conjoint studies" OR "Conjoint choice experiment*" OR "Part-worth utilities" OR "Functional measurement" OR "Paired comparison*" OR "Discrete choice" OR "Discrete choice experiment*" OR "Discrete choice modelling" OR "Discrete choice conjoint experiment*" OR "Choice experiment*" OR "Pairwise choice*" OR "Stated preference*" OR "Binary choice". Where available, filters for studies in humans and health were applied.

2.3.2 Eligibility criteria and Study selection

Inclusion criteria were empirical conjoint analyses (including discrete choice experiments [DCEs], adaptive conjoint analyses [ACAs] and ranking or rating response studies) conducted in the health care arena, and which included two or more types of respondents who may reasonably be viewed to be involved in a principal-agent type relationship (e.g. patients-clinicians, patients-carers, clinicians-other clinicians, national level decision-makers-public, etc.). The review was restricted to fully published, English language studies. Methodological and review papers without an empirical conjoint analysis, and empirical conjoint analyses that included only one type of respondent, were excluded.

An initial screen of the search result titles and abstracts was conducted by the first author (WGL) to remove duplicates and obvious non-health, or non-conjoint analysis studies, or those reporting only one type of respondent. Where only conference abstracts were identified, the primary abstract author was contacted where possible to request a full copy of the study. Full papers were screened independently by both authors (WGL and DAH) against the inclusion and exclusion criteria, with any discrepancies resolved by discussion.

2.3.3 Data extraction

The first author (WGL) abstracted and categorised all data. A proforma was developed to abstract: the study-type; country in which the study was conducted; the health care areas of application; number and types of respondents; actual attributes assessed and the attribute levels; whether or not the attributes were statistically significant within the estimation model; the relative importance of attributes as reported (or as inferred from the marginal rates of substitution or contribution to overall utility estimates); and a summary of the results. Study quality was considered against the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) good practice guidance for conducting conjoint analyses (Bridges et al., 2011).

Attributes were categorised as reflecting: untreated disease burden or underlying Disease characteristics; intervention or service Effectiveness (including quality of life); intervention or service Safety or adverse effects; intervention or service Costs or cost effectiveness; and non-health/Process-related aspects of interventions or services. These categories were selected on the basis that they provide a convenient means of grouping a wide range of attributes, and are likely to cover the key considerations of one or all stakeholders engaged in principal-agent relationships in health care decisions-making.

2.3.4 Data analyses

The rank order of importance of attribute groups, based on their estimated or implied relative importance, was then established for each type of respondent within each study. Emphasis was placed on the rank order of importance for effectiveness and safety attributes, and for health related- (effectiveness, safety) and process-related attributes. Descriptive statistical analyses (two-sample z-test) of the proportions of attribute-groups ranked as being most important for principals compared with agents were conducted where practicable, given the small numbers of experiments for some principle-agent pairs. As the rank order alone does not inform on the extent to which the first ranked attribute is preferred over the second and subsequent ranked attributes, Yates corrected X^2 test, adjusted for the different numbers of attribute groups across experiments, was used to test the statistical significance of any differences in the proportion of effectiveness versus safety attributes, and the proportion of health-related versus process-related attributes, included among the top three most influential attributes for each of principals and agents.

2.4. Results

The search strategy generated 34,555 database records from which 45 fully published studies were eligible for inclusion in this review (Figure 2.1, p41).

2.4.1 Study characteristics

Table 2.1 (p 42) provides a summary of the included study characteristics. Few studies fully met all 10 of the checklist items in the ISPOR good practice guidelines; however, 39 of the 45 studies considered elements of all 10 checklist items and only one study inadequately reported on multiple checklist items (Table 2.2A, Appendix to Chapter 2).

As several studies included more than two respondent types and/or more than one conjoint analysis-type experiment, 95 different principal-agent experiments were available for analysis. Patients and clinicians were by far the most common principal-agent pair studied, and so form the focus of the results.

Figure 2.1. Study identification and selection

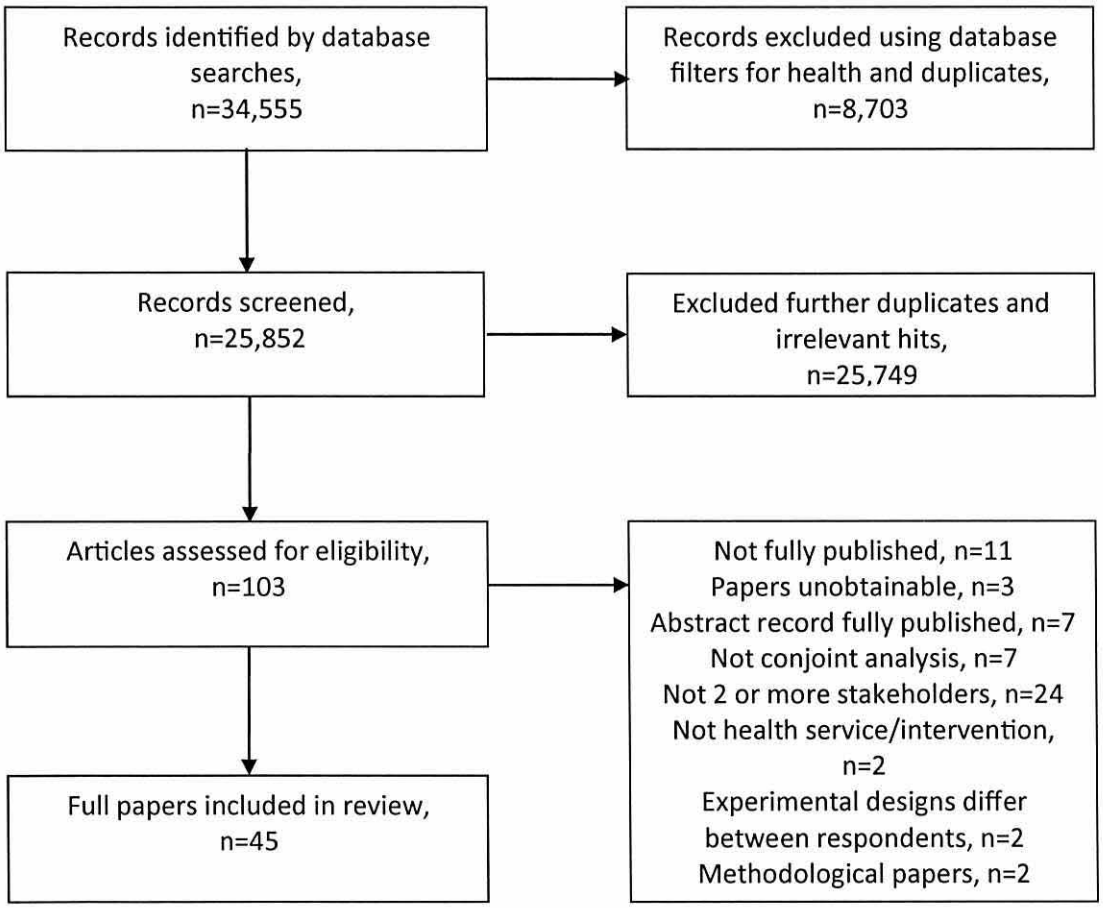


Table 2.1. Summary of study characteristics

	Number of studies (N=45)
Study type*	
DCE, Forced choice	25
DCE, Opt out	10
Adaptive conjoint analysis	4
Ranking conjoint analysis	2
Rating conjoint analysis	10
Country	
Australia	4
Europe (excl. UK)	19
UK	5
Canada/USA	12
Other	5
Year published	
Pre-2000	1
2000 - 2005	7
2006 - 2009	15
2010-2012	22
Area of application	
Pharmaceuticals	14
Screening	6
Surgery	5
Process /outcomes of care	16
Other	5
Respondents in principal-agent relationshipsΔ	
Patients vs. clinicians	35 (59 experiments)
Patients vs. carers	5 (7 experiments)
Patients vs. public/lay	6 (8 experiments)
Patients vs. HLDM	1 (1 experiment)
Public vs. HLDM	1 (1 experiment)
Carers vs. clinicians	2 (2 experiments)
Clinicians vs. clinicians	6 (15 experiments)
Clinicians vs. HLDM	1 (1 experiment)
HLDM vs. HLDM	1 (1 experiment)
Attribute groups (Studies containing)	
Effectiveness (E)†	27
Safety (S)	25
Process-related (P)	34
Baseline disease (D)	4
Cost (C)	17
n/a	1

*Some studies included more than one experiment
†Includes quality of life and functional status
ΔSome studies included more than two respondent-types
DCE=Discrete choice experiment
HLDM=High level decision-maker (e.g. policy makers, HTA committee members, hospital managers)

Table 2.2. Cross tabulation of attribute group ranking

Attribute group	Number of experiments with attribute groups ranked same or different [Study reference numbers]							
	Patients vs. Clinicians		Patients vs. Carers or public		Clinicians vs. other Clinicians		Others vs.. Others	
	Same	Different	Same	Different	Same	Different	Same	Different
Effectiveness vs. other attribute types	11 [1,4,16,19,27,34, 38,39,40]	15 [8,12,15,19,20, 21,23,29,30,32, 37,40]	1 [36]	2 [5]	2 [39,40]	2 [32,40]	1 [41]	2 [41,45]
Effectiveness vs. other benefits	11 [19,20,23,29,32, 37,40]	2 [30,32]	1 [36]	1 [35]	2 [40]	1 [32]	1 [45]	0
Effectiveness vs. safety	11 [1,4,16,21,34, 39,40]	11 [10,12,14,19,20, 29,30,33,37]	0	2 [5]	3 [39,40]	1 [32]	0	0
Safety vs. other attribute types	14 [1,4,6,9,10,11, 12,16,34,39,40]	19 [8,10,12,17,18, 19,20,21,29,30, 32,37,39,40]	1 [6]	3 [5,10]	1 [43]	4 [18,32,39,43]	0	0
Safety vs. other safety	8 [12,18,19,21,29, 32]	8 [6,12,15,17,20, 30,32,37]	0	1 [6]	1 [18]	1 [32]	0	0
Disease vs. other attribute types	2 [31]	0	1 [31]	0	1 [31]	0	0	1 [45]
Disease vs. other disease	4 [26]	9 [2,26,31]	2 [26]	2 [26,31]	3 [28,34]	8 [2,28]	0	1 [45]
Process vs. other attribute types	14 [1,4,6,9,11,16, 27,31,37,38]	21 [8,10,17,18,19, 20,21,23,30,32, 39,40]	4 [6,10,31,36]	4 [5,13,25]	2 [31,40]	4 [18,32,33,40]	2 [41,43]	2 [41,42]

Attribute group	Number of experiments with attribute groups ranked same or different [Study reference numbers]							
	Patients vs. Clinicians		Patients vs. Carers or public		Clinicians vs. other Clinicians		Others vs.. Others	
	Same	Different	Same	Different	Same	Different	Same	Different
Process vs. other process	11 [9,11,19,31,37 38,39]	15 [7,9,10,18,19,20, 23,24,27,28,30, 32,41]	2 [25,31]	4 [5,10,13]	3 [32,39]	1 [18]	1 [42]	3 [41,43]
Process vs. (health outcomes)	17 [1,4,6,9,10,11, 27,31,37,38]	17 [8,17,18,19,20, 21,23,30,32,39, 40,41]	2 [10]	2 [5]	2 [31,40]	3 [18,39,40]	1 [41]	1 [41]
Cost vs. other attribute types	7 [1,32,34]	6 [10,15,39,40]	0	2 [10,13]	2 [32,40]	3 [18,39,40]	0	1 [42]
[Study reference numbers] refer to study reference numbers in Table 2.3								

2.4.2 Importance of attribute groups

Across all experiments, in all principal-agent pairs, the rank order of importance of the attribute groups included in each experiment differed for principals and agents in 71 (75%) experiments, was the same in 17 (18%) and was unclear in the remainder. Table 2.2 (p43) provides a cross tabulation of the frequency with which the rank order of attribute groups was the same or different among different principal-agent pairs, and Table 2.3 (p 53) provides a summary of the data from each published study.

For patients and clinicians, in experiments that included multiple attribute groups, the rank order of Effectiveness, Safety and Process-related attributes differed in 58-60% of instances. In experiments involving multiple Effectiveness attributes, the rank order of importance among these was the same for patients and clinicians in 11/13 (85%) experiments. In contrast, in experiments involving multiple Safety attributes or multiple Process-related attributes, the rank order of these among patients and clinicians was the same in 8/16 (50%) and 11/26 (42%), respectively.

2.4.3 Effectiveness and Safety (Benefit-Risk) trade-offs of patients and clinicians

There was no statistically significant difference in the frequency with which patients and clinicians ranked any attribute group as being the first most important attribute ($p>0.1$ for all proportions). There was also no statistically significant difference in the proportion of Effectiveness attributes and Safety attributes ranked among the top three most important attributes for patients (83% versus 58%, Yates corrected $X^2 = 1.67$, 1 d.f., $p>0.1$) or clinicians (80% versus 62%, Yates corrected $X^2 = 0.909$, 1 d.f., $p>0.1$) (see Figure 2.2, p46).

2.4.4 Health- and Process-related attribute trade-offs of patients and clinicians

Patients were statistically significantly more likely to rank Health-related attributes among their top three most important attributes than they were Process-related attributes (64% versus 42%, Yates corrected $X^2 = 4.55$, 1 d.f., $p<0.05$), as were clinicians (69% versus 44%, Yates corrected $X^2 = 4.93$, 1 d.f., $p<0.05$) (see Figure 2.3, p46).

Figure 2.2. Rank order of importance of Effectiveness and Safety attribute groups among Patients and Clinicians

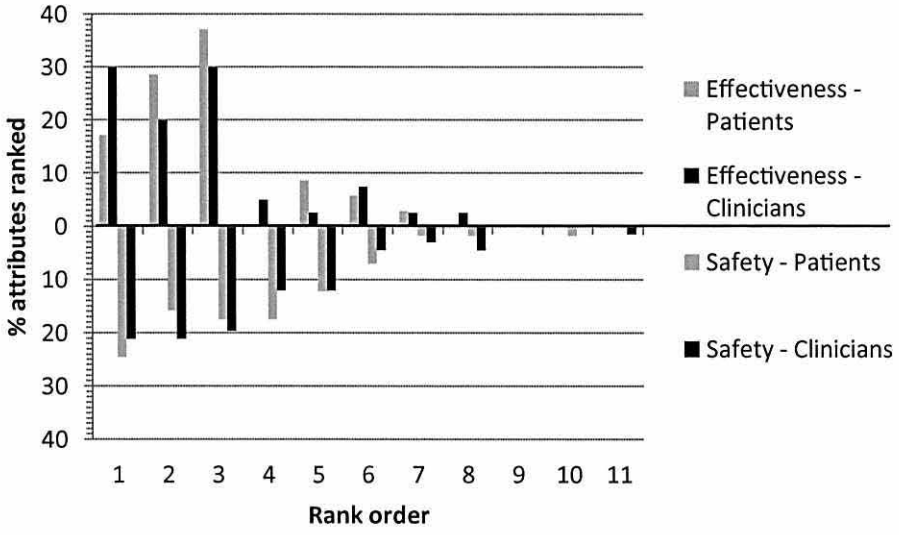
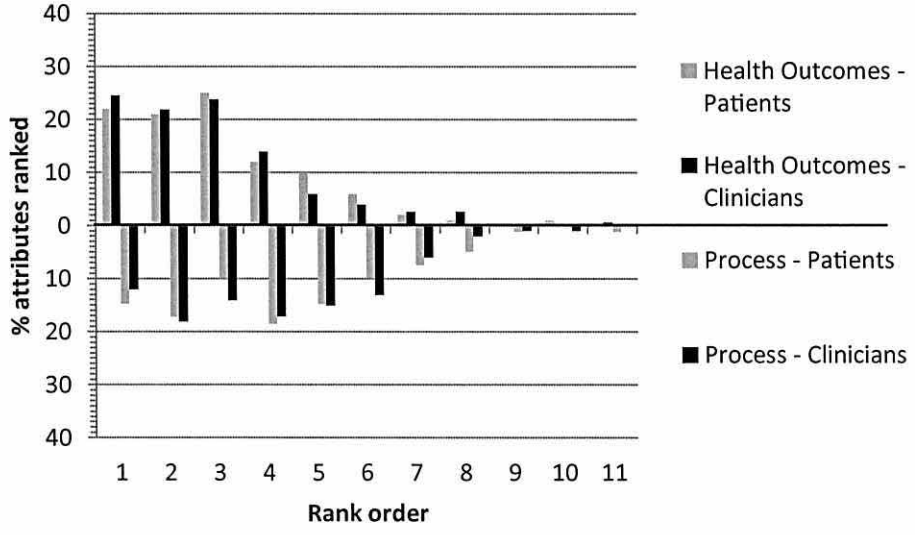


Figure 2.3. Rank order of importance of Health outcomes and Process attribute groups among Patients and Clinicians



2.4.5 Other attribute groups and other principal-agent pairs

There were fewer experiments involving Disease-or Cost-related attributes. The rank order of Disease attributes relative to other Disease-attributes differed between patients and clinicians in 9/13 (69%) experiments, and between clinicians and other clinicians in 8/11 (73%) experiments.

Only four experiments included high-level decision-maker (HLDM) respondents such as policy-makers, HTA committee members, and hospital managers (Table 2.3 study refs: 2, 38, 44, 45), One compared public and HLDM preferences for public reimbursement of pharmaceuticals (Table 2.3 study ref: 45): members of the public ranked quality of life of drug treatment responders (Effectiveness attribute) followed by expected patient survival if untreated (Disease attribute) to be the most important attributes when deciding which medicines should receive public funding in Australia, compared with Pharmaceutical Benefits Advisory Committee (PBAC) appraisal committee members who ranked untreated patient quality of life (Disease attribute) followed by expected patient survival if untreated (Disease attribute) ahead of quality of life of drug treatment responders (Effectiveness attribute). However, the DCE completed by the HLDMs included an additional attribute related to uncertainty in the probability of response to treatment.

2.5 Discussion

2.5.1 General findings

The main finding of this systematic review is that the importance attached to key considerations in health care decision-making differed between principals and agents in 75% of the cases examined. If the results of these conjoint analyses can be extrapolated to real life contexts, then preference structures of decision-makers are likely to differ from those on whose behalf they act more often than not.

There are good reasons why patients, the public, and clinicians, as principals, may need other stakeholders (agents) to make decisions on their behalf. Health care is characterised by uncertainty in both the in the incidence of disease (necessitating health care) and also in the efficacy of its treatment in any given individual (Arrow 1963). It would be impossible for patients or their doctors to undertake robust assessments of all aspects of all possible treatments to determine their net clinical or economic value

(Claxton et al., 2009). Given their different remits, levels of knowledge, skills and experience that place agents in a position of responsibility for making decisions on behalf of other stakeholders, an expectation of perfect agency — in which the agent makes decisions that fully reflect the principal's preferences and needs, unaffected by their own (Mooney and Ryan 1993) — would seem unrealistic.

Therefore, the recent policy focus and efforts to include patients' and the public's views into decision-making seem justified, as the preferences of agents alone are unlikely to fully reflect the views of those affected by the decision they make.

2.5.2 Implications for patient-clinician relationships

As may be anticipated, empirical studies of the patient-clinician relationship were by far the most common among the identified conjoint analyses. Based on implied rank order of importance, patients and clinicians differed in the importance attached to disease/patient characteristics, the benefits and risks of health care interventions/services, and the non-health, process-related aspects of health care in the majority (58%-60%) of experiments that consider these attributes. Given the scale, or frequency with which preferences can differ, shared decision-making would seem to be imperative.

It has been suggested that for shared decision-making to happen, both patients and clinicians must commit to sharing information and decision-making responsibility (Coulter and Collins, 2011). There is evidence that patients make different choices when well informed, and a range of decision-aids have been shown to improve patients' knowledge of available treatment options and their expectations of possible benefits and harms, help them reach choices that are more consistent with their informed values, and increase participation in decision making (Stacey et al., 2012). A recent NICE Clinical Guideline makes evidence-based recommendations on involving patients in decisions and supporting adherence to prescribed medication (Nunes et al., 2009). However, perhaps a first step of patient-centred care should be to determine how and the extent to which individual patients wish to participate in health care decision-making. Whilst almost all patients desire information from clinicians, many may still prefer to delegate decision-making to the clinician (Flynn et al., 2006).

2.5.3 Implications for other stakeholders

Conclusions that can be drawn across other principal-agent type relationships, including patients and their carers, or patients and the public, are limited beyond the general observation that their preferences may differ. The observation that different types of clinicians have different views of disease/patient characteristics in the majority of experiments in which this was assessed could potentially have implications for (inappropriate) referrals from family/primary care clinicians to specialist clinicians (Bederman 2012). However, the findings of differences in the preferences of patients and clinicians have implications for other stakeholders and decision-makers.

Current decision-making processes used by NICE focus on the health outcomes of interventions in the form of quality-adjusted life years (QALYs) when determining their value. NICE's current *Guide to the methods of technology appraisal* (NICE 2008a), and the draft update to these (NICE 2012b), state that significant characteristics of healthcare technologies that have a value to people that is independent of any direct effect on health should be noted (NICE 2008a; NICE 2012b); however, it remains unclear how these are to be valued alongside QALYs when noted. QALYs have long been criticised for failing to capture the full range of impacts of health technologies and interventions (Mooney 1989; Rawles 1989; ECHOUTCOME 2013).

The present review finds that non-health, process-related aspects of health care are important to both patients and clinicians; the statistical significance of health- and process-related attributes among the included studies indicates both were willing to make trade-offs against health outcomes in order to secure non-health aspects of health care. An earlier literature review of conjoint analyses, which pre-dated most of the studies included in the present review, also observed the importance of non-health attributes to patients, but was unable to draw conclusions on the priority they should be given relative to health-related attributes (OHE 2007). The present review reveals that both patients and clinicians are significantly more likely to rank health outcomes-related attributes more importantly than non-health, process-related attributes across a range of health care services and interventions.

These findings, therefore, support proposals to consider the non-health benefits of medicines under the future value-based pricing system [Dept Health 2010a]; however, they suggest that health-related outcomes should still be given more weight in those considerations.

2.5.4 Strengths and Limitations

Several previous reviews have highlighted that the preferences of decision-makers and those affected by their decisions may differ (Montgomery and Fahey, 2001; Mulley et al., 2012); however this is the first systematic review that has attempted to characterise the scale of the differences that may exist in the preferences of principals and agents towards key decision-making considerations. This has been possible by focusing on studies employing conjoint analyses, which are grounded in random utility/consumer theory (Ryan et al., 2008), and permit estimation of the differential importance of characteristics (attributes) making up products and services. The inclusion only of studies that include both principals and agents expressing their preferences for the same attributes, under the same scenarios, provides for more robust comparisons than would be achieved across studies that solicit their preferences in separate exercises. By categorising attribute types and considering their implied rank order of importance it has been possible to make greater inferences on the relative importance of health outcomes and non-health, process-related characteristics of health care interventions and services than was possible in an earlier systematic review (OHE 2007). The collective findings have potential implications for decision-makers at all levels.

There are some potential limitations that must be acknowledged. Study screening and selection for inclusion was conducted in duplicate, independently by both authors; however, due to time and resource constraints, all data extractions were conducted by the first author (WGL) and would require further independent verification prior to submission of the manuscript for full publication. The exclusion of non-conjoint analysis type studies may exclude other sources of informative evidence. Conjoint analyses require respondents to express preferences for hypothetical scenarios, and none of the studies included in this review, and few others in the wider DCE literature (De Bekker-Grob et al., 2010), have tested the external validity of respondents' stated preferences.

It was not feasible to compare or combine directly the parameter estimates across different discrete choice models due to differences in scale and error variance. Therefore, the review focused on the implied rank order of importance of statistically significant attributes included among the different experiments. To achieve this, each attribute was categorised into a discrete group, which is a subjective exercise. In addition, all attributes assigned to a given group are implicitly assumed to have an equal chance of being ranked in any order so, for example, no distinction was made between different types of adverse events that may have very different consequences – they are all simply

categorised as safety attributes. That said, it would be inappropriate to impose a personal view on the importance or consequences of different adverse events in this type of exercise.

It is possible that the greater the number of attributes included in each choice alternative, the greater the potential for these to be ranked differently by respondents. The number of attributes among the experiments ranged from 2 to 11; however, it is of note that the rank order of importance differed between patients and clinicians in the experiment that included only two attributes (Table 2.3 study ref: 15), so it is difficult to draw conclusions in this regard or form a basis for adjusting the analysis to compensate for this. No studies explored all five attribute groups designated in this review– each experiment is designed with specific objectives in mind, and it is not possible to draw any conclusions on the appropriateness of the inclusion or omission of any particular attribute type in any given experiment. Exploration of cost or cost effectiveness considerations was limited due to different types of costs explored in the study (cost to tax payer versus individual out-of-pocket costs) and the need to use the coefficient of the cost attributes as the denominator for determining the relative importance of other attribute types via marginal rates of substitution for some studies.

The wide range of reporting methods and level of detail provided in the published studies presented challenges to data extraction, and precluded some studies from the analyses (e.g. Bejlinga et al., 2011). Few studies fully met all 10 of the checklist items suggested in the ISPOR good practice guide for conducting conjoint analyses (Bridges et al., 2011); however, it is of note that over half of the studies included in this review were conducted before good practices for conducting conjoint analyses had been agreed and published, which may have influenced reporting. Neither the ISPOR checklist, nor other relevant published conjoint analysis guides (Lancsar and Louviere, 2008), provides a hierarchical system for discriminating high quality studies from poor, such as those developed for assessing the risk of bias in clinical studies. Therefore, no studies were excluded on the basis of the extent to which they appeared to meet the ISPOR good practice checklist; however, studies that were underpowered, or generated seemingly implausible parameter estimates, were effectively excluded by default (e.g. Schmitz et al, 1994). As the only study to have inadequately reported on multiple checklist items (Mellor and Green, 2002) provides just one of 59 experiments conducted among patients and clinicians, its retention in the analyses is not anticipated to significantly bias the overall results of this review.

Finally, just as DCEs assume implicitly that the aggregate stated preferences of individual respondents represent the preferences of all respondents as a single entity, our approach to determining the proportion of attribute groups that were ranked as most important effectively aggregates respondent-types. It should be noted that even among those studies where the rank order of Effectiveness : Safety attributes or of Health outcome : Process attributes was the same for principals and agents, the rank order of other attributes included in the experiments often differed.

2.5.5 Conclusions

The preferences of principals and agents can differ at all levels of health care decision-making, supporting the initiatives to encourage decision-makers to actively engage and incorporate the views of those affected by their decisions. Patients and clinicians have different preference structures, particularly with respect to the order of importance attached to benefits and risks of health care services and interventions, although both were numerically more likely to rank benefits of health care interventions and services ahead of safety considerations. The importance of non-health, process-related aspects of health care is confirmed, lending support to proposals to consider the non-health benefits of medicines under the future value-based pricing system (Dept Health 2010a); however, the findings further suggest that health-related outcomes should still be given more weight in those considerations.

Table 2.3. Summary of the data from each published study

Study Ref. Author#, Year	Subject, Country, Study type	Patients' attributes in rank order of importance (unless otherwise stated)*†	Clinicians' rank order (unless otherwise stated)*†	Other stakeholders' rank order(unless otherwise stated)*†
1. Arana, 2006	Cervical cancer screening, Spain, DCE with opt-out (8 choice tasks)	Female students (service users/patients) (n=60): Test interval (P)> Probability of false positive (S)> Probability of cancer death (E)> Test cost (C)	Female physicians (n=60): Same as patients	
2. Bederman, 2010	Spinal surgery, Canada, Rating CA (16 choice tasks)	Patients (n=164): Pain severity (D1)> Walking ability (D2)> Pain duration (D3)> Pain onset (D4)> Neurological symptoms (D5)> Pain location (D6)	i) Family physicians (n=202): D5>D2>D1>D4>D6>D3 ii) Spinal surgeon (n=131): D6>D1>D2>D5>D3>D4	
3. Bijlenga, 2011	Obstetric outcomes, The Netherlands, DCE – Forced choice (9 choice tasks)	Patients (n=24): Unclear ordering - most attributes NS	Health professionals (n=30): Unclear ordering - most attributes NS	Public/lay (n=27): Unclear ordering - most attributes NS
4. Bishop, 2004	Antental screening, UK, DCE with opt-out (9 choice tasks)	Pregnant women (service users/patients) (n=253): Risk of miscarriage due to test (S)> Timing of test (P)> Detection rate (E)	Ostetricians/midwives (n=94): Same as patients	
5. Bridges, 2012	Male circumcision to prevent HIV, South Africa, DCE – Forced choice (6 choice tasks)	Young adult males (n=237): Follow-up required (P1)> Risk/benefit counselling (P2)> Lower infection rate (E)> Pain of surgery (S)> Private waiting room (P3)> Male staff (P4)(NS);>No booking required (P5)(NS)> HIV test required (P6)(NS)> Other attributes all NS		i) Carers: Fathers (n=204): S>P1>P5>P3>all others NS ii) Carers; Mothers (n=204): P1>E>P3>P6>all others NS
6. Cheung, 2012	Vestibular schwannoma treatment, USA, Ranking CA (9 profiles)	Patients (n=61): Risk of permanent deafness (S1)> Risk of temporary facial nerve damage (S2)> Long term cancer risk (S3)> Duration of recovery (P)	Surgeons (n=60): S3>S2>S1>P	Lay/prospective patients (n=61 + n=74): S1>S3>S2>P
7. Davison, 2010	Chronic kidney disease programmes, Canada, DCE – Forced choice (12 choice tasks)	Patients/caregivers (n=198): Donor source (P1)> Care provider (P2)> Basis of donor allocation (P3)> Advanced care planning (P4)> Extent of information provision (P5)> Decision to stop dialysis (P6)	Renal specialist/GPs (n=150): P1>P4>P3>P5>P2>P6	
8. De Bekker-Grob, 2009	Drug prevention of osteoporosis, The Netherlands, DCE with opt-out (16 choice tasks)	Female patients (n=117) Nausea (S)> Administration route and regimen (P1)> 10-yr hip fracture reduction (E)> Treatment duration (P2)> [Cost (C) used for MRS]	GPs (n=39): E>S>P1>P2; [C]	
9. Espelid, 2006	Dental restoration, Norway & Denmark, DCE with opt-out	Norwegian patients (n=196, Danish patients (n=110) (6 choice tasks): Appearance (P1)> Longevity of dental restoration (P2)> Risk of allergic reaction (S)	Norwegian dentists (n=22), Danish dentists (n= 20) (18 choice tasks): P2>P1>S Norwegian dental assist. (n=18),	

Study Ref. Author#, Year	Subject, Country, Study type	Patients' attributes in rank order of importance (unless otherwise stated)*†	Clinicians' rank order (unless otherwise stated)*†	Other stakeholders' rank order(unless otherwise stated)*†
			Danish dental assist. (n=47) (18 choice tasks): Same as patients	
10. Faggioli, 2011	Vascular surgery for aortic aneurysm, Italy, DCE – Forced choice	Treated patients (n=141, 4 choice tasks): Additional costs (C)> Risk of complications (S)> 5-yr risk of repeat (P1)> Recovery time (P2)> Check up requirements (P3)(NS)	Specialist physicians (n=30, 8 choice tasks): S>P1>P3>P2(NS)>C(NS)	Carers (n=78 + n=54, 8 choice tasks): S>C>P1>P3(NS)>P2(NS)
		Untreated patients (n=84, 4 choice task): S>C>P1>P2>P3(NS)		
11. Fiebig, 2009	Cervical cancer screening, Australia, DCE with opt-out (16 choice tasks)	Pap screen experienced women (n=167): Time since last tests (P1)> Test interval (P2)> Probability of false positive (S)> Cost (C)	GPs (n=215): Same as patients	
12. Gregorian, 2010	Opioid drug treatment of pain, USA, ACA (20 choice tasks)	Chronic pain patients (n=316): Incidence of vomiting (S1)> Pain relief achieved (E)> Incidence of nausea (S2)> Incidence of constipation (S3)> Incidence of pruritis (S4)> Incidence of drowsiness (S5)	Chronic pain physician (n=163): S1>S2>E>S5>S3>S4	
		Acute pain patients (n=302): Incidence of vomiting (S1)> Pain relief achieved (E) = Incidence of nausea (S2)> Incidence of constipation (S3) = Incidence of pruritis (S4)> Incidence of drowsiness (S5)	Acute pain physician (n=162): S1>S2>E>S3>S5=S4	
13. Hendrix, 2010	Obstetric care, The Netherlands, DCE – Forced choice (7 choice tasks)	Nulliparous women (n=231): Ability to influence decisions (P1)> Birth setting (P2)> Place of birth (P3)> Assistance provider (P4)> Out of pocket cost (C)> Pain relief possible (P5)		Carers: Partners (n=212): P1>P5>C>P4>P3>P2
14. Johnson, 2010	Crohn's disease drug treatment, USA, DCE – Forced choice (9 choice tasks)	ORs for gastroenterologists (n=315) vs. patients (n=580), not rank order: Severe daily symptoms (E) 0.83; Moderate daily symptoms (E) 1.12; Mild daily symptoms (E) 1.10; Prevent serious complications (E); Need for oral steroids (E) 0.95; 10-yr Risk of death/disability from PML 2%(S) 1.14; 10-yr Risk of death/disability from PML 5% (S) 0.89; 10-yr Risk of death from serious infections 2% (S) 1.20; 10-yr Risk of death from serious infections 5% (S) 0.76; 10-yr Risk of cancer 2% (S) 0.87; 10-yr Risk of cancer 5% (S) 0.91	n/a	
15. Lee, 2005	Post-operative symptoms, China, DCE – Forced choice (7 choice tasks)	Gynaecological surgery patients (n=200): Risk of post-op nausea/vomiting (S1)> Pain level (S2)> Sedation level (S3)(NS)	Anaesthetists/nurses (n=52): S1>S3>S2	
	Antiemetic treatment, China, DCE – Forced choice	Gynaecological surgery patients (n=200): Efficacy of antiemetic (E)> Extra cost (C)	Anaesthetists/nurses (n=52): C>E	
16. Lewis, 2006	Antinatal screening, Australia, DCE – Forced choice (9 choice tasks)	Pregnant women (n=113): Risk of miscarriage (S)> Timing of test (P)> Detection rate (E)	Ostetricians/midwives (n=175): Same as patients	
17. Longacre,	Drug/procedure for	Cirrhosis patients (n=53):	Gastroenterologists /	

Study Ref. Author#, Year	Subject, Country, Study type	Patients' attributes in rank order of importance (unless otherwise stated)*†	Clinicians' rank order (unless otherwise stated)*†	Other stakeholders' rank order(unless otherwise stated)*†
2008	prophylaxis of variceal haemorrhage in cirrhosis patients, USA, ACA (? choice tasks)	Shortness of breath/low blood pressure (S1)> Fatigue (S2)> Procedure-related bleed (S3)> Route of administration (P)> Sexual problems (S4)> Pain/difficulty swallowing (S5)> Risk of perforation (S6)	Hepatologists (n=61): S3>S1>S4>S6>S2>P>S5	
18. Mantovani, 2005	Drug prophylaxis for haemophilia Italy, DCE – Forced choice (9 choice tasks)	Adult haemophilia patients (n=178): Perceived viral safety (S1)> Distribution by community pharmacy (P1)> Distribution home (P2)> Risk of inhibitor development (S2)> Infusion frequency (P3)> [Cost (C) used for MRS]; Need for reconstitution(P4)(NS)	i) Physicians (n=69): S1>P2>S2>P1>P3; [C]; P4(NS) ii) Pharmacists (n=58): S1>S2>P4>P3>P1>P2; [C]	
19. Marshall, 2009	Colorectal cancer screening, Canada & USA, DCE with opt-out (11 choice tasks)	Canadian population sample (prospective service users) (n=454): Sensitivity (E1)> Specificity (E2)> Pain (S1)> Procedure (P1)> Complication rate (S2)> Preparation (P2)> Frequency (P3)> Follow up required (P4); [Cost (C) used for MRS] USA population sample (prospective service users) (n=961): Sensitivity (E1)> Pain (S1)> Specificity (E2)> Complication rate (S2)> Procedure (P1)> Preparation (P2)> Frequency (P3)> Follow up required (P4); [Cost (C) used for MRS]	Canadian physiciansΔ (n=100): E1>E2>P3>S1>S2>P2>P1>P4; [C] USA physiciansΔ (n=99): E1>E2>S1>P1>P2>P3>S2>P4; [C]	
20. Mellor, 2002	Drug prescribing decisions, UK, Rating CA (? rating tasks)	Patients (n=170): Dosing schedule (P1)> Long-term side effects (S1)> Efficacy indicator (E1)> Short-term side effects (S2)> Durability of response (E2)> Psychotic effects (S3)> Dosage convenience (P2)> Metabolic indicator (S4)	Physicians (n=170): S2>S1>S4>P2>P1>E1>E2>S3	
21. Mohamed, 2012	Hepatitis B drug treatment, Turkey, DCE – Forced choice (12 choice tasks)	Patients (n=117): How long medication studied (P)> Probability viral load undetectable (E)> [Unclear rank: 5-yr risk of renal insufficiency (S1); Cost (C)]> 5 year risk fracture (S2)	Physicians (n=159): E>S1>[Unclear rank: P; C]>S2	
22. Morton, 2012	Dialysis in chronic kidney disease, Australia, DCE with opt-out (12 choice tasks)	Adult CKD patients (n=105): ORs for Home dialysis vs. Conservative care, not rank order: Survival (E) 1.68; Travel restrictions (P) 0.37; Nocturnal dialysis (P) 0.07; Time on dialysis (P) 1.37(NS); Dialysis during day/evening (P) 1.28(NS)		Carers (n=73): Survival (E) 1.82; Travel restrictions (P) 0.43; Nocturnal dialysis (P) 0.03; Time on dialysis (P) 1.64; Dialysis during day/evening (P) 7.90
	Dialysis in chronic kidney disease, Australia, DCE with opt-out (12 choice tasks)	Adult CKD patients (n=105): ORs for In-centre haemodialysis vs. Conservative care, not rank order: Time on dialysis (P) 2.02; Dialysis during day/evening (P) 1.28; Travel restrictions (P) 0.34; Nocturnal dialysis (P) 0.23		Carers (n=73): Time on dialysis (P) 2.02; Dialysis during day/evening (P) 7.90; Travel restrictions (P) 0.18; Nocturnal dialysis (P) 0.04
23. Mühlbacher, 2011	Multiple myeloma therapy, Germany, DCE – Forced choice (8 choice tasks)	Patients (n=282): Possibility of further lines of therapy (P1)> Survival gain (E1)> Emotional quality of life (E2)> Periods between treatment lines (P2)> Administration (P3)> Physical quality of life (E3)	Physicians (n=243): E1>P1>P3>E2>P2>E3	

Study Ref. Author#, Year	Subject, Country, Study type	Patients' attributes in rank order of importance (unless otherwise stated)*†	Clinicians' rank order (unless otherwise stated)*†	Other stakeholders' rank order(unless otherwise stated)*†
24. Neuman, 2007	Maternity ward attributes, Israel, DCE – Forced choice	Women who gave birth (n=323, 6 or 7 choice tasks): Staff professionalism (P1)> Staff attitude (P2)> Information provided (P3)> Travel time (P4)> Number of beds in room (P5)(NS)	Medical staff (n=30, 12 choice tasks): P5>P1>P2>P4>P3	
25. Opuni, 2010	Antiretroviral therapy clinics, South Africa, DCE – Forced choice (20 choice tasks)	HIV patients (n=510): Staff rude (P1)> Staff indifferent (P2)> Clinic brand clear (P3) = Price of chosen clinic (C1) > Wait time (P4) > Cllinic brand discrete (P5) > Price alternative clinic1 (C2)= Price alternative clinic2 (C3)		Publi/lay (n=777): P1>P2>C1>P3>P4>P5>C2 = C3
26. Oudhoff, 2007	Surgical waiting lists (Varicose veins), The Netherlands, Rating CA (9 tasks)	Patients (n=60): Physical symptoms (D1)> Impairments in work (D2)> Social limitations (D3)> Psychological distress (D4)	i) Surgeons (n=61): D2>D1>D3>D4	Public/lay (n=74): Same as patients
			ii) Occupational health physicians (n=50): D2>D1>D3>D4	
			iii) GPs (n=34): Same as patients	
	Surgical waiting lists (Inguinal hernia), The Netherlands, Rating CA (9 tasks)	Patients (n=72): Impairments in work (D1)> Physical symptoms (D2)> Psychological distress (D3)> Social limitations (D4)	i) Surgeons (n=59): D1>D2>D4>D3	Public/lay (n=56): D1>D2>D4>D3
			ii) Occupational health physicians (n=51): D1>D3>D4>D2	
			iii) GPs (n=32): Same as patients	
	Surgical waiting lists (Gallstones), The Netherlands, Rating CA (9 tasks)	Patients (n=65): Physical symptoms (D1)> Impairments in work (D2)> Social limitations (D3)> Psychological distress (D4)	i) Surgeons (n=59): Same as patients	Public/lay (n=72): Same as patients
			ii) Occupational health physicians (n=54): D2>D1>D3>D4	
			iii) GPs (n=42): Same as patients	
27. Payne, 2011	Pharmaco-genetic testing services, UK, DCE – Forced choice (16 choice tasks)	Patients (n=159) Level of information (P1)> Results provider (P2)> Predictive ability of test (E)> Time to results (P3)	Health care professionsal (n=139): P2>P1>E>P3	
28. Pedersen, 2012	General practice services (2 alternative GP practices), Denmark, DCE – Forced choice (16 choice tasks)	Patients (n=698): Time to appointment (P1)> Distance (P2)>Telephone wait time (P3)> Extended opening hours (P4)> Length of consultations (P5)> Routine task performer (P6)> Waiting room time (P7)	GPs (n=969): P4>P6>P1>P2>P5>P3>P7	
	General practice services (own GP practice or one of 2 alternatives), Denmark, DCE – Forced choice (16	Patients (n=698): Extended opening hours (P1)> Routine task performer (P2);>Time to appointment (P3)> Distance (P4)> Telephone wait time (P5)> Length of consultations (P6)> Waiting room time (P7)	GPs (n=969): P1>P2>P3>P5>P6>P4>P7	

Study Ref. Author#, Year	Subject, Country, Study type	Patients' attributes in rank order of importance (unless otherwise stated)*†	Clinicians' rank order (unless otherwise stated)*†	Other stakeholders' rank order(unless otherwise stated)*†
	choice tasks)			
29. Pieterse, 2007	Pre-operative radiotherapy for rectal cancer, The Netherlands, ACA (14 tasks)	Disease-free patients (n=66): Probability of faecal incontinence (S1)> 5-yr risk recurrence (E1)> 5-yr survival (E2) = Probability of sexual dysfunction (S2)	Oncologists (n=60): E1=S1>E2>S2	
30. Porzolt, 2010	Drug treatment for diabetes mellitus, Germany; ACA (? tasks)	Patients with diabetes (n=827): No weight change (S1)> Weight loss (S2)> Reduce HbA1c (E1)> Absence of side effects (S3)> No limitations to daily life (E2)> Improved well being (E3)> Branded products (P1)> Flexible administration (P2)> Administration linked to meals (P3)> Weight gain (S4)> Generic products (P4)	GPs and diabetologists (n=60): S2>E1>S1>E3>P4>S3>P3>E2>P 1>P2>S4	
31. Sampietro-Colom, 2008	Elective orthopaedic waiting list, Spain, Ranking CA (? tasks)	Patients & Relatives (n=347): Pain (D1)> Disease severity (D2)> Difficulties with activities of daily living (D3)> Work limitations (D4)> Has carer (P1)> Is a carer (P2)> Probability of recovery (D5)	i) Allied health professional (n=117): D1>D3>D2>D4>P1>P2>D5	Public/lay (n=300): D2>D1>D3>D4>P1>P2 >D5
			ii) Consultant (n=86): D1>D3>D2>D4>P1>P2>D5	
32. Scalone, 2009	Drug treatment for haemophilia, Italy, DCE – Forced choice (16 choice tasks)	Patients/ caregivers (n=37): Tax-payer cost (C)> Risk of anamnestic response (S1)> Major surgery possible (S2)> Perceived risk of viral infection (S3)> Time to stop bleeding (E1)> Frequency of infusions for prophylaxis (P1)> Time to pain recovery (E2)> Number of infusions to stop bleeding (P2);	i) Physicians (n=39): C>S1>P2>S2>E2>E1>P1>S3	
			ii) Pharmacists (n=25): C>E1>E2>S2>S3>P2>P1>S1	
33. Schmitz, 1994	Methadone maintenance clinic privileges, USA, Ranking CA (105 tasks)	Patients (n=12): Reported for patients – not for counsellors	Counsellors (n=4): Not reported - unclear	
34. Shafey, 2011	Second-line lymphoma treatment, Canada, DCE – Forced choice (17 choice tasks)	Patients (50% experienced relapse) (n=81): Toxicity (S)> Remission length (E)> Cost to health system (C)(NS)	Physicians (n=48): Same as patients	
35. Stenek, 2000	Heart failure treatment outcomes, USA, Ranking CA (16 tasks)	Patients (n=51): Survival (E1) > Tiredness (E2) > Shortness of breath (E3) = Depression (E4)		Pubic/lay (n=47): E4>E2>E1>E3
36. Sung, 2012	Febrile neutropenia treatment in cancer, Canada, DCE with opt-out (12 choice tasks)	i) Adult patients (n=78): Mortality risk (E1)> Risk of ICU admission (E2)> Risk of readmission (E3)> Frequency of clinic visits(P)		
		ii) Child patients (n=43): Mortality risk (E1)> Risk of ICU admission (E2)> Risk of readmission (E3)> Frequency of clinic visits(P)		Carers: Parents (n=153): Same as patients
37. Thrumurthy, 2011	Oesophagastriic cancer surgery, UK, DCE – Forced choice (n=20 choice tasks)	Post-op patients (n=81): Quality of life (E1)> Cure rate (E2)> Post-op morbidity (S1)> Surgeon reputation (P1)> Post-op mortality (S2)> Hospital type (P2)	Physicians (n=90): E1>S2>E2>P1>S1>P2	

Study Ref. Author#, Year	Subject, Country, Study type	Patients' attributes in rank order of importance (unless otherwise stated)*†	Clinicians' rank order (unless otherwise stated)*†	Other stakeholders' rank order(unless otherwise stated)*†
38. van Empel, 2011	Fertility clinic process and outcome, Belgium & The Netherlands, DCE – Forced choice (16-17 choice tasks)	Fertility patients (n=925): Pregnancy rate (E)> Physician attitude (P1)> Information (P2);>Travel time (P3)> Continuity of care (P4)	Fertility expert physicians (n=227): Same as patients	
39. Vermeulen; 2007	Wound dressings, The Netherlands, Rating CA (18 vignettes)	Patients (n=74): Pain on dressing change (S)> Hospital duration (P1)> Time to heal (E)> Costs (C)> Frequency of changes (P2)> Patient changes dressing (P3)> Carer changes dressing (P4)	i) Nurses (n=150): S>P1>E>P2>C>P3>P4 ii) Surgeons (n=50): P1>S>E>C>P2>P3>P4	
40. Wellman, 2003	Depression drug treatment (For hypothetical patient), USA, Rating CA (18 choice scenarios)	Patients visiting HMO provider network (n=101): Functional status improvement (E1)> Symptom improvement (E2)> Patient co-payment (C)> Risk of moderate-severe side effects (S)> Frequency of administration (P)	i) Physicians (n=101): E1>E2>P>S>C ii) Pharmacists (n=77): E1>E2>S>P>C	
	Depression drug treatment (For real patient/self), USA, Rating CA (18 choice scenarios)	Patients visiting HMO provider network (n=101): Patient co-payment (C)> Functional status improvement (E1);>Symptom improvement (E2)> Risk of moderate-severe side effects (S)> Frequency of administration (P)	i) Physicians (n=101): E1>C>E2>P>S	
			ii) Pharmacists (n=77): C>E1>E2>P>S	
41. Youngkong, 2010	HIV/AIDS intervention programme, Thailand, DCE – Forced choice (16 choice tasks)	HIV/AIDS patients (n=74): Gender (P1)> Type of intervention (P2)> Effectiveness (E)> Quality of evidence (P3)> Target group (P4)	Community Health volunteers (n=50): E>P4>P1>P2>P3	Policy-makers (n=28): E>P2>P3>P4>P1
Other stakeholder comparisons				
Study Ref. Author#, Year	**Subject, Country, Study type**	**Principal's attributes in rank order of importance (unless otherwise stated)*†**	**Agents' rank order**	
42. Bech, 2003	Hospital reimbursement scheme, Denmark, DCE – Forced choice (4 choice tasks)	Hospital managers (n=92): Incentive to increase patient numbers (P1)> Treatment quality (P2)> Budget safety (C)> Other attributes unclear importance	Politicians (n=57): P1>C>P2> Other attributes unclear importance	
43. Gidman, 2007	Paediatric day case surgery, UK, DCE – Forced choice (8 choice tasks)	Parents of paediatric patients (n=280): Recovery experience (P1)> Shared decision-making (P2)> Parents present at induction (P3)> Staff attitude (P4)> Doses of Pain relief required (P5); [Cost (C) used for MRS]	Anaesthetists (n=193): P2>P1>P3>P4>P5; [C]	
44. Soinin, 2012	Willingness to pay for hypothetical new treatments, Finland, DCE – Forced choice (5 choice tasks)	Complex reporting: No statistically significant differences in WTP for different treatments between clinicians (n=146) and politicians (n=73)		
45. Whitty, 2011¶	Funding of drugs, Australia, DCE with opt-out	Public (n=161; 27 choice tasks): Quality of life for responders (E1)> Baseline survival (D1)> Baseline	Appraisal committee members (n=11; 48 choice tasks): D2>D1>E1>E2>E3; [C]	

Study Ref. Author#, Year	Subject, Country, Study type	Patients' attributes in rank order of importance (unless otherwise stated)*†	Clinicians' rank order (unless otherwise stated)*†	Other stakeholders' rank order(unless otherwise stated)*†
		quality of life (D2);> Length of survival for responders (E2)< Probability of treatment response (E3); [Cost (C) used for MRS]		

#Only first author listed
* Order of importance based on reported order where stated, or implied order based on marginal rate of substitution using cost as denominator where possible
†All attributes statistically significant unless stated (NS). Attributes that are NS for all respondent types not presented.
Attribute groups: (C)=Cost-related; (D)= Disease-related; (E) = Effectiveness-related; (P)= Process-related or non-health-related personal preference; (S)= Safety-related
Δ Clinicians reporting their expectation of patients' preferences, rather than their own professional preferences
¶DCE for decision-makers included additional attribute of uncertainty, models not directly comparable
GPs=General practitioners

2.6 Authors' contributions and Funding

WGL (the candidate) and DAH (supervisor) conceived the study. WGL designed and tested the search strategy and ran the searches. WGL performed initial screen for study eligibility, and obtained full papers. WGL and DAH performed full screen of papers independently and designed abstraction proforma. WGL abstracted data, categorised attributes, analysed data and drafted the manuscript. WGL and DAH revised the manuscript for intellectual content. WGL finalised the manuscript.

Funding for the project was provided by a Bangor University PhD studentship awarded to WGL. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Chapter 3

Revealed preference study of All Wales Medicines Strategy Group decision-making

Published as:

Linley WG, Hughes DA. Cost-effectiveness, clinical factors and policy influences on reimbursement decisions of the All Wales Medicines Strategy Group. *Pharmacoeconomics* 2012; 30 (9): 779-94.

3.1 Abstract

Background: There have been several explorations of factors influencing the reimbursement decisions of the National Institute for Health and Clinical Excellence (NICE) but not of other UK-based Health Technology Assessment (HTA) organisations.

Objective: This study aimed to explore the factors influencing the recommendations of the All Wales Medicines Strategy Group (AWMSG) on the use of new medicines in Wales.

Methods: Based on public data, logistic regression models were developed to evaluate the influence of cost-effectiveness, the quality and quantity of clinical evidence, disease characteristics (including rarity), budget impact, and a range of other factors on the recommendations of AWMSG, and its sub-committee, the New Medicines Group (NMG).

Results: Multivariate analyses of 47 AWMSG appraisals between 2007-9 correctly classified 87% of decisions. The results are suggestive of a positive influence on recommendations of the presence of probabilistic sensitivity analyses (PSA) but, counter-intuitively, a statistically significant negative influence of evidence from high quality randomised controlled trials (RCTs) (odds ratio 0.059, 95%CI 0.005, 0.699). This latter observation may be attributed to our strict definition of high quality, which excluded the use of surrogate endpoints. Putative explanatory variables, including cost-effectiveness, budget impact, underlying disease characteristics and ultra-orphan drug status were not statistically significant predictors of final AWMSG decisions based on our data set. Univariate analyses indicate that medicines with negative recommendations had significantly higher incremental cost effectiveness ratios, consistent with the pursuit of economic efficiency. There is also evidence that AWMSG considers equity issues via an 'ultra'-orphan drugs policy.

Conclusions: Consideration of decision uncertainty via PSA appears to positively influence reimbursement decisions of AWMSG. The significant negative impact of the presence of high quality RCTs, and the lack of a significant positive impact of other expected factors, may reflect issues in the plausibility of supporting evidence for medicines that received negative recommendations. Further, it serves to emphasise the difficulties in applying the usual hierarchies of evidence to the HTA process, and in particular to the appraisal of high cost, specialist medicines close to market launch.

3.2 Introduction

The All Wales Medicines Strategy Group (AWMSG) is a Welsh Government-funded body whose remit includes the appraisal of new medicines for which guidance by the National Institute for Health and Clinical Excellence (NICE) is not imminent (DHSS 2007). There are parallels in the appraisal processes of AWMSG and NICE for individual medicines (AWMSG 2007; NICE 2009a), and the requirement for local NHS organisations in Wales to implement guidance within three months of publication (DHSS 2007). However, the topic selection criteria of AWMSG and NICE differ.

Prior to October 2010, the focus of the AWMSG appraisal programme was new medicines for the treatment of cancer and cardiovascular disease, and certain other high cost medicines (greater than £2,000 per patient, per year). However, the programme has since expanded to include all new medicines, excluding those for which NICE guidance for the same indication is expected within 12 months. In contrast to NICE, the AWMSG routinely appraises "ultra"-orphan drugs (defined by AWMSG as medicines with orphan-designated status by the European Medicines Agency for diseases with prevalence of less than 1 in 50,000 population) [AWMSG 2011b] and other medicines for use in specialist disease areas (e.g. human immunodeficiency virus infection, HIV), which in England may fall under the remit of regional or national specialised commissioning groups [NCGHSS 2010]. Therefore, although AWMSG guidance is superseded by subsequent NICE guidance should it be issued (DHSS 2007), the AWMSG programme has the potential to provide more timely national guidance to NHS Wales for some medicines within the current remit of NICE and national coverage decisions for those medicines outwith NICE's remit.

The AWMSG appraisal process requires pharmaceutical companies to submit a dossier of evidence in support of their new medicine, from the payer perspective of NHS Wales and Personal and Social services. This submission is subjected to expert technical review, and wider stakeholder input is obtained by inviting independent medical expert opinion and relevant patient organisations to submit information. These are considered collectively by a New Medicines Group (NMG), which is a multidisciplinary sub-committee of AWMSG charged with making a preliminary recommendation based on evidence of clinical and cost-effectiveness (AWMSG 2007). AWMSG considers the preliminary recommendation of NMG alongside wider societal factors and budget impact in making its final recommendation to the Minister for Health and Social Services (AWMSG 2007).

Once ratified, the medicine may be recommended for routine use within its appraised licensed indication, for restricted use, or not for use within NHS Wales (AWMSG 2007).

There is an increasing body of evidence on the factors that influence policy-makers' decisions on the reimbursement or approval of new medicines (George et al., 2001; Grégoire et al., 2001; PausJenssen et al., 2003; Tilson et al., 2010; Mason et al., 2010; Chim et al., 2010), including the decisions of NICE in the UK (Devlin and Parkin, 2004; Dakin et al., 2006; Tappenden et al., 2007). Based on logistic regression modelling of 33 available recommendations, Devlin & Parkin (2004) concluded that the incremental cost effectiveness ratio (ICER), the uncertainty surrounding the ICER, the number of patients affected by the disease and the availability of alternative therapy significantly influenced the decision of NICE to accept or reject a health technology. Dakin et al. (2006) also identified the number of supporting randomised controlled trials (RCTs) and the presence of a submission from a patient organisation to be significant determinants of NICE recommendations, based on 60 appraisals. In addition, a stated preference experiment conducted in members of the NICE appraisal committees found baseline health-related quality of life of potential beneficiaries to be a significant influence on the decision to recommend health technologies (Tappenden et al., 2007).

This paper adds to the literature on decision-making by health technology appraisal (HTA) organisations in the UK by exploring the factors that influence AWMSG recommendations for new medicines in Wales. We aimed to build upon the factors identified previously to influence decisions, using AWMSG-specific factors in relation to policy, and a novel method for considering the influence of health burden of potential beneficiaries. Previous works in this area have focused on the final recommendations of appraisal bodies (Devlin and Parkin, 2004; Dakin et al., 2006). Our data set also provides a useful insight to the preliminary recommendations of NMG, which inform the final AWMSG recommendations.

3.3 Methods

Data from all 80 AWMSG Final Appraisal Reports and minutes from associated public meetings, issued up to December 2009, were accessed from the AWMSG website (http://www.wales.nhs.uk/awmsg). Permission was sought from pharmaceutical companies to access non-public submission documentation. However, this was not granted in many cases which, combined with issues of consistency in data reporting before creation of the NMG in March 2007, meant our sample was based on publically available data from 60 submissions with recommendations made between March 2007

and December 2009. The submissions covered 54 different medicines, and included five for multiple indications and one re-submission for the same indication (see Appendix to Chapter 3). Univariate analyses were based on this sample. Data from submissions that did not contain cost-utility analyses or cost minimisation analyses were excluded from multivariate regression analyses on the basis that the measures of effectiveness employed in their economic analyses precluded meaningful pooling with those from the remaining submissions (Figure 3.1, p66).

3.3.1 Data extraction and variable construction

Table 3.1 (p67) includes all the variables and rationale for their consideration in models of AWMSG decision-making. Data for medicines with recommendations as options for routine or restricted use were pooled together as 'positive' recommendations (41/60), and were compared against those with negative recommendations (19/60).

An incremental net monetary benefit (INMB) statistic at an assumed threshold for cost-effectiveness of £20,000 (or £30,000) per QALY gained was constructed in preference to the incremental cost effectiveness ratio (ICER), to preserve the use of data from submissions that included only cost minimisation analyses (CMAs, assumed QALY gain of zero). The basis of this cost effectiveness threshold was AWMSG guidance to submitting companies, which indicates that there is no fixed ICER threshold but below a most plausible ICER of £20,000 per QALY gained, judgements about the acceptability of a medicine as an effective use of NHS resources are to be based primarily on the cost-effectiveness estimate (AWMSG 2007).

The "health burden" of the underlying disease state upon patients was considered by mapping the appraised licensed indication of the medicines to those conditions included in the World Health Organisation Global Burden of Disease (WHO GBD) project (WHO 2008). This quantifies, on a population-standardised regional basis, the health effects of a wide range of diseases and conditions in terms of Disability-Adjusted Life Years (DALYs), which are a construct of years of life lost (YLL) from premature death (and years of life lost due to disability (YLD). From this we defined a categorical health burden variable (BurdenYLL), which takes a value of 1 if YLL is greater than YLD, based on statistics for 2004 for Europe region A (includes the UK) (WHO 2008).

Figure 3.1. Submissions providing data for univariate and logistic regression model analyses

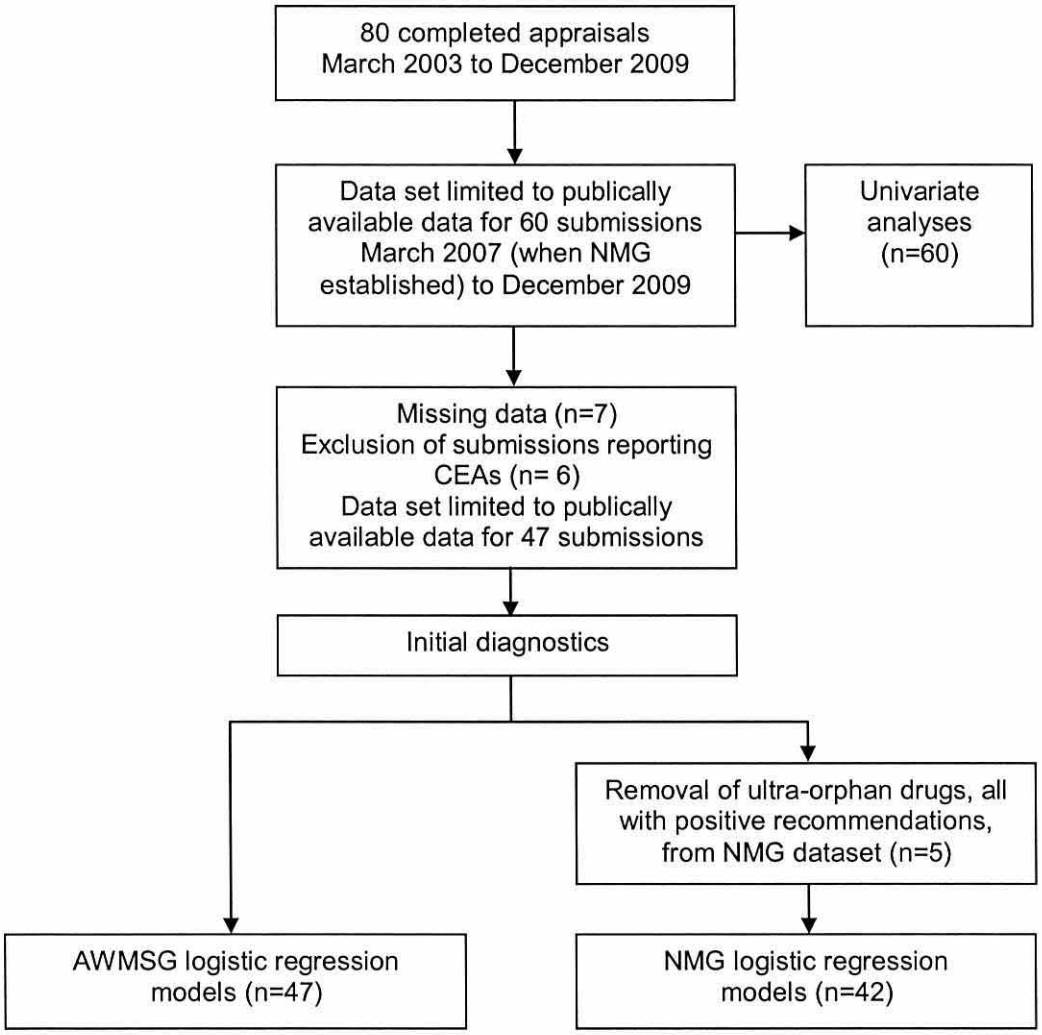


Table 3.1. Summary of variables used in regression analyses

Variable	Description	Rationale
Recommendation (AWMSG or NMG)	Takes the value 1 if positive recommendation (recommendation for routine or restricted use), otherwise 0.	Separate models developed for final AMWSG and preliminary NMG recommendations. Recommendations for routine or restricted use combined due to small number of restricted recommendations and the fact that some pharmaceutical companies submitted for a restricted recommendation from the outset.
INMB20 (INMB30)	Incremental net monetary benefit at a cost-effectiveness threshold of £20k (£30k) per QALY. Cost minimisation analyses assume an incremental QALY of zero.	INMB20 (INMB30) chosen as AWMSG guidance makes specific reference to ICERs below or above £20k-£30k per QALY in making judgements about the acceptability of a medicine as an effective use of NHS resources (AWMSG 2007). INMB used in preference to the ICER to preserve data from CMAs. ICERs and INMBs were as reported in the base case analysis in the company submission, or in the case of multiple comparators, the mean of reported base case ICERs.
No.Studies	Number of studies (of any kind) discussed as providing clinical efficacy data.	Indicator of quantity of available clinical efficacy data.
UltraOrphan	Takes the value of 1 if the new drug meets the criteria for ultra-orphan status, otherwise 0.	AWMSG policy permits consideration of additional factors when judging the cost effectiveness of ultra-orphan drugs (AWMSG 2011b).
Innovative	Takes the value of 1 if the new drug is first in class and/or there are no relevant comparators, otherwise 0.	AWMSG guidance encourages innovation in medicines that will benefit patients (AWMSG 2007). As AWMSG does not have a formal definition of innovativeness, this is a pragmatic definition that eliminates the need for subjective assessment of need or therapeutic advantage (Kennedy 2009; Puntmann et al., 2010).
PtGrpSub	Takes the value of 1 if a patient group submission was received, otherwise 0.	AWMSG and NMG may consider information from patient group organisations when judging clinical effectiveness and patient needs (AWMSG 2007).
BurdenYLL	Takes the value of 1 if the impact of the disease is greater on survival than on disability and 0 if not, as indicated by mapping the licensed indication to WHO GBD project DALY construct (WHO 2008).	AWMSG guidance states that consideration may be given, selectively, to the degree of clinical need of patients, and the particular features of the population and the condition being treated (AWMSG 2007). AWMSG ultra-orphan drugs policy permits consideration of the degree of severity of the condition in terms of quality of life and survival [AWMSG 2011b]. A minority of medicines licensed for conditions not directly considered in the WHO GBD project were mapped to those judged by the authors to be closely representative of patient experience.

Variable	Description	Rationale
RCTqual	Takes the value of 1 if supporting clinical efficacy data included double-blind RCTs that assessed patient-orientated outcomes against an appropriate comparator (which may include placebo/no treatment if relevant to the decision context).	Proxy indicator of the quality of available clinical efficacy data, chosen to reflect studies with high internal and external validity. The requirement for an appropriate comparator adds to the external validity of the trial data required for HTA.
PSA	Takes the value of 1 if probabilistic sensitivity analysis was undertaken as part of the economic evidence, otherwise 0.	AWMSG guidance states that consideration may be given to uncertainty generated in clinical and economic evidence. AWMSG has a preference for PSA for exploring the impact of uncertainty [AWMSG 2008].The presence or absence of PSA is an indicator of the company's consideration of combined uncertainty in clinical and economic data, and may serve as a proxy for the thoroughness of the modelling approach.
BIK	Estimated mean of the company-assessed annual budget impact over 5 year period (in £000s).	AWMSG may consider budget impact in making judgements about the acceptability of a medicine as an effective use of NHS resources (AWMSG 2007); however, NMG is advised not to consider budget impact.

AWMSG = All Wales Medicines Strategy Group; CMA = Cost minimisation analyses; DALY = Disability-adjusted life year; HTA = Health Technology Assessment; ICER = Incremental cost effectiveness ratio; NMG = New Medicines Group; PSA = Probabilistic sensitivity analysis; QALY = Quality-adjusted life years; RCT = Randomised controlled trial; WHO GBD = World Health Organisation Global Burden of Disease

The RCTqual variable was included as a proxy indicator of the quality of supporting evidence, and was constructed to reflect the evidence requirements of Health Technology Assessment (HTA) bodies. Components of the variable were informed by the Strength of Recommendation Taxonomy (SORT) evidence grading system (Ebell et al., 2004), which considers patient-orientated evidence (i.e. evidence of an impact upon morbidity, mortality or quality of life) derived from randomised controlled trials (RCTs) with high internal validity to be among the highest quality for making treatment recommendations. As the licensed medicines included in our dataset had been appraised by AWMSG close to market launch and are supported by mainly regulatory trial data, we pragmatically assumed that any significant internal validity issues would have been adequately addressed during the licensing process. We added the requirement for an appropriate comparator to the definition of the variable to capture external validity issues relevant to the HTA process. We therefore defined a categorical RCTqual variable, which takes a value of 1 if supporting evidence is available from double-blind RCTs that assessed patient-orientated outcomes against an appropriate comparator, or else 0. Each of the three components of this variable (double-blind RCTs; patient-orientated outcomes; appropriate comparator) was tested in the models individually in sensitivity analyses.

3.3.2 Univariate statistical analyses

All numeric descriptor data were non-normally distributed (Shapiro-Wilk W tests for non-normality $p<0.0001$) and so were analysed using the non-parametric, 2-sided Mann Whitney U test. All categorical variables were assessed using 2-sided Fisher's exact test, as expected values for several analyses were small and precluded the use of chi-squared tests.

3.3.3 Logistic regression modelling

AWMSG model development was conducted in three stages. In the first model, AWMSG decisions were regressed against the variables considered to be most analogous to those used in the preferred models of NICE decisions (Devlin and Parkin, 2004; Dakin et al., 2006): INMB20, PSA, Innovative, BIK, No.Studies, PtGrpSub, and RCTqual. In the second (full) model, two further variables, BurdenYLL and UltraOrphan, which are specific to AWMSG policy (AWMSG 2007; AWMSG 2011b); were added. For the third model, stepwise elimination of variables (when $p>0.2$, as

suggested by Menard (1995) was undertaken to identify the most influential variables and determine a most parsimonious model that fitted the data.

Evaluation of the models included measures of goodness of fit and pseudo-R^2, along with predictive accuracy, sensitivity and specificity, and area under the Receiver Operating Characteristic (ROC) curve. We tested for collinearity and correlation among the dependent and independent variables, and for overly influential cases among the data sample.

The variables in the full AWMSG models were used to specify a model of the preliminary recommendations of NMG. The NMG model was not estimable with inclusion of ultra-orphan drugs in the dataset, as all NMG recommendations for ultra-orphan drugs were positive; therefore, this variable was excluded from the NMG dataset. A parsimonious NMG model was constructed, as above.

Extensive sensitivity and scenario analyses were planned for the full AWMSG model, including deconstruction of composite variables and the use of alternative thresholds for cost effectiveness, categorical INMB variables, and incremental cost effectiveness (utility) ratios (ICERs). In addition, we explored the removal of the BIK variable from the NMG model to reflect the fact that, in contrast to AWMSG, NMG is advised not take into consideration the budget impact of the use of new medicines when making its recommendations (AWMSG 2007). Exploratory analyses included restriction of the dataset to non-ultra orphan drugs and to non-HIV drugs. All statistical analyses and modelling were conducted using StatsDirect statistical software version 2.7.7, 2009 (StatsDirect Ltd, England).

3.4 Results

Of the 60 submissions analysed, 41 (68%) received a positive recommendation for use by AWMSG (28 for routine use and 13 for restricted use), and 19 (32%) received a negative recommendation. There was agreement between the preliminary recommendations of NMG and the final recommendations of AWMSG for 50 (83%) submissions; four preliminary negative recommendations were changed to routine use, two preliminary negative recommendations were changed to restricted use, one preliminary recommendation for restricted use was changed to routine use, two preliminary recommendations for restricted use were changed to negative recommendations, and one preliminary recommendation for routine use was changed

to restricted use (see Appendix to Chapter 3). All final AWMSG recommendations were endorsed by the Minister for Health and Social Services.

3.4.1 Univariate analyses

The median of ICERs was more than two-fold higher for drugs that received a negative recommendation compared with those that received a positive recommendation (£28,563 versus £12,390; p=0.0099; Table 3.2, p72). The incremental costs for drugs that received a negative recommendation were significantly greater than for those that received a positive recommendation (£4,971 versus £684; p=0.0076); however, so were the reported QALY gains (0.33 versus 0.10; p=0.0465), which indicates that the incremental costs are driving the difference in ICER estimates. There were no significant differences in the proportions of submissions categorised as having a positive incremental net monetary benefit at cost-effectiveness thresholds of £20,000 or £30,000 per QALY gained, although it should be noted that the number of submissions providing data differed across these analyses.

Annualised budget impact estimates were three-fold greater for medicines that received a negative recommendation compared with those that received a positive recommendation. This finding was despite a lower annual eligible patient population, although neither difference achieved statistical significance.

Ten (24%) medicines that received a positive recommendation had orphan drug status (i.e. were indicated for serious conditions that affect not more than 5 in 10,000 persons, as defined in the European Union [European Parliament 1999]), compared with 10 (53%) of those with a negative recommendation (odds ratio for a positive recommendation [OR] 0.29; 95% CI 0.08 to 1.07; p=0.0417). However, there was no significant difference in rates for drugs with ultra-orphan status. Medicines considered to be innovative were also less prevalent among positive recommendations than among negative recommendations (24% versus 53%; p=0.0417).

In terms of the quality of supporting evidence, only the use of an appropriate comparator among the supporting clinical studies appeared to positively influence recommendations.

Table 3.2. Results of univariate analyses

	Positive recommendation (n=41)	Not recommended (n=19)	Median difference (CI) / Odds Ratio (95%CI)	p-value[b]
Numeric descriptors[a]; Median (Interquartile range) [Mean]				
ICER[c]	£12,390 (£0, £26,292) [£24,549][j]	£28,563 (£12,684, £48,528) [£63,836][k]	-£11,786 (95.1% CI -£28,184, -£3,657)	**0.0099**
Incremental costs	£684 (-£612, £9,000) [£9,711][l]	£4971 (£2,454, £42,273) [£195,115][m]	-£4,404 (95% CI -£18,602, -£1,668)	**0.0076**
QALY gain[d]	0.100 (0.00185, 0.360) [0.387][n]	0.330 (0.110, 0.877) [0.921][o]	-0.207 (95% CI -0.502, 0)	**0.0465**
Annual budget impact[e]	£48,495 (£9,810, £171,762) [£158,563]	£154,226 (£43,942, £491,168) [£403,653]	-£83,669 (95.2% CI -£294,826, £4,552)	0.0772
Eligible patient population size per year[f]	41 (14, 190) [198]	29 (22, 45) [130]	6 (95% CI -14, 67)	0.4840
Number of clinical studies in submission	3 (2, 4) [3.049]	3 (2, 4) [3.158]	0 (95.2% CI -1,1)	0.6677
Categorical descriptors				
CUA	31 (75.6%)	15 (78.9%)	0.8267 (0.1626, 3.5124)	>0.9999
CEA	3 (7.3%)	3 (15.8%)	0.4211 (0.0515, 3.5406)	0.3697
CMA	7 (17.1%)	1 (5.3%)	3.706 (0.4134, 176.1442)	0.4157
INMB @£20k >0	21 (55.3%)[p]	6 (42.9%)[q]	1.6471 (0.4044, 6.9369)	0.5364
INMB @£30k >0	29 (76.3%)[p]	10 (71.4%)[q]	1.2889 (0.2357, 6.0078)	0.7290
Orphan drug status:	10 (24.4%)	10 (52.6%)	0.2903 (0.0791, 1.0659)	**0.0417**
Ultra-orphan drug status[g]	5 (12.2%)	3 (15.8%)	0.4839 (0.0767, 3.7799)	0.3939
Innovative drugs[h]:	10 (24.4%)	10 (52.6%)	0.2903 (0.0791, 1.0659)	**0.0417**
First in class	8 (19.5%)	10 (52.6%)	0.2182 (0.0566, 0.8365)	**0.0150**
No comparator	5 (12.2%)	2 (10.5%)	1.1806 (0.1705, 13.5553)	>0.9999
Patient group submission	34 (82.9%)	14 (73.7%)	1.7347 (0.3647, 7.5996)	0.4927
Health burden mainly survival	29 (70.7%)	12 (63.2%)	1.4097 (0.3718, 5.0839)	0.5658
High quality RCT data available[i]:	6 (14.6%)	7 (36.8%)	0.2939 (0.0678, 1.2744)	0.0893
Efficacy data from DBRCTs	23 (56.1%)	13 (68.4%)	0.5897 (0.1534, 2.0983)	0.4095
Main studies assessed POOs	17 (41.5%)	13 (68.4%)	0.3269 (0.0854, 1.1703)	0.0946
Appropriate comparator used	34 (82.9%)	10 (52.6%)	4.3714 (1.0991, 17.4896)	**0.0258**
PSA conducted	26 (63.4%)	7(36.8%)	2.9714 (0.8426, 10.8593)	0.0930

[a] Shapiro-Wilk W test for non-normality p<0.0001; [b]2-sided p-value assessed by Mann-Whitney U test (numeric variables) or Fisher's exact test (categorical variables); [c] 7 dominant products assumed to have ICER zero;[d] Includes CMA where QALY gain is zero; [e] Based on mean of the company-assessed annual budget impact over 5 year period; [f] Based on company-estimated mean annual net number of patients eligible for treatment and predicted uptake over 5 years (from 55 submissions with accessible data); [g] versus non-orphan drugs; [h] Defined as first in class and/or without relevant comparators; [i] Supporting trials included DBRCTs to assess POOs against appropriate comparator; [j] 30 submissions providing 44 ICER estimates; [k] 14 submissions, providing 16 ICER estimates; [l] 36 submissions, providing 59 incremental cost estimates; [m] 16 submissions, providing 19 incremental cost estimates; [n] 35 submissions, providing 58 QALY gain estimates; [o] 12 submissions, providing 13 QALY gain estimates; [p] 38 submissions; [q] 14 submissions; CEA = cost effectiveness analysis; CMA = cost minimisation analysis; CUA = cost utility analysis; DBRCT = Double-blind randomised controlled trial; ICER=Incremental cost effectiveness ratio (incremental cost per QALY gained); INMB = Incremental net monetary benefit; QALY = Quality-adjusted life years; POOs= Patient-orientated outcomes (i.e. not surrogate outcome); RCT=Randomised controlled trial.

3.4.2 Logistic regression analyses

The first AWMSG model, based on the variables identified from the main models of Devlin and Parkin (2004) and Dakin et al. (2006), had a McFadden pseudo-R^2 of 0.22 and correctly classified 85% of recommendations (data not shown). The addition of AWMSG-specific variables (BurdenYLL and UltraOrphan) to produce our full AWMSG model resulted in a better fit of the data, and correctly classified 87% of recommendations (Table 3.3, page 75). Problematic levels of collinearity among the independent variables were not apparent (mean tolerance statistic across all nine independent variables 0.67, range 0.54 to 0.86) (Menard 1995). The model X^2 statistic failed to reach statistical significance for the both the initial and the full AWMSG models, probably due to the high number of variables in the models relative to the size of the data set (Menard 1995; Peng et al., 2002). However, we can reject the null hypothesis that AWMSG recommendations are unrelated to the independent variables in the parsimonious model, the findings of which are consistent with the main findings of the full AWMSG model.

When considered multivariately, only the presence of high quality supportive RCTs exerted a statistically significant (but counter-intuitive) effect at the level of $p<0.05$ in the full AWMSG model (Table 3.3). The odds of a medicine being recommended by AWMSG decreased significantly by 94% (95% CI, 30% to 100%) if supporting clinical evidence for the drug included double-blind RCTs that assessed patient-orientated outcomes against an appropriate comparator. It should be noted that only nine (19%) of the 47 submissions in this dataset met all three criteria required for "high quality" as presently defined. Along with cost-effectiveness, budget impact and innovativeness did not exert a significant effect upon AWMSG recommendations in multivariate analysis. The parsimonious AWMSG model confirms the significant, negative effect of the RCTqual variable upon AWMSG recommendations, and retains the variables PSA and No.Studies, which were not statistically significant within the full AWMSG model.

The full NMG model had a statistically significant X^2 statistic, indicating a good fit of the data, and there was no evidence of problematic collinearity among the variables (mean tolerance statistic 0.72, range 0.59 to 0.85) (Menard 1995). None of the independent variables were statistically significant at the level of $p<0.05$. The budget impact variable approached significance ($p=0.06$) and was retained in the parsimonious model, along with cost effectiveness represented as INMB20, although their influence appears to be small.

Table 3.3. Logistic regression analyses of AWMSG and NMG recommendations

	AWMSG Full Model			AWMSG Parsimonious Model			NMG Full Model			NMG Parsimonious Model		
Variables and coefficients												
	Coeff.	**Odds ratio (95% CI)**	***p***	**Coeff.**	**Odds ratio (95% CI)**	***p***	**Coeff.**	**Odds ratio (95% CI)**	***p***	**Coeff.**	**Odds ratio (95% CI)**	***p***
INMB20	4.11E-07	1 (1.0000, 1.0000)	0.7365				0.0001	1.0001 (1.0000, 1.0002)	0.1446	0.0001	1.0001 (1.0000, 1.0002)	0.1569
PSA	1.4364	4.2055 (0.6833, 25.8833)	0.1213	1.4147	4.115 (0.8477, 9.9800)	0.0793	1.4147	4.1153 (0.6830, 24.7976)	0.1226	1.1482	3.1524 (0.6581, 15.1013)	0.1509
BIK	-0.0009	0.9991 (0.9976, 1.0006)	0.2655				-0.0050	0.9950 (0.9898, 1.0002)	0.0593	-0.0040	0.9960 (0.9917, 1.0004)	0.0719
Innovative	-1.1827	0.3064 (0.0487, 1.9303)	0.2078				-1.1403	0.3197 (0.0429, 2.3806)	0.2656			
No.Studies	0.4187	1.5200 (0.7489, 3.0852)	0.2463	0.4115	1.5091 (0.8149, 2.7947)	0.1906	0.0864	1.0902 (0.6542, 1.8169)	0.7403			
RCTqual	-2.8259	0.0593 (0.0050, 0.6994)	**0.0248**	-2.6608	0.0699 (0.0088, 0.5546)	**0.0118**	-1.5377	0.2149 (0.0216, 2.1389)	0.1897			
PtGrpSub	-0.8934	0.4093 (0.0143, 11.7411)	0.6019				-0.3479	0.7062 (0.0502, 9.9266)	0.7964			
BurdenYLL	1.0839	2.9561 (0.3972, 22.0000)	0.2899				1.5041	4.5002 (0.5571, 36.3545)	0.1582			
Ultra-Orphan	1.2396	3.4542 (0.1637, 72.8820)	0.4256									
Intercept	0.3420			-0.2175			-0.1922			0.2689		

Model evaluation				
	AWMSG Full Model	**AWMSG Parsimonious Model**	**NMG Full Model**	**NMG Parsimonious Model**
N	47	47	42	42
Deviance goodness of fit X^2	39.4987; df=37; p=0.3589	13.2737; df=13; p=0.4269	38.6445; df=33; p=0.2297	42.9189; df=38; p=0.2685
Model X^2	13.9031; df=9; p=0.1258	9.4011; df=3; **p=0.0244**	17.175921; df=8; **p=0.0283**	12.9016; df=3; **p=0.0049**
McFadden Pseudo-R^2	0.2603	0.4146	0.3077	0.2311
Sensitivity	100%	97.14%	84.62%	84.62%
Specificity	50%	41.67%	62.50%	56.25%
Area under ROC curve	0.8131	0.7667	0.8209	0.8041
Correctly classified	87.23%	82.98%	76.19%	73.81%
Parsimonious Models = Stepwise-reduced parsimonious models, using cut-off p>0.2 (Menard 1995)				

3.4.3 Sensitivity and scenario analyses

Substitution of the INMB20 variable with INMBs at cost effectiveness thresholds of £30,000 to £50,000 per QALY gained, or with categorical INMBs greater than zero, had no material impact upon the full AWMSG model, and was not significant within the full NMG model. However, in the full NMG model, the budget impact variable reached significance when the threshold for cost effectiveness is increased to £30,000 per QALY gained (and above) in the continuous and categorical INMB variables.

Deconstructing the INMB into incremental costs and QALYs had no impact upon model outputs; neither variable was significant in the AWMSG or the NMG models. None of the individual components of the RCTqual variable were significant when tested in either the AWMSG or the NMG model. Removal of the budget impact variable from the NMG model resulted in a poor fit of the data (Model X^2 statistic p=0.0985) and the remaining variables explained less of the variability compared with the full NMG model (McFadden's pseudo-R^2 0.22).

Restricting the analyses to submissions including only CUAs, to remove the contribution of CMAs to the net monetary benefit variable, supports the findings of the full AWMSG (n=39) and NMG (n=34) models. When further restricted to submissions reporting a positive base-case ICER, the continuous ICER variable also did not exert a statistically significant effect on AWMSG (n=33; p=0.4504) or NMG (n=28; p=0.1130) recommendations. Exclusion of ultra-orphan drugs from the AWMSG dataset (n=42, to reflect the NMG dataset), improved the AWMSG model fit (McFadden pseudo-R^2 0.31) although the model X^2 statistic remained non-significant (p=0.0784). Variable significance remained as per the full AWMSG model, and an exploratory parsimonious AWMSG model (X^2 statistic p=0.0076) retained the RCTqual (p=0.0674), PSA (p=0.1487) and the budget impact (p=0.0482) variables. Finally, exploratory analyses with exclusion of HIV drugs from the dataset (to explore the finding that all 12 HIV drugs received positive recommendations, despite none being supported by patient-orientated evidence) produced poor fitting AWMSG (n=35) and NMG (n=30) models. Interpretation of analyses conducted on specific restricted datasets is limited by the small sample sizes.

AWMSG decisions were expected to be highly influenced by the preliminary recommendations of NMG, and inclusion of the NMG recommendation as an explanatory variable resulted in an inordinately large coefficient estimate.

3.5 Discussion

Our multivariate analysis of appraisals made by the All Wales Medicines Strategy Group revealed that only the RCTqual variable exerted a significant effect upon recommendations and, furthermore, that this was a negative effect. Other putative explanatory variables, including the cost-effectiveness of medicines, their budget impact, and underlying disease characteristics, were not significant predictors of decisions.

One plausible explanation for the unexpected negative influence of RCT quality upon AWMSG recommendations is that medicines supported by higher quality evidence have attracted higher overall costs (e.g. are priced at a premium), which has resulted in high ICERs. Indeed, unadjusted univariate analyses indicate that incremental costs are driving increased ICER estimates for medicines with negative recommendations. However, there was no evidence of significant collinearity or correlation to suggest an association between RCT quality, ICERs, or incremental cost variables.

A more plausible explanation may be found in our definition of RCT quality as applied to the range of medicines in our dataset. Based on unadjusted univariate analyses, when the RCT quality variable was disaggregated, the use of appropriate comparators in supporting RCTs was statistically significantly lower among medicines with negative recommendations. In addition, our dataset is dominated by high cost specialist medicines, which may present methodological difficulties to the generation of patient-orientated evidence. For example, twelve (20%) were for the treatment of HIV, all of which received positive AWMSG recommendations and none of which were supported by evidence from RCTs that would be defined as "high quality" using the SORT criteria, due to the fact they typically assessed surrogate (virological and/or immunological) outcomes. Multivariate sensitivity analyses indicate that none of the three components of the RCTqual variable exterted a statistically significant effect on AWMSG recommendations. Our composite definition of high quality RCT evidence may, therefore, have been too strict for the range of specialist medicines within our dataset, by excluding the use of well established or validated surrogate endpoints (Taylor and Elston, 2009). It is also possible that our analyses have failed to capture judgements on the plausibility of the evidence base for medicines with negative recommendations.

We were guided in our definition of the RCTqual variable by the SORT criteria for grading evidence quality (Ebell et al., 2004), as we felt these criteria reflected well the considerations of effectiveness and external validity involved in HTA and reimbursement-

decision processes. However, the need to make medicines reimbursement decisions ever closer to market authorisation is likely to be associated with greater reliance on evidence generated primarily to meet the needs of licensing authorities. Our findings serve to emphasise the differences in the evidence requirements of licensing authorities and HTA bodies. Whilst the former generally focus on measures of efficacy and safety in distinct patient populations in the context of RCTs, HTA bodies focus on effectiveness in the real world, which includes wider considerations and evidence from a wider range of sources (Breckenridge et al., 2010). There are clearly difficulties in applying the usual hierarchies of evidence to the HTA process (Rawlins 2008).

Dakin et al. (2006) employed the average Jadad score as a measure of RCT quality within their models. This focuses on a limited number of internal validity aspects of RCTs (Jadad et al., 1996) but of itself takes no account of the appropriateness of outcomes measures and comparators for informing reimbursement decisions. Interestingly, increasing RCT quality (i.e. increasing Jadad score) was associated with a numerically (but not statistically significant) increased risk of receiving a restricted or a negative recommendation rather than a recommendation for routine use (Dakin et al., 2006). Devlin and Parkin (2004) did not specifically consider the quality or validity of the available clinical evidence in their analyses of NICE recommendations.

The full NMG model provided a good fit of the data, but no variables achieved significance at the level of $p<0.05$. However, the point estimates of their coefficients and odds ratios indicated that the direction of influence of the variables was consistent between the AWMSG and the NMG models. The NMG dataset excluded ultra-orphan drugs, which may have reduced the power of the NMG analyses versus the AWMSG analyses, but use of the NMG dataset in the AWMSG model yielded similar results to the full AWMSG model.

Multivariate sensitivity and scenario analyses using the full AWMSG and NMG models indicate that neither the assumed threshold for cost effectiveness when constructing the INMB variables, nor ICERs, influenced recommendations. This conflicts with the unadjusted univariate analyses, which suggest the ICER is statistically significantly greater for medicines with negative recommendations compared with positive recommendations. However, the analyses need to be interpreted in the context of the limited sample size and our approach to preserve this. Inclusion of seven submissions supported by CMAs in the regression dataset, six of which received positive recommendations, would potentially impact upon incremental costs and QALY gains.

Scenario analyses to explore the impact of the exclusion of CMAs were consistent with the full regression dataset analyses, but inevitably with reduced power to detect a true difference.

Whilst cost effectiveness *per se* was not observed to influence recommendations, at thresholds for cost effectiveness of £30,000 per QALY gained and above, budget impact was observed to exert a statistically significant negative impact upon NMG recommendations. This is an interesting finding, as AWMSG guidance to submitting companies indicates that NMG is advised not to consider budget impact when making its preliminary recommendations (AWMSG 2007). Removal of the budget impact variable from the full NMG model, to reflect this AWMSG guidance, produced a poorly fitted model and the remaining variables explained less of the variability compared with the full NMG model, providing further support of an influence of budget impact upon recommendations. However, the odds ratio indicates that influence is very small.

Parsimonious models were developed in recognition of the limited data sample size and have greater power to identify influential factors. These support the general findings of the full models, although only RCTqual in the parsimonious AWMSG model achieved significance.

The odds of a positive recommendation were increased four-fold for drug submissions supported by economic modelling that considered the combined uncertainty in clinical and economic parameters via PSA. This finding lends further support to the possibility that there were issues with plausibility and uncertainty in the evidence available in support of medicines that received negative recommendations. Although the presence of PSAs to represent consideration of uncertainty is simplistic, AWMSG guidance to submitting companies indicates a preference for PSA (AWMSG 2008). We therefore felt our approach had advantages over that of Devlin & Parkin (2004), who expressed uncertainty as a function of the reported range of cost effectiveness ratios relative to the mean or base case cost-effectiveness ratio estimate. Dakin et al. (2006) did not specifically consider economic uncertainty in their models.

A positive influence of a greater body of clinical evidence has also been apparent in NICE recommendations (Dakin et al., 2006). In the present analysis, the increase in the odds of a positive recommendation with each additional supporting clinical study was not significant at the $p<0.05$ level. However, whereas NICE submissions in the analyses of Dakin et al. (2006) included, on average, 14 RCTs, AWMSG submissions included a

mean of three clinical studies. This is likely to be a consequence of AWMSG appraising medicines closer to the time of market launch when the body of supporting evidence is less mature.

The parsimonious NMG model also found PSA to be among the more influential variables. INMB20 and budget impact were also retained in the parsimonious NMG model, although both appear to have only a small influence based on their odds ratios, which is lost in the transition from the preliminary recommendations of NMG to final recommendations of AWMSG. Possible explanations include the wider remit of AWMSG to consider societal factors and budget impact, alongside evidence of clinical and cost effectiveness (AWMSG 2007).

Other observations of note within the full AMWSG model include the positive (but not significant) influence upon recommendations for medicines indicated for the treatment of diseases that impact to a greater extent upon survival than upon quality of life, and medicines with ultra-orphan status. Previous revealed preference analyses have not considered the influence of the severity of the underlying disease upon recommendations of NICE (Devlin and Parkin, 2004; Dakin et al., 2006), but an international comparison of drug appraisals, including those conducted by NICE, found no association between recommendations and whether or not the underlying condition was life-threatening (Clement et al., 2009). Our regression analyses provide tentative evidence of consideration by AWMSG of equity-related issues consistent with a special case for patients with rare disease; univariate analyses indicate that AWMSG policy has elevated the recommendations for ultra-orphan drugs to similar rates as non-orphan medicines (Linley and Hughes, 2010).

There are a number of caveats which may limit interpretation of our findings. First, our models are, by necessity, simple representations of complex decision-making processes. As our analyses were restricted to data available within the public domain, it is plausible that other influential factors have not been captured. Omitted variable bias in this type of study is difficult to explore and yet may account for some unexplained findings, such as the 100% positive recommendation rate observed for medicines used in the treatment of HIV.

Second, we developed a binary choice model of "positive" or "negative" recommendations, although Dakin et al. (2006), suggested that modelling routine and restricted NICE recommendations separately may provide a better representation of

decision-making. However, in contrast to the NICE appraisals considered in that study, several companies submitted for a restricted AWMSG recommendation for their medicine from the outset, probably due to difficulties in demonstrating cost effectiveness across the whole appraised licensed indication. Although we anticipate that submitting companies would prefer an unrestricted AWMSG recommendation for their medicines, in these circumstances it would be inappropriate to assume that a restricted recommendation represents a lesser recommendation as a result of the decision-making process *per se*.

Third, model development was systematic, being based around previous works (Devlin and Parkin, 2004; Dakin et al., 2006), with further refinement for AWMSG-specific factors, but this was at the expense of statistical power. The limited sample size may have contributed to the observed lack of influence of putative explanatory variables, including cost effectiveness and budget impact. Therefore, as was observed by Dakin et al. (2006), in their analyses of NICE decisions, it is not possible to conclude with confidence that a non-significant variable is not an influential factor upon AWMSG and NMG recommendations; however, we can be more confident that those variables found to have a significant effect are important determinants of recommendations. Our full AWMSG model had greater predictive power than did the Dakin et al. (2006) models of NICE recommendations. However, in recognition of the limited statistical power of our full model analyses, we also developed parsimonious models via appropriate stepwise elimination of variables. The model X^2 statistics of the resultant parsimonious models were significant, indicating a good fit of the data, and the ratio of observations to independent variables, being in the range 14 to 16, exceeded the general rule of at least 10 that is often considered appropriate (Peng et al., 2002). The findings of the parsimonious models were consistent with the general observations from the full models.

Finally, the construction and definition of several variables are potential limitations. Our use of the incremental net monetary benefit statistic as an explanatory variable, in an attempt to preserve sample size, implicitly assumes there are no qualitative differences between submissions that include CUAs and those that include CMAs. Given that submissions supported by CMAs tend to relate to formulation changes or "me-too" drugs, rather than new chemical entities, this may not be the case. However, removal of CMA-supported submissions from the dataset in exploratory analyses had little effect on the model outputs. Our definition of innovative drugs may be challenged as it does not consider the context of need within the target patient population or the extent to which the new medicine offers a therapeutic advantage (Kennedy 2009; Ferner et al., 2010). Although novel, our approach to include an indicator of the health burden on patients of

the underlying condition inevitably involved a degree of subjectivity for a minority of medicines with licensed indications that could not be mapped directly to the conditions included within the WHO GBD project (e.g. treatment of invasive candidiasis was mapped to meningitis on the basis that both are acute, life-threatening conditions requiring specialist treatment in an intensive care setting). For these latter two variables we adopted a pragmatic and relatively standardised approach aimed at minimising our subjective assessment of therapeutic advantage and need.

3.5.1 Conclusion

There are parallels in the results of the present study of AWMSG decision-making and previous explorations of NICE decision-making. Both our study and Dakin et al. (2006), reveal a preference for medicines supported by a greater body of clinical evidence, and whilst uncertainty in the ICER influenced NICE recommendations in the study of Devlin and Parkin (2004), consideration of combined uncertainty in economic model parameters had a positive influence on the recommendations of AWMSG (and preliminary recommendations of NMG). All three studies have noted that ICERs for drugs with negative recommendations are, on average, greater than for drugs with positive recommendations (Devlin and Parkin, 2004; Dakin et al., 2006), consistent with the pursuit of economic efficiency. However, only Devlin and Parkin (2004), found the cost effectiveness ratio to be a significant influencing factor in NICE recommendations when tested multivariately; neither our study nor that of Dakin et al. (2006), found cost effectiveness to be a consistent significant determinant of final recommendations. The negative influence of RCT quality in both our study and that of Dakin et al. (2006), is also of interest, and would seem to highlight difficulties in the application of evidence hierarchies and grading schemes to the complex decision-making processes involved in HTA. However, the results of our analyses and those of others (Devlin and Parkin, 2004; Dakin et al., 2006) need to be interpreted in the context of a number of limitations, including limited sample sizes.

Our dataset was confined to a specific range of medicines, defined as high cost or for the treatment of cardiovascular disease or cancer. Expansion of the AWMSG programme from October 2010, to cover all new medicines not subject to imminent review by NICE, will significantly increase the range and number of medicines covered. This will provide opportunities for further research into the factors influencing the decision-making process for new drugs in Wales and may facilitate further comparisons with the decisions of other HTA organisations.

3.6 Authors' contributions

WGL (the candidate) and DAH (supervisor) conceived the study. WGL obtained the final published AWMSG recommendation documentation, extracted data and compiled the database, and mapped the appraised licensed indication of the medicines to those conditions included in the World Health Organisation Global Burden of Disease (WHO GBD) project [WHO 2008]. DAH audited the data. WGL analysed the data, interpreted the results and drafted the manuscript. WGL and DAH revised the manuscript for intellectual content. WGL finalised the manuscript.

Chapter 4

Stated preference, discrete choice experiment among All Wales Medicines Strategy Group appraisal committee members

Published as:

Linley WG, Hughes DA. Decision-makers' preferences for approving new medicines in Wales: a discrete choice experiment and exploration of external validity. *Pharmacoeconomics* 2013. doi: 10.1007/s40273-013-0030-0. [Epub ahead of print]..

4.1 Preface

Chapter 3 presented a revealed preference study of AWMSG decision-making, using logistic regression methods to explore the impact of putative variables on the odds of recommendation, based on a sample of 47 submissions made to AWMSG between 2007 and 2009. That study was the first exploration of the preferences of a UK HTA body other than NICE.

This Chapter presents a stated preference study, which has been designed with reference to the revealed preference study of AWMSG decision-making discussed in Chapter 3. This ensures face validity, and enables comparison of stated preferences against actual historical AWMSG decisions as an assessment of external validity. This is one of only a few studies to explore the external validity of stated preferences in the health economics arena, and is the first to explore the external validity of the stated preferences of a HTA body.

4.2 Abstract

Background: Few studies to date have explored the stated preferences of national decision-makers for health technology adoption criteria, and none of these have compared stated decision-making behaviours against actual behaviours. Assessment of the external validity of stated preference studies, such as discrete choice experiments (DCEs), remains an under-researched area.

Objectives: The primary aim was to explore the preferences of All Wales Medicines Strategy Group (AWMSG) appraisal committee and appraisal sub-committee (the New Medicines Group) members ("appraisal committees") for specific new medicines adoption criteria. Secondary aims were to explore the external validity of respondents' stated preferences and the impact of question choice options upon preference structures in DCEs.

Methods: A DCE was conducted to estimate appraisal committee members' preferences for incremental cost effectiveness, quality-adjusted life years (QALY) gained, annual number of patients expected to be treated, the impact of the disease on patients before treatment, and the assessment of uncertainty in the economic evidence submitted for new medicines compared with current NHS treatment. Respondents evaluated 28 pairs of hypothetical new medicines, making a primary forced choice between each pair and a more flexible secondary choice, which permitted either, neither or both new medicines to be chosen. The performance of the resultant models was compared against previous AWMSG decisions.

Results: Forty-one out of a total of 80 past and present members of AWMSG appraisal committees completed the DCE. The incremental cost effectiveness of new medicines, and the QALY gains they provide, significantly ($p<0.0001$) influence recommendations. Committee members were willing to accept higher incremental cost effectiveness ratios and lower QALY gains for medicines that treat disease impacting primarily upon survival rather than quality of life, and where uncertainty in the cost effectiveness estimates has been thoroughly explored. The number of patients to be treated by the new medicine did not exert a significant influence upon recommendations. The use of a flexible choice question format revealed a different preference structure to the forced choice format, but the performance of the two models was similar. Aggregate decisions of AWMSG were well predicted by both models, but their sensitivity (64%, 68%) and specificity (55%, 64%) were limited.

Conclusions: A willingness to trade cost effectiveness and QALY gains against other factors indicates that economic efficiency and QALY maximisation are not the only considerations of committee members when making recommendations on the use of medicines in Wales. On average, appraisal committee members' stated preferences appear consistent with their actual decision-making behaviours, providing support for the external validity of our DCEs. However, as HTA involves complex decision-making processes, and each individual recommendation may be influenced to varying degrees by a multitude of different considerations, the ability of our models to predict individual medicine recommendations is more limited.

4.3 Introduction

Within the constraints of a fixed health care budget, such as exists for the UK National Health Service (NHS), the funding of new medicines requires that other existing medicines and services are displaced. Policy makers therefore have difficult decisions to make that balance the most effective, sustainable use of finite resources, and the legal and moral obligations to provide fair, comprehensive care for the populations they serve. Consequently, many countries have established centralised appraisal systems and Health Technology Assessment (HTA) bodies to make national recommendations on the use of new medicines (Stafinski et al., 2011a).

Several studies have analysed previous decisions of various HTA bodies to determine factors that influence recommendations (George et al., 2001; Grégoire et al., 2001; PausJenssen et al., 2003; Devlin and Parkin, 2004; Dakin et al., 2006; Harris et al., 2008; Tilson et al., 2010; Mason et al., 2010; Chim et al., 2010). These revealed preference (RP) studies generally make use of documented evidence within the public domain and, whilst such sources can be informative, they may not fully capture all influences on the decision-making process. Indeed, a review of appraisal systems world-wide has concluded that, whilst policies and decision-making criteria to guide appraisal committees generally appear to be transparent, it remains unclear how such policies and criteria are operationalised during committee deliberations (Stafinski et al., 2011b). An alternative approach, then, is to use stated preference (SP) techniques, such as discrete choice experiments (DCEs), to explore decision-making.

Compared to RP studies, DCEs allow greater control over experimental design to permit estimation of the impact, and willingness to trade-off, various characteristics (attributes) making up hypothetical choice alternatives. However, if DCE-based stated preferences are to reflect real-life decision-making, it is imperative that attributes and levels employed to describe choice alternatives have face validity, and that choice tasks represent as far as possible the true nature of decision-problems (Lancsar and Louviere, 2008).

There has been growing interest in the application of DCEs in a range of health care contexts (Ryan and Gerard, 2003; De Bekker-Grob et al., 2010); however, few to date have explored the impact of cost effectiveness and other health technology adoption criteria upon the recommendations of national decision-makers. Tappenden et al. (2007), conducted a binary choice experiment among 37 appraisal committee members of the National Institute for Health and Clinical Excellence (NICE) in the UK and, more recently,

Koopmanschap et al. (2010), conducted a forced-choice DCE among 66 Dutch health professionals, 40% of whom were policy makers. Both found respondents were willing to make trade-offs between the incremental cost effectiveness ratio (ICER) of health technologies and other attributes. Whitty et al. (2011), did not explore the impact of cost effectiveness *per se*, but observed increased survival and quality of life, and reduced costs and uncertainty, to increase the likelihood of drug reimbursement in a pilot study among 11 members of the Australian Pharmaceutical Benefits Advisory Committee (PBAC). Unfortunately, none of these, nor many other health-related DCEs to date, have compared stated decision-making behaviours against actual behaviours, and assessment of the external validity of DCEs remains an under-researched area (Ryan and Gerard, 2003; De Bekker-Grob et al., 2010).

The All Wales Medicines Strategy Group (AWMSG) is a Welsh Government-funded body with a remit to appraise new medicines for use in NHS Wales when NICE guidance is not imminent (AWMSG 2011a). The primary aim of the current DCE was to explore the preferences of AWMSG appraisal committee members for specific new medicines adoption criteria and their efficiency trade-offs. As HTA involves complex decision-making processes, the attitudes of appraisal committee members towards a range of other potential influencing factors, including UK medicines resource allocation policies, were also explored. Secondary aims included exploration of the external validity of the DCE by comparing appraisal committee members' hypothetical recommendations against actual AWMSG recommendations for the use of new medicines in Wales. As forced choice pairwise comparisons may not provide the most accurate preference estimates (Johnson and Backhouse, 2006), we further explored the impact upon preference structures of permitting more flexible choices among the alternatives in our DCE.

4.4. Methods

4.4.1 Participants and administration

The AWMSG appraisal process involves a preliminary recommendation made by the New Medicines Group (NMG) sub-committee, based on evidence of clinical and cost effectiveness, and a final recommendation made by AWMSG, based on consideration of the preliminary NMG recommendation, and budgetary and broader societal impacts. With the assistance of the All Wales Therapeutics and Toxicology Centre (AWTTC), which provides professional and administrative support for the appraisal process, contact details

were obtained for 72 out of a total of 80 current and past voting members and deputies of AWMSG and NMG up to May 2011. All 72 were invited to participate in the study.

Questionnaires were completed anonymously online. Invitees received three reminders, and a small incentive, of entry to a prize draw for a £50 high street gift voucher, was offered for those who confirmed via email they had completed the questionnaires within four weeks of invitation. Participants were provided with an explanation of the study, a detailed description of the attributes and levels and an example choice task to complete before the DCE proper (see Appendix to Chapter 4). In an effort to encourage completion of each choice task independently of all other choice tasks, respondents were not permitted to view or amend their previous choices.

4.4.2 DCE — Attributes and levels

The profiles of hypothetical new medicines consisted of five attributes, which were selected based on consideration of AWMSG process documentation (AWMSG 2011a), results of previous RP studies of NICE (Devlin and Parkin, 2004; Dakin et al., 2006) and AWMSG decisions (Chapter 3), and the attributes included in previous relevant SP studies (Tappenden et al., 2007; Koopmanschap et al., 2010; Whitty et al., 2011) (Table 4.1, p92).

To provide a standardised representation of the main impact of the disease before treatment (IMPACT) for our RP study of AWMSG decision-making (Chapter 3), the licensed indications of the medicines appraised by AWMSG between 2007–9 were mapped to those conditions included in the World Health Organisation Global Burden of Disease project, which includes estimates of their associated impact on years of life lost (Survival) and years of life lost due to disability (used as a proxy for Quality of Life) (WHO 2008). For consistency, the IMPACT attribute in the current DCE was similarly defined in terms of Survival and Quality of Life.

Three levels for each continuous attribute were adopted. To ensure face validity, these were chosen specifically to reflect the distribution of values observed in previous AWMSG submissions (Chapter 3).

Table 4.1. Attributes and levels of the DCE choice tasks

Attributes	Description	Levels and coding	Rationale
Main impact of disease *before* treatment (IMPACT)	The main health burden due to the underlying disease or condition (before treatment with either new medicine) may be to reduce survival (Survival) or may be to reduce health-related quality of life (Quality of Life) of patients, compared with age-matched people without the condition.	Effects coded: Survival 1 Quality of life -1	*Attribute:* Previous SP studies found baseline disease severity (Whitty et al., 2011) and HRQoL (Tappendent et al., 2007; Koopmanschap, et al., 2010) to significantly influence recommendations. AWMSG process documentation makes specific reference to the underlying severity of disease in terms of baseline HRQoL and prognosis (AWMSG 2011). *Levels:* For standardisation and consistency with previous RP study of AWMSG decision-making (Chapter 3), IMPACT was defined in terms of whether disease mainly impacted upon Survival or Quality of Life.
Annual number of patients to be treated (No_PTS)	The number of patients anticipated to be treated with the new medicine each year, if recommended. Note, this is not necessarily the number of patients afflicted by the disease or condition; it is the number of patients anticipated to receive treatment with the new medicine in that indication each year.	Continuous (unit 100 patients): 40 500 1000	*Attribute:* Number of patients affected by the condition was found to influence past NICE recommendations [Devlin and Parkin, 2004] and budget impact was found to influence Dutch policy-makers' stated preferences [Koopmanschapp et al., 2010]. AWMSG may consider budget impact, but NMG excludes budget impact from its considerations [AWMSG 2011], although both appraisal groups may consider patient numbers under the AWMSG ultra-orphan drugs policy [AWMSG 2011]. Patient numbers to be treated represents a compromise between the use of budget impact and number of patients affected by the condition. *Levels:* Estimates of annual number of patients to be treated for 39 AWMSG drug submissions ranged from 2 to 1,313 (Chapter 3). The 20th, 50th and 80th percentiles produced insufficient spacing of levels, therefore the chosen upper level reflects 80% of the maximum estimate, and the lower level was informed by AWMSG policy for appraising medicines for very rare diseases (ultra-orphan drugs), defined implicitly as those licensed for severe diseases affecting fewer than 60 patients in Wales [AWMSG 2011].

Attributes	Description	Levels and coding	Rationale
QALYs gained per treated patient (QALYg)	The average number of quality-adjusted life years gained per patient treated with the new medicine versus the current standard of care for that disease or condition.	Continuous (unit 1 QALY): 0.1 0.8 1.6	*Attribute:* Putative variable included as a measure of treatment benefit that is familiar to appraisal committee members and to understand ICER influence. Study objectives dictate inclusion. *Levels:* Lower and upper levels reflect the 20th and 80th percentiles of QALY gains reported in AWMSG submissions for medicines supported by CUAs (Chapter 3). The middle level is a simple mid-point of these, as the median (0.4) was viewed to be too close to the lower level by pilot respondents.
Incremental cost per QALY gained (ICERk)	The cost effectiveness of the new medicine presented as the incremental cost per quality-adjusted life year (QALY) gained for the new medicine versus the current standard of care for that disease or condition.	Continuous (unit £1,000): £4,000 £18,000 £40,000	*Attribute:* Previous SP studies found ICER to have a significant influence on recommendations (Tappenden et al., 2007; Koopmanschapp et al., 2010). AWMSG appraisal guideline makes specific reference to the role of cost effectiveness in decision-making (AWMSG 2011). Study objectives dictate inclusion. *Levels:* Reflect the 20th, 50th and 80th percentiles of ICERs reported for 39 AWMSG submissions supported by CUAs (Chapter 3). Middle and upper levels aim to capture effects of the £20–£30k threshold range quoted in AWMSG process documentation (AWMSG 2011).
Uncertainty in cost effectiveness is thoroughly explored (UNCERTAINTY)	This indicates whether or not the degree of uncertainty in cost effectiveness estimates has been explored by assessing the combined uncertainty arising from several data sources (known as probabilistic sensitivity analysis, PSA).	Effects coded: Yes 1 No -1	*Attribute:* AWMSG appraisal guideline makes specific reference to uncertainty in cost effectiveness estimates (AWMSG 2011). *Levels:* Reporting of PSA was found to influence AWMSG/NMG recommendations in a RP study (Chapter 3).

AWMSG=All Wales Medicines Strategy Group; CUA=Cost utility analysis; HRQoL=Health-related quality of life; ICER=Incremental cost effectiveness ratio; NMG=New Medicines Group; PSA=Probabilistic Sensitivity Analysis; QALY=Quality-adjusted life-year; RP=Revealed preference; SP=Stated preference

4.4.3 DCE — Experimental design

Applying the 'rule of thumb' that each main effect level of interest should be represented across the design at least 500 times (Orme 2010), we estimated that a 50% response rate (36 participants) to a two-alternative, forced-choice format with a maximum of three levels per attribute would require a design with at least 21 choice tasks. To generate an efficient fractional factorial design we therefore used an orthogonal main effects plan consisting of 27 pairwise choices (Hann and Shapira, 1996), created by collapsing columns to accommodate the two-level attributes and using a fold over design to avoid overlap of attribute levels. One of the medicines profiles among the 27 included a clearly dominated profile, and we introduced a further dominated choice task as an internal validity test for rational trading behaviours. Respondents were randomised to one of four versions of the questionnaire, created by a simple shift in the order of the choice tasks blocked into seven questions to reduce potential issues of ordering and learning effects (Bateman et al., 2008).

Each pair-wise choice task required respondents to make a primary forced choice to recommend for use in NHS Wales one of the two hypothetical new medicines, each of which were compared incrementally against usual care. In a secondary question, respondents were presented with a more flexible choice between recommending either, neither or both medicines. An example choice task is displayed in Figure 4.1.

Figure 4.1. Example choice task

		New Medicine A		New Medicine B	
Main impact of disease *before* treatment		Survival		Quality of life	
Annual number of patients to be treated		40		500	
QALYs gained per treated patient		1.6		0.1	
Incremental cost per QALY gained		£40,000		£4,000	
Uncertainty in cost effectiveness is thoroughly explored		No		Yes	
Primary question	Which medicine would you prefer to recommend for approval (please tick **one** box, ✓)				
Secondary question	Given the choice, would you recommend approval of (please tick **one** box, ✓):	A ONLY	B ONLY	Both A & B	Neither

4.4.4 Piloting

We piloted the DCE among five members of research staff at the Centre for Health Economics and Medicines Evaluation at Bangor University, and six members of AWTTC who are directly involved in the assessment of pharmaceutical industry submissions for consideration by the AWMSG and regularly attend NMG and AWMSG committee meetings. This led to minor clarifications of attribute wording, rearrangement of the order of attribute presentation within each hypothetical new medicine profile and adjustment of the spacing of levels for the attributes relating to QALYs gained and number of patients to be treated.

4.4.5 DCE — Statistical analysis

Given the panel nature of the data, a random effects logit model was used to analyse the binary responses to the primary forced-choice questions. Following Johnson and Backhouse (2006), a conditional logit model, which accommodates multiple selections among the alternatives (e.g. both new medicine A and new medicine B), was employed to analyse responses to the secondary flexible-choice questions. As the option of recommending neither medicine was present within each flexible choice task, an alternative specific constant for recommending a medicine rather than rejecting a medicine (neither) was specified in that model. Main effects models were estimated as none of the interaction terms, incorporated to explore the possibility of interaction between the continuous attributes, were statistically significant. A finite population correction was applied to calculate confidence intervals and p-values. All statistical analyses were conducted using STATA® SE v10.1, 2009.

4.4.6 Supplementary questionnaire

Likert rating scales were used to explore attitudes towards a range of non-standard appraisal criteria, including current policies of AWMSG relating to treatments for patients at the end of their lives (AWMSG 2011c) and those with very rare diseases (AWMSG 2011b). Views on other policies and principles relevant to medicines reimbursement in the UK (e.g. the Department of Health's Cancer Drugs Fund in England (Dept Health 2010c), reasons identified by NICE for departing from conventional thresholds for cost effectiveness (Rawlins et a., 2010), and the proposed criteria for rewarding new medicines with higher prices under the future value-based pricing framework (Dept Health 2010d), were also solicited.

4.5. Results

Forty-one past and current voting members of AWMSG (n=20; 16 current) and NMG (n=21; 16 current) completed the DCE (providing 2296 observations in the primary forced-choice model). All respondents chose to recommend the two dominant medicine profiles included to test for rational trading behaviours. Of 1,148 completed choice tasks, 1,122 (97.7%) responses were consistent between the primary forced choice and the secondary flexible choice questions, indicating rational choices were being made for the secondary flexible choice question given primary forced choice selections. Removal of the two respondents responsible for the majority of the 26 inconsistent responses made no qualitative difference to the coefficients or their significance, and so results are presented based on the full dataset. Overall goodness of fit of both the forced choice and the flexible choice models was good (Pseudo-R^2 values 0.33–0.35, Model X^2 p-values <0.0001).

4.5.1 Primary forced-choice DCE model

New medicine A was preferred in 55.2% of the forced choice tasks (see Appendix to Chapter 4, Table 4.1A). The observed sign for each attribute coefficient was as expected and all, except the annual number of patients to be treated, exerted a highly significant influence ($p<0.0001$) on recommendations for the use of new medicines (Table 4.2, p97). All else being equal, the odds of a positive recommendation decreased by 8% for every £1,000 increase in the ICER. An increase of one QALY increased the odds of recommendation three-fold; thorough consideration of uncertainty using PSA increased the odds of recommendation more than two-fold; and medicines intended for the treatment of diseases that impact mainly upon survival rather than quality of life increased the odds of recommendation by 73%.

All else being equal, and assuming a 50% probability threshold for recommendation of a new medicine over the existing standard of care, the threshold for cost effectiveness (i.e. maximum willingness to pay) for medicines used in the treatment of diseases that impact mainly on patient survival was £27,000 per QALY gained, and a minimum QALY gain of 0.38 was required. In comparison, for medicines used in the treatment of diseases that impact mainly on health-related quality of life, the maximum willingness to pay per QALY gained was reduced by half and the minimum QALY gain required for recommendation was increased more than three-fold (Table 4.3, p98).

Table 4.2. Econometric models

	Forced choice model (A vs.B)		**Flexible-choice model (A vs. B vs. Both vs. Neither)**	
Model type	Random effects logit model		Conditional (fixed effects) model	
Attribute	Coefficient (95% CI)	OR (95% CI)	Coefficient (95% CI)	OR (95% CI)
ICER (/£1000)	-0.0792 (-0.0848, -0.0736)*	0.9239 (0.9187, 0.9290)*	-0.0657 (-0.0709, -0.0605)*	0.9364 (0.9314, 0.9414)*
QALYg	1.1069 (0.9790, 1.2348)*	3.0250 (2.6382, 3.4118)*	0.7257 (0.6075, 0.8439)*	2.0662 (1.8222, 2.3102)*
UNCERTAINTY	0.9178 (0.8381, 0.9975)*	2.5038 (2.3042, 2.7034)*	0.5717 (0.5031, 0.6403)*	1.7712 (1.6496, 1.8928)*
IMPACT (Survival)	0.5494(0.4743, 0.6245)*	1.7322 (1.6021, 1.8424)*	0.0587 (-0.0638, 0.1812) (NS)	1.0604 (0.9958, 1.1250) (NS)
No_PTS (/100)	-0.0047 (-0.0240, 0.0146) (NS)	0.9953 (0.9761, 1.0146) (NS)	-0.0083 (-0.0263, 0.0097) (NS)	0.9917 (0.9739, 1.0095) (NS)
_constant	0.6937 (0.5145, 0.8728)*		1.0378 (08361, 1.2395)*	
Number of observations	2296		3444	
Log likelihood	-1059.74		-823.76	
Model X^2	572.88 (5 d.f); p < 0.0001		874.89 (6 d.f); p < 0.0001	
Pseudo R^2	0.3341†		0.3468	

Model X^2 p-value calculated on LL ratio test of full model vs. constant-only model
†Pseudo-R^2 calculated as 1-(LL full model/LL constant-only model) = 1-(-1059.74/-1591.47)
*p<0.0001; NS=not statistically significant at level p≤0.05;
CI=Confidence interval; LL=log likelihood; OR=Odds ratio;
Finite Population Correction Factor=0.7026, based on N=80, n=41

Table 4.3. Thresholds for cost effectiveness and QALY gains under scenarios

Scenario (all else being equal)	Thresholds - Forced choice model		Thresholds - Flexible choice model†	
	ICER (x£/QALY)	QALYg	ICER (x£/QALY)	QALYg
Main disease impact before treatment: Survival	<27,000	>0.88	<14,200	>0.52
Main disease impact before treatment: Quality of Life	<13,100	>1.38	<12,400	>0.68
Uncertainty thoroughly explored using PSA	<31,700	>0.05	<24,600	*
Uncertainty not thoroughly explored with PSA	<8,500	>1.71	<7,200	>1.15
Main disease impact before treatment: Survival AND Uncertainty thoroughly explored using PSA	<38,600	*	<25,800	*
Main disease impact before treatment: Survival AND Uncertainty NOT thoroughly explored using PSA	<15,400	>1.21	<8,400	>1.05
Main disease impact before treatment: Quality of Life AND Uncertainty thoroughly explored using PSA	<24,700	>0.55	<24,000	*
Main disease impact before treatment: Quality of Life AND Uncertainty NOT thoroughly explored using PSA	<1,500	>2.20	<6,600	>1.21

ICER=Incremental cost effectiveness ratio; PSA=Probabilistic sensitivity analysis; QALYg=Quality-adjusted life-year gained
†Main disease impact before treatment was not statistically significant in the flexible choice model
*=No minimum QALY gain required in this scenario

4.5.2 Secondary flexible-choice DCE model

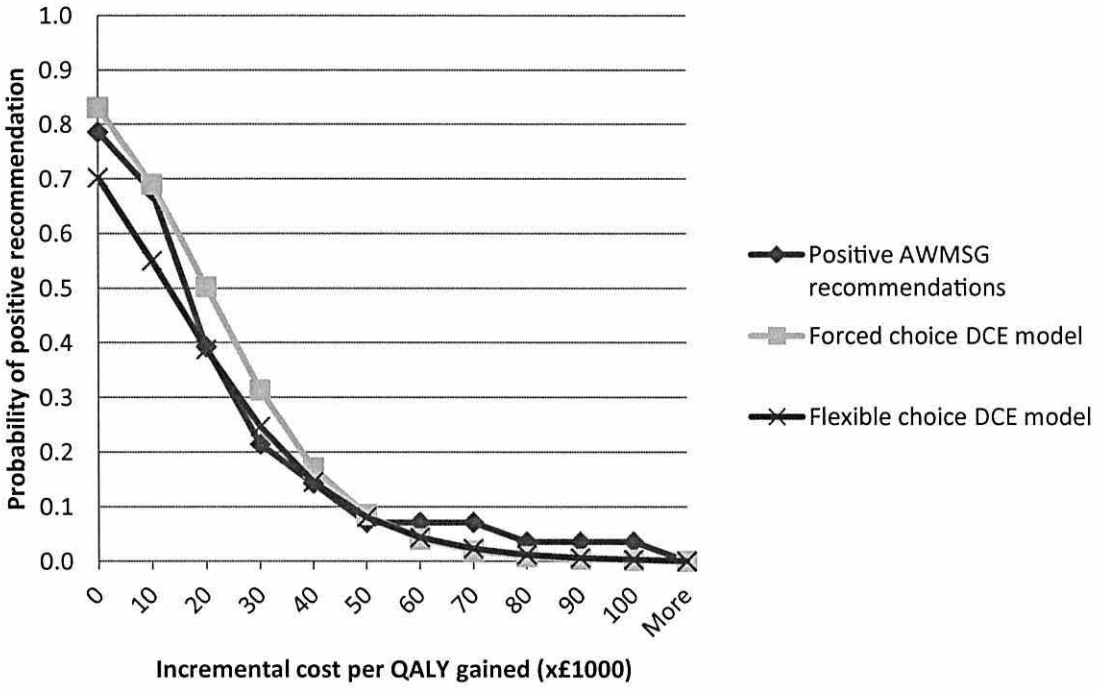
New medicine A was preferred in 32.4% of the flexible choice tasks, compared with 26.6% for new medicine B, 21.1% for both medicines and 19.9% for neither. The ratio of respondents favouring new medicine A to new medicine B was maintained (1.2:1) but the introduction of more flexible alternatives permitted a mean of 43.3% of respondents to vary their primary fixed choice responses, and there was no choice task for which all 41 respondents retained their original response to the forced choice question. There appeared to be less homogeneity in responses compared with the forced choice tasks (see Appendix to Chapter 4, Table 4.1A).

As in the forced choice model, the observed sign for each attribute coefficient was as anticipated. The positive constant ($p<0.0001$) indicates that respondents were more likely to recommend approval of either or both new medicines than to select neither as their preferred choice. However, in addition to the annual number of patients to be treated, the impact of the disease on patients before treatment was also no longer statistically significant. The effect of a unit change in the QALYg and UNCERTAINTY attributes upon the odds of a positive recommendation was reduced compared with the forced-choice model (Table 4.2), as was the maximum willingness to pay per QALY gained and the minimum QALY gain required for a positive recommendation (using the 50% probability threshold) (Table 4.3).

4.5.3 External validity of the DCE models

Of the 39 AWMSG appraisals of medicines supported by cost-utility analysis (2007-9), 28 (72%) received a positive recommendation. The mean probabilities of recommendation of all 28 medicines with positive AWMSG recommendations when their attribute values were tested in the forced-choice and flexible-choice models were above the assumed 50% probability threshold for recommendation (forced-choice model, 63%; flexible-choice model, 56%). For the 11 medicines with negative AWMSG recommendations the respective probabilities were below 50% (42% and 39%) (see Appendix to Chapter 4, Table 4.2A). The influence of the ICER on predicted probabilities of positive recommendations derived from both DCE models are presented in Figure 4.2 (p100), with the cumulative probability of actual positive AWMSG recommendations (i.e. any recommendation for use in NHS Wales) superimposed for reference.

Figure 4.2. Cumulative probability of positive recommendation based on incremental cost effectiveness for the stated preference models and actual positive AWMSG recommendations



On an individual basis, the forced choice model correctly predicted the AWMSG recommendation for 25 (64%) submissions, with a sensitivity of 68% and specificity of 55%. The flexible choice model also correctly predicted 64% of AWMSG recommendation, but with sensitivity and specificity of 64%. Among 18 medicines with a positive AWMSG recommendation and for which parameter uncertainty in cost effectiveness had been assessed by PSA, the forced and flexible choice models correctly predicted 90% and 85% of recommendations, respectively. Among 22 medicines with positive recommendations and which are indicated for the treatment of disease that mainly impact upon patient survival, the forced and flexible choice models correctly predicted 77% and 68% of recommendations, respectively.

4.5.4 Supplementary questionnaire

Based on a Likert scale of 1=no importance, to 5=utmost importance, each of the factors captured within the attributes included in the DCE choice tasks received a median score of 4. However, other factors (e.g. availability of alternative treatment options, budget impact) also had a median score of 4 and the quality of clinical evidence scored 5 (Figure 4.3, p101).

Figure 4.3. Committee members' ratings of importance of various factors

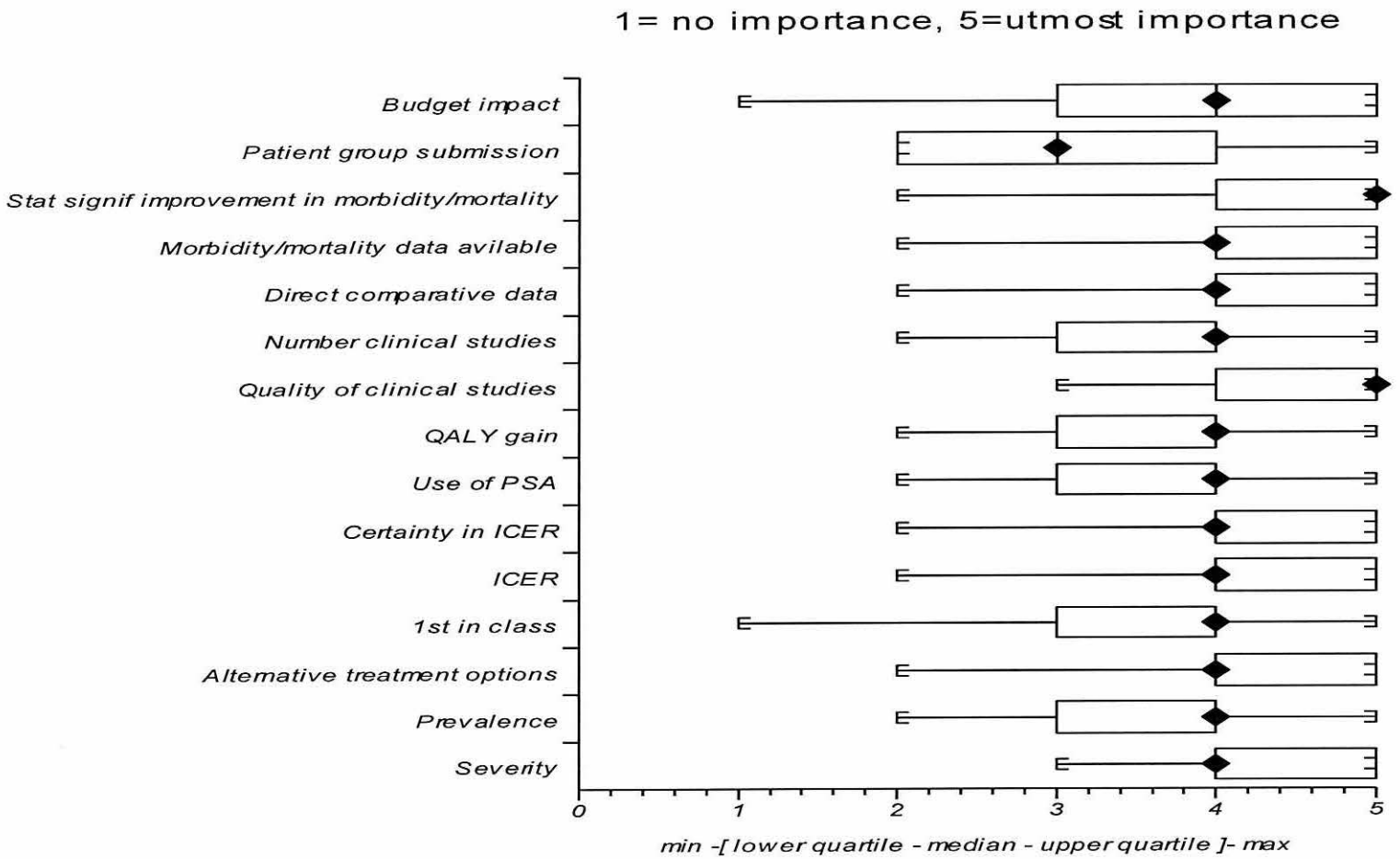
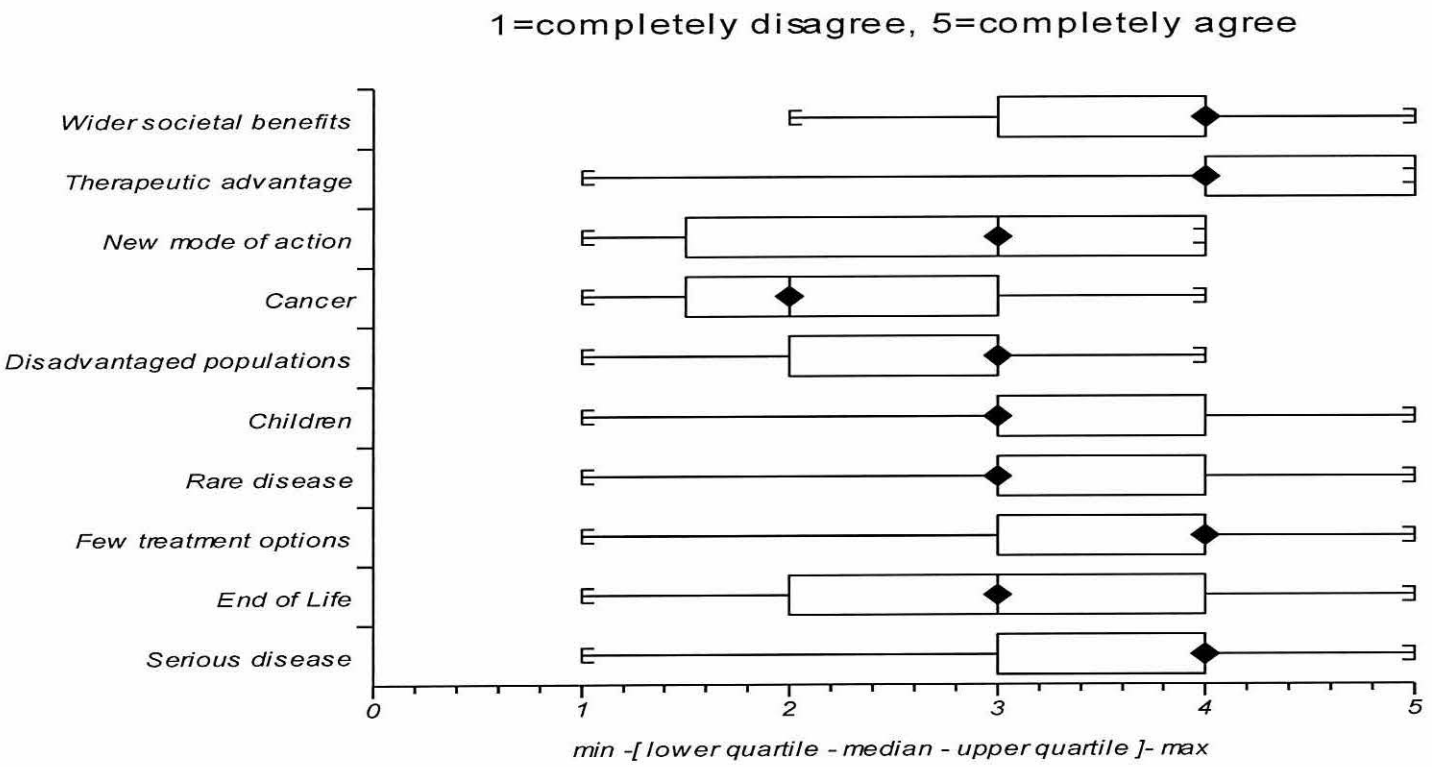


Figure 4.4. Committee members' levels of agreement with policies for paying premium medicine prices



Appraisal committee members agreed, with a median score of 4 on a Likert scale of 1=strongly disagree to 5=strongly agree, that all else being equal, the NHS should be prepared to pay a premium for medicines that: are intended for treatment of very severe diseases; are intended for treating diseases with few existing treatment options; offer a significant therapeutic advantage; and offer wider societal benefits (e.g. reduced reliance upon carers). There was indifference (median score 3) towards policies that favour premium prices for medicines aimed at patients at the end of life, those with very rare diseases, children, and disadvantaged populations, and there was a level of disagreement (median score 2) with a policy that would favour the NHS paying premium prices for cancer medicines compared with non-cancer medicines, all else being equal (Figure 4.4, p102).

4.6 Discussion

Our DCE models indicate the incremental cost effectiveness, QALY gains, and whether or not uncertainty surrounding cost effectiveness estimates has been thoroughly explored, all significantly influence AWMSG appraisal committee members' recommendations for new medicines reimbursement by NHS Wales. Whether or not the impact of the disease on patients before treatment influences decisions is less clear, being dependent on the choice options available to respondents. The annual number of patients to be treated by the new medicine did not exert a significant influence on recommendations in either model.

Appraisal committee members are willing to make trade-offs against the ICER and QALY gains. The implied cost effectiveness threshold was greater, and the minimum QALY gain required for recommendation was lower, where uncertainty in cost effectiveness estimates has been considered in PSA and when considering treatments for diseases that mainly impact upon survival compared with those that mainly impact upon health-related quality of life. Assuming good experimental design and high face validity, our findings are broadly consistent with AWMSG's appraisal guidelines, which state that a fixed cost effectiveness threshold is not employed and the case for supporting medicines with ICERs above the range £20-£30,000 per QALY gained needs to be increasingly strong (AWMSG 2011). This is encouraging from a process perspective and goes some way to addressing recent concerns about how policies and decision-making criteria are operationalised by HTA committee members (Stafinski et al., 2011b).

Our tests of the external validity of the stated preference models have produced mixed findings. The predicted probabilities of recommendation of new medicines are good when aggregated across all AWMSG recommendations, and the concordance between the model-derived predicted probabilities of positive recommendations and the cumulative probability of actual positive AWMSG recommendations (Figure 4.2, p100) is remarkable. However, based on their sensitivities and specificities, the ability of the models to discriminate between individual new medicines recommendations appears limited.

Although the performance of the forced, and flexible choice models was similar in terms of their ability to predict AWMSG recommendations, actual preference structures within each model differed. Imposing a forced choice on appraisal committee members resulted in acceptance of higher ICERs and a requirement for greater QALY gains than was the case when given the option to accept or reject both new medicines. Our study therefore also provides further empirical evidence in support of Johnson and Backhouse (2006), who, in their DCE of health technology adoption criteria conducted primarily among pharmaceutical industry personnel, concluded that the option to reject both alternatives in a paired choice comparison may be important for obtaining accurate preference estimates.

Our supplementary questions revealed indifference towards AWMSG-specific policies that permit more lenient considerations of the cost effectiveness of ultra-orphan medicines intended for treatment of very rare diseases (AWMSG 2011b), or those that extend life of patients who are at the end of life (AWMSG 2011c). However, it should be noted that AWMSG had not had cause to implement its end of life policy prior to our conducting the DCE. There was agreement with the proposed criteria for rewarding new medicines with higher prices under the imminent UK value-based pricing framework (Dept Health 2010a; Hughes 2011), but disagreement with a suggestion that, all else being equal, the NHS in Wales should prioritise cancer treatments over others, as implied by the Cancer Drugs Fund in England (Dept Health 2010c). Importantly, these policy views of appraisal committee members were broadly consistent with those of the UK general public (including 213 in Wales) who responded to a choice-based questionnaire that explored societal preferences for NHS resource allocation (see Chapter 5).

4.6.1 Strengths and limitations

We aimed to achieve high face validity for a parsimonious DCE. Attributes were selected with reference to those shown previously to be relevant to policy-makers based on RP

and SP studies, and tempered with our previous analyses of factors influencing AWMSG decision-making (Chapter 3). The levels for each attribute were carefully selected to reflect those usually seen by AWMSG appraisal committee members. The experimental design was generated with particular reference to the limited potential number of respondents, and although most DCEs conducted to date have typically used fewer choice tasks (De Bekker-Grob et al., 2010), our number was comparable to that used by Koopmanschap et al. (2010), and benefited from a lower cognitive burden due to fewer attributes.

Our use of a secondary, more flexible choice question ensured preference estimates were not unnecessarily restricted and permitted further examination of respondents' preference structures. We achieved a respectable response rate of 57%, representing 51% of all past and present appraisal committee members. We encouraged independent completion of each choice task, and made efforts to reduce the potential impact of learning effects by randomising participants to one of four versions of the questionnaire. Our results demonstrate internal validity and generally rational trading behaviours, with a low rate of obvious inconsistent choices. We directly explored the external validity of our DCE models, which has been done rarely in the health care context (Ryan and Gerard, 2003; De Bekker-Grob et al., 2010). We have determined that the influence of putative attributes is, on average, consistent with actual AWMSG decision-making behaviours (criterion validity). The influence of these attributes is also consistent with that of similar attributes observed to influence national decision-makers in other published DCEs (Tappenden et al., 2007; Koopmanschap et al., 2010; Whitty et al., 2011), and lends further credibility (convergent validity) to our findings. However, the lack of exploration of external (criterion) validity in other published DCEs precludes a robust assessment of their choice models.

There are some caveats, however. Our choice tasks include simple, generic attributes that we felt are applicable to most reimbursement decisions; however, they relate only to reimbursement decisions based on cost utility analyses, and around 25% of submissions to AWMSG in the period 2007-9 relied on alternative types of economic analyses (Chapter 3). The 39 AWMSG-appraised medicines included within our external validity tests represent a heterogeneous mix, including high cost, highly specialist medicines and those intended for very rare conditions with high unmet needs. Such medicines often exceed the usual thresholds of cost-effectiveness, and it is plausible that other influencing factors, applied on an individual basis to each medicine, may override the generic attributes we considered in the DCE. As no other comparable DCEs have considered the

external validity of their findings to date, we are unable to determine the relative performance of ours. However, we would anticipate similar challenges in demonstrating their external validity as they too use simplified attributes to model complex decision-making processes.

In contrast to our pilot study, 41% of respondents felt the profiles of hypothetical medicines may not be sufficiently descriptive. The results of our supplementary questionnaire suggest the quality of clinical evidence to be more important than the attributes contained within our DCE, and the degree of uncertainty around cost effectiveness estimates to be of the same importance as whether uncertainty had been thoroughly explored using PSA. We excluded quality of clinical evidence from our DCE attributes on the basis of the results of our RP study (Chapter 3), in which we concluded that the usual hierarchies of evidence quality are difficult to apply to the high cost, highly specialist medicines appraised by AWMSG close to market launch. Previous, related DCEs (Tappenden et al., 2007; Koopmanschap et al., 2010; Whitty et al., 2011) did not include clinical evidence quality as a separate attribute, and an earlier RP study of NICE decision-making failed to find a positive influence of randomised controlled trial quality (Dakin et al., 2006). Nonetheless, it remains possible that an important influencing factor has been omitted.

Regarding representation of uncertainty in cost effectiveness, previous DCEs categorised uncertainty as 'high' or 'low' (Tappenden et al., 2007; Whitty et al., 2011), or incorporated uncertainty as the probability that the ICER in each hypothetical treatment profile was at least double (Koopmanschap et al., 2010). We considered neither of these approaches to be satisfactory, as we have no standard means with which to judge what constitutes a high or low level of uncertainty, and we felt the latter approach would be cognitively challenging and of uncertain face validity. We were also concerned that the inclusion of a 'high' level of uncertainty could potentially dominate choice tasks and therefore elected to include whether or not uncertainty in cost effectiveness had been thoroughly explored, using PSA as a proxy. This was observed to be of importance in our RP study of AWMSG decision-making (Chapter 3) and would allow the external validity of our models to be tested.

We included both NMG and AWMSG appraisal committee members in our sample, 78% of whom were currently serving committee members. Whilst we do not anticipate preferences for what are putative, generic attributes to have changed greatly over the five years since the institution of NMG, we acknowledge that NMG makes only preliminary

recommendations and it is AWMSG that makes the final decision to the Welsh Government on the use of new medicines in the NHS in Wales, taking account of wider factors. It is plausible that NMG and AWMSG appraisal committee members have different preference structures, resulting from their different remits, but our sample was too small to assess sub-groups. We are, however, reassured by the 83% level of agreement observed between 60 preliminary NMG and final AWMSG recommendations made in the period 2007-9 (Chapter 3), and the 82% level of agreement among the 39 medicines used to test the external validity of our models.

Finally, as in other DCEs, we implicitly assume that the aggregate stated preferences of individual respondents represent the preferences of the decision-making body as a single entity. Within the forced choice DCE, responses appear to be reasonably homogeneous with obvious majority votes; however, the inclusion of additional alternatives within the choice sets provides a greater opportunity for individuals to express different preferences. The conditional logit model we used to analyse responses to the flexible choice model relies on the assumption of independence of irrelevant alternatives (IIA), which implies that the alternatives within the choice tasks compete with each other equally. This may not hold upon the introduction of an opt-out (Neither medicine) alternative to our choice tasks (Ryan et al., 2008), as medicines A and B are likely to compete with each other more closely than the alternative of Neither medicine. It is therefore possible that coefficient estimates for the flexible choice model are biased. However, the external validity of the flexible choice model is similar to that of our fixed choice model, and the usual alternative approaches that relax the assumptions of IIA (e.g. multinomial probit, nested logit and mixed logit models) cannot accommodate respondent selections of multiple alternatives, as required for our flexible choice model. Our adoption of the conditional logit model was therefore a pragmatic decision but is a limitation of the flexible choice model.

4.6.2 Conclusion

We performed a DCE that, despite limitations, permitted AWMSG voting members to provide rational stated preferences for putative new medicines' reimbursement criteria. The incremental cost effectiveness of new medicines, and the QALY gains they provide, significantly influence decisions to varying degrees depending on whether or not the uncertainty in cost effectiveness has been thoroughly explored, and whether or not the primary impact of the disease is on survival or quality of life. A willingness to trade the cost effectiveness and QALY gains against these other factors indicates that economic

efficiency and QALY maximisation are not the only considerations of AWMSG when making recommendations on the use of medicines in Wales. On average, appraisal committee members' stated preferences appear consistent with their actual decision-making behaviours, providing support for the external validity of our DCEs. Committee members' stated and revealed preferences are also broadly in line with appraisal process guidance. However, as HTA involves complex decision-making processes, and each individual recommendation may be influenced to varying degrees by a multitude of different considerations, the ability of our models to predict individual medicine recommendations is more limited.

4.7. Authors' contributions

WGL (the candidate) and DAH (supervisor) conceived the study. WGL designed the survey, managed questionnaire administration and data collection, analysed responses, interpreted the results and drafted the manuscript. WGL and DAH revised the manuscript for intellectual content. WGL finalised the manuscript.

Chapter 5

Empirical analysis of public preferences for medicines prioritisation criteria

Published as:

Linley WG, Hughes DA. Societal views on NICE, Cancer Drugs Fund and Value-based Pricing criteria for prioritising medicines: a cross sectional survey of 4118 adults in Great Britain.

Health Economics 2012. doi: 10.1002/hec.2872.

[Epub ahead of print]

5.1 Abstract

The criteria used by the National Institute for Health and Clinical Excellence (NICE) for accepting higher incremental cost effectiveness ratios for some medicines over others, and the recent introduction of the Cancer Drugs Fund (CDF) in England, are assumed to reflect societal preferences for NHS resource allocation. Robust empirical evidence to this effect is lacking. To explore societal preferences for these and other criteria, including those proposed for rewarding new medicines under the future value-based pricing (VBP) system, we conducted a choice-based experiment in 4,118 UK adults via web-based surveys. Preferences were determined by asking respondents to allocate fixed funds between different patient and disease types reflecting nine specific prioritisation criteria. Respondents supported all four criteria proposed for rewarding new medicines under the VBP system (tackle severe diseases, address unmet needs, are innovative, and have wider societal benefits), but did not support the end-of-life premium or the prioritisation of children or disadvantaged populations as specified by NICE, nor the special funding status for treatments of rare diseases, nor the CDF. Policies introduced on the basis of perceived –and not actual– societal values may lead to inappropriate resource allocation decisions with the potential for significant population health and economic consequences.

5.2 Introduction

The UK National Health Service (NHS) has legal and moral obligations to provide fair, comprehensive, needs-based care for all (Dept Health 2012a). Given the unprecedented efficiency savings demanded across the NHS in recent and coming years (Dept Health, 2009; Institute for Fiscal Studies/Nuffield Trust, 2012), it is imperative that resource allocation decisions provide the most effective, sustainable use of finite resources. The National Institute for Health and Clinical Excellence (NICE) makes compulsory recommendations on the use of medicines and other health technologies in the National Health Service (NHS) in England and Wales. The funding of new medicines requires that other existing medicines or services are displaced, the opportunity cost of which is reflected in NICE's cost-effectiveness threshold, set at £20,000 to £30,000 per quality-adjusted life-year (QALY) gained (NICE 2008a). However, several medicines with incremental cost-effectiveness estimates in excess of this threshold range have been approved by NICE for use via the NHS (e.g., sunitinib for advanced renal cancer, riluzole for motor neurone disease) (Rawlins et al., 2010).

Justification for this departure from the usual cost effectiveness threshold range includes the social value judgements of NICE's Citizen Council. Based on its views, six specific criteria besides clinical and cost effectiveness have recently been put forward as reflecting societal preferences in the allocation of health resources (Rawlins et al., 2010). Despite these laudable efforts to incorporate societal views into the NICE work programme, the extent to which this group of 30 lay persons can reflect the views and preferences of the public as a whole regarding the allocation of scarce health resources has been questioned (Buxton and Chambers, 2011), and it is suggested that access to some new medicines, such as those to treat cancer, may still be inappropriately restricted (Dept Health 2010c).

The Cancer Drugs Fund (CDF) was introduced in England in 2011 to facilitate access specifically to cancer medicines that have received a negative opinion from NICE on the grounds they do not represent a good use of NHS resources, or which have not yet been appraised. The government justified the CDF – set at £200m per annum – on the basis that: "*...it is possible that society values health benefits to patients with cancer more highly, all else being equal, than benefits to patients suffering other conditions*" (Dept Health 2010c). Although disease severity is consistently viewed as a valid criterion for prioritising health resources (Dolan et al., 2005; Shah 2009), we are unaware of any empirical evidence for the preferential funding of cancer treatments.

From 2014, all new branded medicines in the UK will be priced according to their therapeutic value and wider benefits they may deliver (Dept Health 2010d; Dept Health 2011). Under this value-based pricing (VBP) system, it is proposed that explicit weightings be attached to the health benefits (QALY gains) provided by medicines to reflect a broader range of relevant criteria (Dept Health 2010a) (see Table 5.1, p113). Again, with the exception of severity of disease, empirical evidence of the desirability of these criteria for rewarding new medicines with premium prices seems lacking (Dept Health 2010d).

It is apparent, therefore, that current prioritisation criteria used in pricing and reimbursement systems in the UK, and recent initiatives to address their perceived short-comings, are without robust supporting empirical evidence that they reflect societal preferences. This may lead to inappropriate resource allocation decisions, which take on a greater importance in the context of the increasing financial pressures under which the NHS is operating.

Our study explored societal preferences for the prioritisation criteria used by NICE, those proposed under the VBP system, and the UK government's assumptions used to justify the introduction of the CDF. In addition, we explored whether a societal preference exists for treating rare diseases over more common diseases, given that funds are top-sliced for certain treatments of very rare diseases in England (NCGHSS 2011), and both the All Wales Medicines Strategy Group (AWMSG) and the Scottish Medicine Consortium (SMC) permit additional considerations in the appraisal of medicines for the treatment of such diseases (AWMSG 2011b; SMC 2010).

5.3 Methods

5.3.1 Questionnaire design

We reviewed relevant documents and policies [AWMSG 2011b; Dept Health 2010a; Dept Health 2010c; Dept Health 2010d; Dept Health 2010e; Dept Health 2011; NCGHSS 2011; NICE 2008a; NICE 2008b; NICE 2008c; NICE 2009b; Rawlins et al., 2010; SMC 2010] to identify nine specific prioritisation criteria (besides clinical- and cost-effectiveness) for exploration within our study (Table 5.1, p113).

Table 5.1. Current and proposed medicines prioritisation criteria explored

Prioritisation criteria explored	Reported rationale for use of criteria within UK NHS	Scenario construct
Severity of underlying disease	***NICE***: Society would generally give priority to the expensive relief of a very serious condition than to the inexpensive relief of a mild condition [Rawlins et al., 2010]. ***VBP***: Society may place a greater weight on treating particularly severe or life threatening conditions [Dept Health 2010a].	***All else being equal***: Medicine for severe disease compared against medicine for moderate disease, as mild disease may be viewed as not necessitating treatment (in cohorts 1 & 2). ***Trade-off scenarios***: As severe disease is the criterion in question, trade-off scenarios explored smaller health gains (in cohort 1) and higher costs for treatment (in cohort 2) of severe disease compared with moderate disease.
Unmet need	***VBP***: The current system [of appraisal] may not fully reflect society's preferences if there are no existing alternative treatments, and so a significant unmet need (Dept Health 2010a).	***All else being equal***: Medicine for disease with several other treatments available via the NHS compared against medicine for a disease with only one medicine available (in cohorts 1 & 2). ***Trade-off scenarios***: As unmet need is the criterion in question, trade-off scenarios explored smaller health gains (in cohort 1) and higher costs for treatment (in cohort 2) for the disease with only one medicine available compared with the disease with several treatments available.
Significant innovation	***NICE***: Product produces a demonstrable and distinct benefit of a substantial nature that may not be adequately captured in the quality of life measure used (Rawlins et al., 2010). ***VBP***: A treatment representing a significant breakthrough and an important advance over existing therapies would provide a large QALY benefit. It could also be represented by a qualitative assessment of the innovation reported by a new medicine reflecting e.g. new modes of action (Dept Health 2010a).	***All else being equal***: Medicine that works in a new way compared against medicine that works in similar way to existing alternatives for treatment of the same disease (in cohorts 1 & 2). ***Trade-off scenarios***: As innovation should deliver an advantage, the only plausible trade-off scenarios are an improvement in health for the medicine that works in a new way (in cohort 1) and an improvement in health for the medicine that works in a new way accompanied by an additional cost for the medicine that works in a new way (in cohort 2).
Wider societal benefits	***VBP***: Impacts of a product beyond direct health effects. These might include benefits related to reduced reliance on carers, and other wider societal	***All else being equal***: Medicine for disease which causes patients to be reliant upon carers (e.g. family members) for day-to-day needs, and reduces that reliance on carers, compared against medicine for

Prioritisation criteria explored	Reported rationale for use of criteria within UK NHS	Scenario construct
	factors (Dept Health 2010a).	disease that does not cause patients to be reliant upon carers (in cohort 2 only). ***Trade-off scenario***: Medicine for disease which causes patients to be reliant upon carers (e.g. family members) for day-to-day needs, and reduces that reliance on carers, is more costly compared with medicine for disease that does not causes patients to be reliant upon carers (in cohort 2 only).
Disadvantaged populations	***NICE***: The NHS gives special priority to improving the health of the most disadvantaged members of the population, particularly poorer people and ethnic minorities (Rawlins et al., 2010).	***All else being equal***: Medicine for treatment of disease that typically affects disadvantaged populations (e.g. those form low income families) compared against medicine for disease that does not typically affect disadvantaged populations (in cohorts 1 & 2). ***Trade-off scenarios***: As disadvantaged populations is the criterion in question, trade-off scenarios explored smaller health gains (in cohort 1) and higher costs for treatment (in cohort 2) of disadvantaged populations compared with non-disadvantaged populations.
Children	***NICE***: Compilation of evidence and assessment of improvements in the quality of life in children are methodologically challenging. Society would generally favour 'the benefit of the doubt' being afforded to sick children (Rawlins et al., 2010).	***All else being equal***: Medicine for treatment of children compared against medicine for treatment of adults (in cohorts 1 & 2). ***Trade-off scenarios***: As children is the criterion in question, trade-off scenarios explored smaller health gains (in cohort 1) and higher costs for treatment (in cohort 2) for children compared with adults.
End of life treatments	***NICE***: The public generally places special value on treatments that prolong life – even for a few months – at the end of life as long as that extension of life is of reasonable quality (Rawlins et al., 2010). The end-of-life policy specifies that patients should have a short life expectancy, normally of less than 2 years, and the gain in life expectancy over currently available NHS treatment should normally exceed 3 months (NICE, 2009b).	***All else being equal***: Medicine for treatment of fatal disease that leads to death in 18 months without treatment compared against medicine for treatment of fatal disease that leads to death in 60 months without treatment. Both medicines extend life by 6 months (in cohorts 1 & 2). ***Trade-off scenarios***: As patients meeting NICE's end-of-life policy reflect the criterion in question, trade-off scenarios explored smaller life extension of 3 months (in cohort 1) and higher costs for treatment (in cohort 2) for patients with life expectancy of 18 months without

Prioritisation criteria explored	Reported rationale for use of criteria within UK NHS	Scenario construct
		treatment compared with patients with life expectancy of 60 months without treatment.
Cancer treatments	***CDF***: It is possible that society values health benefits to patients with cancer more highly, all else being equal, than benefits to patients suffering other conditions (Dept health 2010c).	***All else being equal***: Medicine for treatment of potentially fatal cancer compared against medicine for treatment of potentially fatal non-cancer disease (in cohorts 1 & 2). ***Trade-off scenarios***: As cancer is the criterion in question, trade-off scenarios explored smaller health gains (in cohort 1) and higher costs of treatment (in cohort 2) for patients with cancer compared with patients with non-cancer disease.
Rare diseases	***AGNSS***: Top-sliced funds for treatments of exceptionally rare diseases in England (NCGHSS 2011). ***SMC***: Policy for appraising orphan drugs (SMC 2010). ***AWMSG***: Policy of appraising ultra-orphan drugs (AWMSG 2011b).	***All else being equal***: Medicine for treatment of rare disease compared against medicine for treatment of common disease (in cohorts 1 & 2). ***Trade-off scenarios***: As rarity is the criterion in question, trade-off scenarios explored smaller health gains (in cohort 1) and higher costs of treatment (in cohort 2) for rare disease compared with common disease.
Stakeholder persuasion	***NICE***: Patients and their advocates...can explain where symptomatology of their condition is poorly reflected in clinical trials and health-related quality of life measure [Rawlins *et al.*, 2010].	No practical scenario construct possible. Not explored in this study.

AGNSS=Advisory Group for National Specialised Services; AWMSG=All Wales Medicines Strategy Group; CDF=Cancer Drugs Fund; NHS= National Health Service; NICE=National Institute for Health and Clinical Excellence; SMC=Scottish Medicines Consortium; VBP=Value-Based Pricing

We used a choice-based format in which adult members of the general public were asked to express their preferred way for the NHS to allocate resources between two competing hypothetical populations. Respondents selected one of 11 alternative resource configurations ranging from all money to be spent on one population, through an equal distribution, to all money to be spent on the alternative population, as illustrated in Figure 5.1 (p118) using disease severity as an example criterion (see Appendix to Chapter 5 for further details).

The descriptions of the populations and their treatments were constructed to isolate as far as possible the influence of the criterion in question. We initially constructed a single questionnaire consisting of three-part questions: a scenario of all else being equal and two subsequent trade-offs. This was piloted amongst a convenience sample of 23 adults with a broad range of educational and occupational backgrounds. None of the pilot respondents reported difficulties in understanding the question framing, terminology or task required; however, to reduce respondent burden and completion time, the questionnaire was subsequently divided into two versions.

Each question of the final version of the questionnaires consisted of two parts. Part 1, common to both questionnaires, represented a scenario of "all else being equal" in which only the criterion in question differed between the competing populations; the costs and effectiveness of treatment and all other aspects of the underlying condition were identical. Part 2 differed between the two versions of the questionnaire, and was included to test if any preferences for the criterion under a scenario of "all else being equal" were retained under less favourable effectiveness and/or cost conditions. In cohort 1 we explored preferences when faced with a trade-off in total health benefits but retained the assumption of equal costs; whereas in cohort 2 we explored preferences when faced with a two-fold change in costs, which in the context of a fixed NHS budget represents a trade-off in the total number of patients who could be treated.

5.3.2 Administration

We commissioned VisionCritical Research Solutions (UK) Ltd to administer the two web-based questionnaires simultaneously to a broadly UK-representative sample of its active survey panel based in England, Scotland and Wales in August 2011. There are no formal methods of sample size calculation for this type of study. We therefore determined our target sample size by reference to those reported in the empirical ethics literature (Dolan et al., 2005; Shah 2009) and available resources. As the survey was closed when our

target of 2000 complete responses was achieved in each independent cohort, it is not possible to determine a response rate.

The choice-based questions were presented to participants in random order to minimise the impact of ordering and learning effects across the cohorts (McColl et al., 2001). An initial 100 panellists acted as an internal pilot to confirm respondents were able to complete the questionnaires properly and within a reasonable timeframe before the survey invite was distributed more widely.

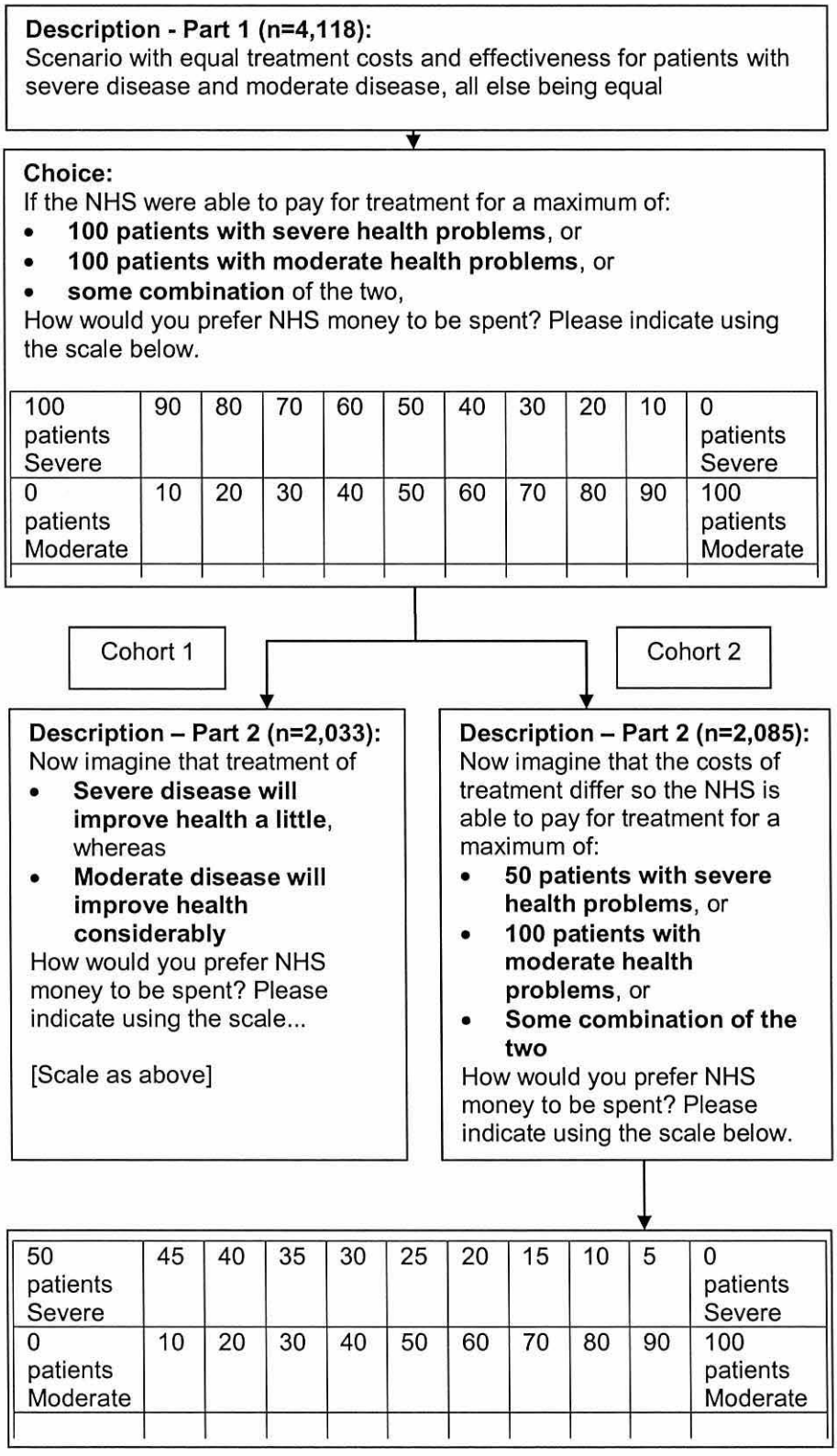
5.3.3 Analysis

Our primary null hypothesis was that, all else being equal, there would be no societal preferences for any of the criteria explored, i.e. most respondents would prefer the NHS to divide resources equally between the competing populations. Secondary hypotheses were that, when faced with a trade-off in total health benefits, respondents in cohort 1 would prefer the NHS to fund treatment that resulted in greater overall health benefits, and when faced with a trade-off in costs, respondents in cohort 2 would prefer the NHS to fund the treatment that enabled most patients to be treated. Societal preferences were inferred from absolute majority response.

Given the consistency of the findings across the two cohorts (see Appendix to Chapter 5 Figure 5.1A, p218), responses to part 1 questions were meta-analysed using a conservative random effects model. Analyses of part 2 questions were conducted separately by categorising responses into three groups: respondents favouring either one of the two competing populations, or respondents favouring an equal division between the two competing populations. Liddell exact test for matched pairs was used to determine the statistical significance of any relative shifts in preferences between both parts of each question.

Socio-demographic data were collected to assess the generalisability of the sample. Logistic regression modelling using age, health status, working status, country of residence and scenario-specific explanatory variables was conducted to determine their impact on respondents' expressed preferences. Analyses were performed in StatsDirect statistical software version 2.7.7, 2009 (StatsDirect Ltd, England).

Figure 5.1. Summary of questionnaire format, using disease severity as an example criterion



5.4. Results

A total of 4,118 adults completed the questionnaires. Respondents' demographics were well balanced across the two cohorts, and were representative of the population of Great Britain, with the exception of a lower proportion of respondents describing themselves as being in very good or good health and a higher proportion describing themselves as being in fair health. People aged 65 years and older were possibly under-represented, and people aged 45 to 64 years slightly over-represented (Table 5.2, p121). Residents of Northern Ireland, who represent less than 3% of the UK population, were not included amongst those surveyed.

5.4.1 Preferences under assumption of "all else being equal"

Pooled responses to part 1 questions are presented in Table 5.3 (p122). All else being equal, a societal preference (based on an absolute majority) for allocation of NHS funds exists for treating patients with severe rather than moderate disease; for treating diseases where there are no alternative treatments available rather than diseases where several alternative treatment options exist; and for treating diseases that cause patients to be reliant upon carers rather than diseases that do not. For all other criteria, between 62% and 85% of respondents' allocations did not support a value premium.

5.4.2 Preferences under health gain trade-offs

Using each cohort's preferences under the assumption of all else being equal as a baseline, when faced with a trade-off in effectiveness, there was a statistically significant shift in preferences for all criteria towards the populations that gained a considerable improvement in health and away from the population that gained a little improvement in health (Liddell exact test $p<0{\cdot}0001$ in each case) (Table 5.3). A preference for treating diseases where there are no alternative treatments available remained present (proportion of respondents, [95% CI]) (41.4% [39.3% to 43.6%]), despite the assumption of a little health gain in that patient group compared to a considerable health gain in patients with several treatment options available (22.3% [20.5% to 24.2%]). A preference in favour of medicines that work in a new way was only apparent when coupled with a considerable improvement in health gains (63.1% [60.0% to 65.2%]). Treatment of a common disease that produces considerable improvements in health gains was also strongly preferred (57.3% [55.1% to 59.4%]) to treatment of a rare disease that produces

a little improvement in health (10.4% [9.1% to 11.8%]). There was no evidence of support of a value premium for any other criteria under effectiveness trade-off conditions.

5.4.3 Preferences under cost trade-offs

When faced with a trade-off in costs, there was a statistically significant shift in preferences for all criteria towards the populations that were more costly to treat (Liddell exact test $p<0.0001$ in each case), with the exception of severity of disease (60.9% vs. 60.0%; RR=1.12; p=0.3193). This resulted in a significantly greater proportion of respondents expressing a preference for the most costly population than expressed a preference for either an equal division of resources or for the less costly population, with the exception of rarity of disease.

5.4.4 Impact of respondents' characteristics on preferences

Logistic regression analyses suggest that respondents' preferences are influenced by their individual characteristics and circumstances (see Table 5.4, p124). For example, those with children in their household were more likely to express a funding preference for treating children over adults than those without (OR 1.63 [95% CI 1.41 to 1.89]), those with a household reliance on carers were more likely to express a funding preference for medicines with wider societal benefits (e.g. reduced reliance upon carers) than those without (OR 1.30 [95% CI 1.03 to 1.64]), and those in social grade C2DE were more likely to prioritise disadvantaged populations (e.g. those on low incomes) than those in social grade ABC1 (OR 1.36 [95% CI 1.19 to 1.55]).

Other observed funding preferences are less easy to explain; for example, compared with respondents rating themselves as in good/very good health, respondents rating themselves as in bad/very bad health were significantly less likely to favour the funding of medicines for severe diseases, medicines for conditions with no other treatment options, and medicines for children. There were no significant differences in preferences for any criterion based on country of residence.

Table 5.2. Socio-demographic characteristics of respondents and adult population of Great Britain

Characteristics	Cohort 1 (n,(%))	Cohort 2 (n,(%))	Great Britain (%)
Gender*			
Male	1026 (50.5)	1000 (48.0)	48.8
Female	1007 (49.5)	1085 (52.0)	51.2
Age*			
18-44	926 (45.5)	903 (43.3)	48.8
45-64	846 (41.6)	900 (43.1)	31.7
65 and over	261 (12.8)	282 (13.5)	19.5
Social grade†			
ABC1	1049 (51.6)	1067 (51.2)	55
C2DE	984 (48.4)	1018 (48.8)	44
Working statusΔ‡			
Employed	1089 (53.6)	1120 (53.7)	58.0
Unemployed	152 (7.5)	144 (6.9)	8.1
Economically inactive¶	792 (38.9)	821 (39.8)	36.9
General health*			
Very good / good	1321 (65.0)	1369 (65.7)	79.0
Fair	514 (25.3)	536 (25.7)	15.0
Bad / very bad	198 (9.7)	180 (8.6)	6.0
Household composition*			
With children	582 (27.9)	603 (29.7)	25.0
Without children	1503 (72.1)	1430 (70.3)	75.0
Household reliance on long-term informal care			
Yes	442 (21.7)	435 (20.9)	Unknown
No	1591 (78.3)	1650 (79.1)	Unknown
Country‡			
England	1749 (86.0)	1761 (84.5)	86.4
Scotland	186 (9.1)	209 (10.0)	8.6
Wales	98 (4.8)	115 (5.5)	5.0

*Figures for Great Britain based on adults aged 16 years and over in Office for National Statistics General Lifestyle Survey 2009 (ONS 2009)
†Figures for Great Britain based on NRS 2010 population data (NRS 2010)
‡Figure for Great Britain based on Office for National Statistics mid-2010 population estimates for adults aged 16 years and over (ONS 2011)
ΔFigures for Great Britain based on nomis official labour force statistics, seasonally-adjusted percentage of people aged 16 years and over, June-August 2011 (covering our survey period) (ONS 2010)
¶People who are neither in employment nor unemployed (e.g. those who were looking after a home or retired)

Table 5.3. Preferences of respondents under assumption of all else being equal and when faced with trade-offs in health gains and costs

Scenario population 1	Choice	Prioritise population 1	Equal allocation to both populations	Prioritise population 2	Choice	Scenario population 2
		% Respondents (95% CI)				
Severe disease	All else being equal*	**59.6 (58.1 to 61.1)**	31.0 (28.0 to 34.0)	9.4 (6.0 to 13.5)	All else being equal*	**Moderate severity disease**
	Little health improvement	28.2 (26.2 to 30.2) RR 0.12; p<0.0001	**42.6 (40.5 to 44.8)** RR 2.84; p<0.0001	29.2 (27.2 to 31.2) RR 5.76; p<0.0001	Improves health considerably	
	Twice the cost of population 2	**60.9 (58.8 to 63.0)** RR 1.12; p=0.3193	30.2 (28.2 to 30.2) RR 0.75; p=0.0101	8.9 (7.7 to 10.2) RR 1.53; p=0.0176	Half the cost of population 1	
No other medicines available	All else being equal*	**56.5 (53.8 to 59.1)**	31.1 (28.9 to 33.1)	12.5 (11.5 to 13.5)	All else being equal*	**Several other medicines available**
	Little health improvement	**41.4 (39.3 to 43.6)** RR 0.22; p<0.0001	36.3 (34.2 to 38.4) RR 1.86; p<0.0001	22.3 (20.5 to 24.2) RR 3.11; p<0.0001	Improves health considerably	
	Twice the cost of population 2	**60.4 (58.3 to 62.5)** RR 2.22; <0.0001	27.2 (25.3 to 29.2) RR 0.48; p<0.0001	12.4 (11.0 to 13.9) RR 0.93; p=0.6753	Half the cost of population 1	
Medicine works in a new way	All else being equal*	24.4 (22.1 to 26.9)	**62.2 (60.7 to 63.6)**	13.4 (12.3 to 14.5)	All else being equal*	**Medicine works in similar way to existing medicines**
	Improves health considerably	**63.1 (60.9 to 65.2)** RR 16.83; p<0.0001	29.1 (27.2 to 31.1) RR 0.10; p<0.0001	7.8 (6.7 to 9.1) RR 0.35; p<0.0001	Little health improvement	
	Improves health considerably and twice the cost of population 2	**53.8 (51.6 to 56.0)** RR 8.98; p<0.0001	32.8 (30.8 to 34.9) RR 0.13; p<0.0001	13.4 (11.9 to 13.9) RR 0.93; p=0.5399	Little health improvement and half the cost of population 1	
Patients reliant on informal carers	All else being equal†	**50.0 (47.8 to 52.1)**	40.6 (38.5 to 42.8)	9.4 (8.2 to 10.7)	All else being equal†	**Patients not reliant on informal carers**
	Twice the cost of population 2	**54.8 (52.7 to 57.0)** RR 1.57; p<0.0001	34.3 (32.3 to 36.4) RR 0.53; p<0.0001	10.8 (9.5 to 12.3) RR 1.34; p=0.0443	Half the cost of population 1	
Disadvantaged population	All else being equal*	34.5 (32.7 to 36.2)	**59.5 (57.6 to 61.5)**	6.0 (5.3 to 6.7)	All else being equal*	**Not disadvantaged population**
	Little health improvement	23.2 (21.3 to 25.1) RR 0.33; p<0.0001	**47.9 (45.6 to 50.1)** RR 0.44; p<0.0001	29.0 (27.0 to 31.0) RR 14.43; p<0.0001	Improves health considerably	
	Twice the cost of population 2	**52.8 (50.6 to 54.9)** RR 6.33; <0.0001	38.5 (36.4 to 40.6) RR 0.14; p<0.0001	8.8 (7.6 to 10.1) RR 2.27; p<0.0001	Half the cost of population 1	

Scenario population 1	Choice	Prioritise population 1	Equal allocation to both populations	Prioritise population 2	Choice	Scenario population 2
		% Respondents (95% CI)				
Children	All else being equal*	37.5 (36.1 to 39.0)	**57.0 (55.5 to 58.5)**	5.5 (4.6 to 6.5)	All else being equal*	Adults
	Little health improvement	19.1 (17.4 to 20.9) RR 0.15; p<0.0001	**44.5 (42.3 to 46.7)** RR 0.48; p<0.0001	36.4 (34.4 to 38.6) RR 20.97; p<0.0001	Improves health considerably	
	Twice the cost of population 2	**54.8 (52.6 to 56.9)** RR 4.86; p<0.0001	38.3 (36.2 to 40.4) RR 0.18; p<0.0001	6.9 (5.9 to 8.1) RR 1.93, p=0.0005	Half the cost of population 1	
18 months survival without treatment	All else being equal*	34.4 (30.4 to 38.6)	**47.6 (46.1 to 49.2)**	17.9 (15.5 to 20.5)	All else being equal*	60 months survival without treatment
	3 month survival gain	23.3 (21.5 to 25.2) RR 0.30; p<0.0001	**50.5 (48.3 to 52.7)** RR 1.21; p=0.0550	26.2 (24.3 to 28.2) RR 2.3; p<0.0001	6 month survival gain	
	Twice the cost of population 2	42.1 (40.0 to 44.2) RR 1.78; p<0.0001	39.2 (37.1 to 41.3) RR 0.47; p<0.0001	18.8 (17.1 to 20.5) RR 1.41; p<0.0001	Half the cost of population 1	
Cancer	All else being equal*	30.8 (28.1 to 33.5)	**64.1 (61.5 to 66.7)**	5.1 (4.5 to 5.8)	All else being equal*	Non-cancer disease
	Little health improvement	20.8 (19 to 22.6) RR 0.30; p<0.0001	**42.0 (39.8 to 44.1)** RR 0.23; p<0.0001	37.3 (35.2 to 39.4) RR 17.74; p<0.0001	Improves health considerably	
	Twice the cost of population 2	**47.5 (45.4 to 49.7)** RR 4.82; p<0.0001	42.2 (40.1 to 44.4) RR 0.15; p<0.0001	10.3 (9.0 to 11.6) RR 3.13; p<0.0001	Half the cost of population 1	
Rare disease	All else being equal*	15.1 (14.0 to 16.2)	43.2 (40.5 to 45.9)	41.7 (38.2 to 45.3)	All else being equal*	Common disease
	Little health improvement	10.4 (9.1 to 11.8) RR 0.45; p<0.0001	32.4 (30.3 to 34.4) RR 0.39; p<0.0001	**57.3 (55.1 to 59.4)** RR 5.54; p<0.0001	Improves health considerably	
	Twice the cost of population 2	23.7 (21.9 to 25.6) RR 3.00; p<0.0001	38.0 (35.9 to 40.1) RR 0.52; p<0.0001	38.3 (36.2 to 40.4) RR 0.82; p=0.0784	Half the cost of population 1	

* Pooled results of cohorts 1 and 2 (n=4,118) using proportion meta-analysis, random effects model; † Reliance on carers explored in cohort 2 only, (n=2,085); RR=Relative risk point estimate based on Liddell exact test for matched pairs, used to compare the proportion of responses under trade-off conditions versus each cohort's responses to part 1 of each question.
Bold figures represent allocations with clear absolute majority and no overlap of confidence intervals; 95% CI = 95% Confidence interval

Table 5.4. Logistic regression analyses under assumption of equal health gains and costs, odds ratios and 95% confidence intervals

Explanatory variables	Dependent variables – Favoured versus (Equal & Not Favoured)								
	Severe disease	No alternative treatment options	Medicines work in new way	Reliance on carers	Disadvantaged populations	Children	18 months survival	Cancer	Rare disease
General explanatory variables – considered in all scenarios									
Age: 45–64 yrs	**1.15 (1.01 to 1.32)**	**1.16 (1.02 to 1.33)**	0.99 (0.85 to 1.16)	**1.23 (1.02 to 1.49)**	0.90 (0.78 to 1.04)	1.05 (0.91 to 1.21)	0.89 (0.77 to 1.03)	1.14 (0.98 to 1.32)	0.98 (0.81 to 1.18)
Age: ≥65 yrs	**1.35 (1.10 to 1.67)**	**1.26 (1.03 to 1.56)**	1.00 (0.79 to 1.27)	**1.51 (1.13 to 2.02)**	1.07 (0.8668 to 1.32)	**1.48 (1.19 to 1.83)**	0.85 (0.69 to 1.06)	**1.39 (1.12 to 1.72)**	1.10 (0.84 to 1.44)
General Health: Fair	0.97 (0.84 to 1.13)	0.91 (0.78 to 1.05)	0.95 (0.80 to 1.12)	0.82 (0.67 to 1.01)	1.04 (0.89 to 1.21)	0.90 (0.78 to 1.05)	0.96 (0.82 to 1.11)	1.06 (0.84 to 1.35)	1.20 (0.99 to 1.47)
General Health: Bad/Very bad	**0.80 (0.64 to 1.00)**	**0.60 (0.48 to 0.75)**	0.78 (0.59 to 1.02)	0.74 (0.53 to 1.04)	1.22 (0.97 to 1.53)	**0.77 (0.61 to 0.98)**	0.84 (0.66 to 1.07)	1.08 (0.86 to 1.37)	1.18 (0.87 to 1.59)
Working status: Employed	0.98 (0.85 to 1.12)	**0.77 (0.68 to 0.88)**	0.99 (0.85 to 1.16)	1.07 (0.88 to 1.29)	**0.85 (0.74 to 0.98)**	**0.85 (0.75 to 0.98)**	1.01 (0.88 to 1.16)	0.93 (0.81 to 1.07)	0.85 (0.71 to 1.02)
Country: Wales	0.98 (0.74 to 1.31)	0.81 (0.61 to 1.07)	0.98 (0.71 to 1.35)	1.30 (0.88 to 1.90)	0.96 (0.71 to 1.28)	1.13 (0.85 to 1.51)	0.97 (0.72 to 1.30)	0.80 (0.59 to 1.10)	0.90 (0.60 to 1.35)
Country: Scotland	0.97 (0.79 to 1.20)	1.20 (0.97 to 1.48)	1.04 (0.82 to 1.32)	1.16 (0.87 to 1.55)	0.90 (0.72 to 1.13)	1.10 (0.89 to 1.37)	0.84 (0.67 to 1.05)	1.01 (0.80 to 1.26)	1.25 (0.95 to 1.64)
Scenario-specific explanatory variables									
Children in household: Yes	-	-	-	-	-	**1.63 (1.41 to 1.89)**	-	-	-
Reliance on carers: Yes	-	-	-	**1.30 (1.03 to 1.64)**	-	-	-	-	-
Social grade: C2DE	-	-	-	-	**1.36 (1.19 to 1.55)**	-	-	-	-

	Severe disease	No alternative treatment options	Medicines work in new way	Reliance on carers	Disadvantaged populations	Children	18 months survival	Cancer	Rare disease
Model X^2, p-value	0.0575*	<0.0001	0.7908*	0.0161	<0.0001	<0.0001	0.1990*	0.0152	0.0683
Deviance goodness of fit X^2, p-value	0.0885	0.2043	0.1082	0.1220	0.1516	0.0561	0.2765	0.0105*	0.5012

Base values of explanatory variables:
Age: 18–44yrs; General health: Good/Very good; Working status: Unemployed and Economically inactive; Country: England; Children in household: No; Reliance on carers: No; Social grade: ABC1;
Each Favoured dependent variable regressed against all general explanatory variables. Scenario specific variables added and retained in respective models if provided a good fit based on Model X^2 and/or Deviance goodness of fit X^2 and McFadden's pseudo-R^2
*Ideally Model X^2 $p<0.05$ and Deviance goodness of fit X^2 $p>0.05$
No evidence of collinearity among independent variables as assessed by tolerance statistics

5.5. Discussion

Our study suggests, all else being equal, that severity of disease, diseases for which no other available treatments exist (representing unmet needs), and medicines that reduce reliance on informal carers (representing wider societal benefits) are supported by society as valid NHS resource prioritisation criteria. In the absence of other differences in patient or disease characteristics, or treatment effectiveness or costs, there were no preferences for any of the other prioritisation criteria we explored.

Under health benefit trade-off conditions there was, in all cases, a statistically significant shift in preferences towards the populations that gained a considerable improvement in health and away from the population that gained a little improvement in health, as we hypothesised. However, counter to our hypothesis, under cost (patient number) trade-off conditions, there was, with the exception of severity of disease, a statistically significant shift in preferences for all criteria towards the populations that were more costly to treat. Unless a preference was apparent under the scenario of "all else being equal", the most plausible interpretation of these cost trade-off findings is that respondents are expressing a general preference for fairness in access to treatment based on need, irrespective of ability to benefit or cost, rather than a preference for the criterion in question *per se*. This is evident in the distributions of actual budget allocations made by respondents (see Appendix for Chapter 5 Figures 5.2A).

For those criteria for which a societal preference was found under conditions of all else being equal, the distribution of budget allocations when costs were doubled were similar to when costs were equal between the competing populations, consistent with a clear preference for these criteria. However, for the remaining criteria, the actual budget allocations suggest that the cost difference causes a shift in budget allocations that peaks where around 70% of the budget is allocated to the most costly population. This is the nearest point to an equal split in patient numbers that our budget allocation scale would permit.

Our study therefore demonstrates that preferences are sensitive to the health gains that may be realised and the number of patients who may be treated, in contrast to our primary hypothesis that was grounded in the utilitarian view of population health (QALY) maximisation. Equity-efficiency trade-offs are being made by respondents, which may be driven by genuine specific social (or private) value judgements and/or more general, egalitarian principles of fairness.

5.5.1 Policy Implications

5.5.1.1 Value-Based Pricing

Our study provides compelling evidence of societal support for all four proposed value-based pricing criteria for rewarding new medicines with higher prices. Given that the Government's consultation on the VBP system generated only eight (4%) responses from individual members of the public (Dept Health 2011), the findings of support for the VBP criteria in our study, based on a sample of over 4,000 members of the public, is reassuring.

Our study was not intended to provide specific weights or to define the levels at which the proposed VBP criteria should be applicable. These are among many other operational issues that remain to be resolved before adoption of VBP in the UK in 2014 (Hughes 2011; Webb 2011). However, our study does confirm the importance of, and societal support for, the proposed criteria for which such weights and levels may need to be determined.

Based on our findings, medicines that work in new ways are only valued above others when they produce a substantial health gain, and society is, at least in principle, supportive of the NHS paying more for innovative medicines that deliver substantial additional health benefits. But with median QALY gains observed in past AWMSG (Chapter 3) and SMC (Walker et al., 2009) submissions being of the order of only 0.1 QALYs, and evidence from published cost utility analyses suggesting incremental benefits of new interventions are decreasing over time [Greenberg et al., 2010], it remains to be seen how many new medicines that are declared to be innovative by manufacturers will be rewarded as such under VBP.

5.5.1.2 NICE criteria

NICE suggests that the six criteria it has identified to date as warranting special consideration in resource allocation decisions reflect societal preferences, as they are based predominantly on the views of its Citizen's Council [Rawlins et al., 2010]. Although disease severity and significant innovation were supported in our study, we observed no compelling evidence for the three other prioritisation criteria we explored (disadvantaged populations, children and patients at the end of life). In reference to its end-of-life criterion, NICE states that the public generally places special value on treatments that prolong life at the end of life as long as that extension is of reasonable quality [Rawlins et al., 2010]. However, the Citizens Council report *Departing from the threshold*, 2008, indicates that,

whilst 24 out of 29 (83%) council members favoured special consideration for treatments that are life saving, only 10 (34%) supported this view for treatments that are life extending (NICE 2008c). This is consistent with our study, in which only 34% of respondents favoured prioritising patients with a reduced life expectancy in the absence of any other differences. Calls for a more systematic and transparent appraisal of medicines (Dept Health 2010c) therefore, seem justified.

5.5.1.3 Cancer Drugs Fund

Based on the anticipated annual costs (8,000 QALYs) and returns (4,000 QALYs) of the CDF (Dept Health 2010c), the Government assumes society values health benefits to cancer patients at least twice as highly, all else being equal, than benefits to patients suffering other conditions. There was no robust empirical evidence in support of this assumption when the CDF was introduced and our study now provides empirical evidence to refute this assumption.

Several reports and studies have highlighted a lower uptake of new cancer medicines (Mason et al., 2010; Richards 2008) and evidence of lower survival rates in the UK compared with other countries (Coleman et al., 2011). However, the Government's consultation document on the CDF points to delayed diagnosis as the main cause of poorer outcomes for cancer patients (Dept Health 2010e). A recent King's Fund report agrees, adding that it is more important to improve access to surgery and radiotherapy, and noting that accessibility of cancer drugs is unlikely to have a significant overall impact (Foot and Harrison, 2011). The consequence of a ring-fenced CDF is that funds are diverted away from services that overall may serve the wider population better, including many patients with cancer (Hughes and Duerden, 2011). Our study, therefore, challenges the rationale for the CDF, which was introduced in England at a time when austerity measures were being actively imposed on other areas of the NHS (Dept Health 2009).

5.5.1.4 Orphan drugs

In addition to being intended for the treatment of rare diseases, medicines that meet requirements for orphan drug designation (Fontain and Hemila, 2000) should address an unmet need, be indicated for life-threatening or seriously debilitating (i.e. severe) conditions, and may also meet a definition of significant innovation, all of which are supported as valid priority-setting criteria in our study. However, new medicines for the treatment of common, serious diseases may also address unmet needs and represent significant innovations (McCabe et al., 2005), so the issue of whether orphan drugs

warrant special funding status would seem to rest on the value attached to rarity of disease.

Our study supports evidence from Norway (Desser et al., 2010), in finding no evidence of a societal preference for treating rare diseases over common diseases. In the absence of other compelling reasons for awarding special funding status to rarity per se (McCabe et al., 2005), the premise of specific orphan and ultra-orphan drug policies appears open to question.

5.5.2 Strengths and limitations

Our study is the first comprehensive empirical analysis of societal views on issues that are central to UK policies on medicines reimbursement. Given the current austerity measures imposed on the NHS, and the imminent reforms of appraisal and reimbursement systems in the UK that have the potential to impact pharmaceutical pricing in other countries, our study is timely and informative for policy-makers and national decision-makers in the UK and further afield.

Our study has a number of methodological strengths. Our sample is broadly representative of the population of Great Britain and uses a larger sample than the UK population survey used to derive the EQ-5D tariff underpinning QALY calculations used by NICE (n=3,395) (Kind et al., 1999). The format adopted for eliciting preferences has potential advantages over a simple binary choice question, by making participants more cognisant of the consequences of their decisions under trade-off conditions. As we explored trade-offs in both health gains and costs, a more complete picture of respondent trade-off behaviours is provided than using either health gains or costs alone.

There are some important caveats, however. Non-completion rates and details of non-responders were unavailable for analysis. This precluded any assessment for potential bias (Johnson and Wislar, 2012). As in all choice-based experiments, participants were asked to make choices between hypothetical scenarios, which inevitably involves simplification of complex decision problems. To avoid imposing our own interpretation of the prioritisation criteria, we constructed scenarios to reflect as closely as possible their definitions in guiding policies and criteria, but these may also simplify decision problems. For example, we defined unmet need in the context of no alternative treatments available, as per the VBP consultation document (Dept Health 2010a); however, the NHS would always provide some level of care, even if that is limited to symptomatic and palliative

care. We also cannot be certain that respondents' preferences are not confounded by their own interpretations of the hypothetical scenarios.

It is possible that a central tendency bias exists in responses; however, we are reassured by the fact respondent preferences, and shifts in preferences under trade-off conditions, are clearly differentiated across the nine criteria explored using the same elicitation method. Our study excludes preferences for situations where multiple criteria may coincide; however, none of the resource allocation criteria identified by NICE and proposed in the VBP consultation document are presented as being contingent on each other, and we still capture all criteria relevant to orphan drug designation, albeit separately.

As the UK NHS is a social insurance scheme that provides health care that is largely free to all at the point of access, the extent to which UK societal preferences would reflect the values of populations in other jurisdictions with other health care systems is unclear. We framed questions to encourage expressions of societal rather than private views, although our analyses suggest that respondent preferences may still be influenced by their personal circumstances. Some of these relationships have obvious, plausible explanations but others, such as those observed for respondents rating themselves as in bad/very bad health, are difficult to explain. Adaptation, comparison processes and cognitive dissonance (Stiggelbout and de Vogel-Voogt, 2008) may confound responses of those experiencing ill health and, as such, these findings should be interpreted with caution.

As may be anticipated due to the web-based research methods, our sample was slightly under representative of people aged over 65 years. Our sample was also possibly under representative of people in employment. Given that those over 65 years of age account for a greater proportion of health and social care spending, and those in employment may feel they contribute to the funding of the NHS via taxation to a greater extent than those not in employment, under-representation of these groups' views could be potentially important. However, the degree to which these groups are under-represented is small. Our conclusions on societal preferences are based on majority views, which may not reflect the intensity or ethical desirability of views. In mitigation, our large sample of respondents permits a broad range of potential views to be expressed.

Lastly, some commentators may consider that responses elicited via self-administered internet-based questionnaires are unreliable; however, cold elicitation methods, such as

ours, may provide more representative accounts of preferences than other methods such as face-to-face interviews or deliberative process (Dolan and Tsuchiya, 2006), which have the potential to distort respondent preferences due to interviewer or group pressure bias (McColl *et al.*, 2001).

5.5.3 Conclusions

The UK NHS has legal and moral obligations to provide fair, comprehensive, needs-based care for all. In doing so it must also provide the best value for tax-payers' money and the most effective, sustainable use of finite resources (Dept Health 2012a). Our study indicates that the public supports trade-offs in equity and efficiency in the allocation of scarce health care resources, but several prioritisation criteria currently imposed upon the NHS by NICE and the government do not reflect societal preferences. This may lead to inappropriate resource allocation decisions with significant population health and economic consequences, the benefactors being pharmaceutical manufacturers who are rewarded with higher prices for their medicines than may be warranted by the benefits they deliver. VBP aims to address these issues, and the proposed criteria for rewarding medicines with higher prices under this system do appear to have societal support.

5.6 Authors' contributions

WGL (the candidate) and DAH (supervisor) conceived the study. WGL designed the survey, commissioned data collection, managed the data and analysed the data. WGL and DAH interpreted the results. WGL drafted the manuscript. WGL and DAH revised the manuscript for intellectual content. WGL finalised the manuscript.

Chapter 6

Review of the effectiveness of NICE's Citizens Council in representing public views and influencing decision-making

6.1 Preface

Chapter 5 presented a large, empirical analysis of public views for the current and proposed prioritisation criteria described above. Based on a choice-based survey of over 4,000 adults in Great Britain, this provided the first comprehensive analysis of current and proposed UK medicines reimbursement criteria.

This Chapter describes a review of the effectiveness of NICE's Citizens Council in reflecting the views of the public and influencing NICE process guidelines. Although there is no gold standard against which to judge the views of the Citizens Council members, the public survey in Chapter 5 provides a basis for comparison of views elicited from two methodological approaches. The findings are expected to be informative for all decision-making bodies planning or undertaking public engagement exercises.

6.2 Abstract

Background: There is growing interest in deliberative public engagement to inform health care decision-making; however evidence of effectiveness of Citizens juries and other deliberative processes is generally lacking. In 2002, the National Institute for Health and Clinical Excellence (NICE) established a Citizens Council, an adaptation of the Citizen jury model, to input public views and inform its decision-making processes. An early review found participation processes were good, but an absence of evidence of effectiveness.

Aims: To explore the effectiveness of NICE's Citizens Council in providing public input to, and informing, NICE decision-making processes.

Methods: Citizens Council reports were examined to determine the level of consistency of members' deliberation-based views towards common themes. These were compared against the results of the large internet-based public survey in Chapter 5, which explored specific prioritisation criteria considered in Citizens Council reports. NICE process documentation was examined to determine the extent to which Citizens Council views are reflected.

Results: Citizens Council views were reliable across reports for 72% of common themes; strict inconsistencies (i.e. polar opposite views) were evident in only 10%. Members' views were consistent with those of the survey-based wider public for three of seven prioritisation criteria, and mixed towards the remainder; there was no prioritisation criterion for which the views of the Citizens Council were strictly at odds with those of the surveyed public. Evidence that Citizens Council views have influenced NICE decision-making processes is lacking.

Conclusions: Without some means against which to check the consistency of views obtained from a given Citizens Council meeting, questions on the reliability of those views as representative of the public's views will remain. Citizens Council meetings are costly to host, yet evidence that they have influenced NICE decision-making processes is lacking. Further research is warranted to explore the use of alternative or supplementary approaches to public engagement which, based on the findings of this review, could reasonably include the use of relatively inexpensive internet-based surveys. Greater transparency on the basis of adoption or rejection by NICE of public views is required.

6.3 Introduction

There are increasing expectations and efforts directed towards public engagement in health care decision-making and priority-setting (Abelson et al. 2003; Mitton et al., 2009; Whitty 2013). While a wide range of methods exists for eliciting public views and preferences, each of which has strengths and limitations, there is no single best method (Ryan, et al., 2001; Rowe and Frewer, 2005). Large, competently designed and conducted public surveys, for instance, can elicit the public's immediate preferences; however, they have been challenged on the point that they do not allow for discussion, reflection and deliberation on the complexities inherent in healthcare prioritisation (Rawlins 2005; Buxton and Chambers, 2011). In contrast, Citizen's juries, involving small numbers of lay volunteers being presented face-to-face with evidence from experts, followed by facilitated deliberation over a number of days, score highly on this point; however, the extent to which the views of a necessarily small group of self-selected volunteers can reflect the diverse views of the general population is open to challenge (Ryan, et al., 2001; Abelson et al., 2003; Rawlins 2005; Buxton and Chambers, 2011), and the deliberation process itself may move the views of respondents away from those of the 'lay' citizen and more towards those of the 'professional' (Abelson et al., 2003; Dolan and Tsuchiya 2006).

The National Institute for Health and Clinical Excellence (NICE) makes compulsory recommendations on the use of medicines and other health technologies in the National Health Service (NHS) in England and Wales. In 2002, NICE established a Citizen's Council, an adaptation of the Citizens jury model, to incorporate public views and inform on the social values that should be applied in its decision-making (Rawlins 2009). The Citizens Council is a standing committee, composed of 30 lay members of the public who meet periodically for three days at a time to deliberate on topics around which NICE requires a public view. Fifteen Citizens Council meeting reports have been published to date, covering topics such as the appropriate definitions of clinical need (CC Report 1), and the circumstances in which NICE should depart from its usual decision rules on cost effectiveness to prioritise the treatment of some patient groups over others (CC Report 11).

An early evaluation of NICE's Citizens Council provided evidence of successful participation of members and their ability to contribute to national debate (Davis et al., 2005). However, as noted in evaluations of other public engagement exercises (Lenaghan, et al., 1999; Mitton et al., 2009), there was little in the way evidence of added

value, or effectiveness in terms of enhanced legitimacy or better decision-making (Davis et al., 2005). Despite increased efforts directed towards public engagement, there has been little focus on the effectiveness of public engagement exercises, or on comparisons of different methods (Lenaghan 1999; Abelson et al., 2003; Mitton et al., 2009).

It is apposite, therefore, to consider the effectiveness of NICE's Citizens Council beyond participation and debate. There is no gold standard against which to judge the views of Citizens Council members, but its purpose is to provide public input to and inform NICE decision-making processes (Rawlins 2009), and it is in these areas that its effectiveness may be explored. The internet-based public survey of NICE prioritisation criteria described in Chapter 5 provides a basis for comparison of two methodological approaches to public engagement. This study therefore aims to ascertain the reliability of Citizens Council views, the congruence of these views with those derived from the public survey, and their incorporation in NICE process guidelines.

6.4 Methods

All 15 of the Citizens Council Reports (CC Reports) published to date were examined for the recommendations or majority views of the members on each separate theme of the topic under discussion (CC Reports 1-15). Themes that were common or closely related and included in two or more CC Reports were compared for the majority view of council members. Themes for which multiple CC Reports provide the same majority views, with no alternative CC reports providing opposing majority views, were labelled as 'consistent' views. Themes for which there were an equal number of CC Reports in which the majority views were polar opposite were labelled as 'inconsistent' views. Themes for which some CC Reports were 'consistent' but others expressed a conflicting majority view were labelled as 'mixed' views. Themes for which it was unclear whether or not views were reflected in more than one CC Report were labelled as such.

NICE's process documents (NICE 2008a; NICE 2008b; NICE 2009b; NICE 2009c; NICE 2012c; NICE 2012d]) were examined for themes common to those included in CC Reports. These themes were labelled as either 'consistent' or 'inconsistent' with CC Reports, depending on whether or not they reflected the majority views expressed in one or more CC Reports, for which there were no alternative CC Reports providing opposing majority views. Themes for which there were 'mixed' views among the CC Reports are necessarily labelled as such for the comparison against NICE process documents.

Circumstances in which NICE may depart from the usual thresholds for cost effectiveness when making health technology recommendations, stated to be based predominantly on the views of NICE's Citizens Council (Rawlins et al., 2010), were compared with the views expressed in CC Reports, in a similar manner as were NICE's process documents, to determine their consistency. Finally, Citizens Council views were further compared with the results of the large internet-based survey presented in Chapter 5, that was conducted specifically to elicit public preferences for these specific 'special circumstances', and other prioritisation criteria.

6.5 Results

On examination, CC Reports 3, 12, 13, and 14 provided no relevant themes to compare the views of the Citizens Council against those expressed in other CC Reports or other documentation, and so were not further considered. CC Report 15 reports on a November 2011 meeting that discussed the circumstances in which the application of different discount rates to future health benefits and costs may be appropriate. This post-dates a clarification to the Guide for the methods of technology appraisal that was issued by NICE to the same effect in July 2011 (NICE 2011), and so was not considered further. From the remaining 10 CC Reports, 90 relevant questions or themes to which the Citizens Council had provided views were identified as relevant to the objectives of this study (see Appendix to Chapter 6).

6.5.1 Consistency among NICE Citizen Council reports

Of the 90 relevant themes, exploration of consistency was possible for 46 (Table 6.1, p139). Among these, Citizens Council members' views were consistent for 33 (72%), and mixed for 8 (17%). The themes around which there were inconsistent or mixed views included: rarity of the condition, the age of patients, the socio-economic characteristics of patients, and the impact of the disease and treatment on length of life (Table 6.2, p140).

6.5.1.1 Rarity of the condition

Citizens Council members in 2002, when deliberating on the definitions of clinical need, concluded that diseases that affect only a small number of patients should be less of a priority than other disease areas, even though their need may be severe (CC Report 1). In contrast, the majority (16/27, 59%) of members in 2004, when deliberating on ultra-orphan drugs, felt that the NHS should be prepared to pay premium prices for drugs to

treat patients with very rare diseases in some circumstances (including the severity of disease) (CC Report 4), and the majority (20/29, 69%) of members in 2008 felt that NICE should be more lenient in its considerations of the cost effectiveness of treatments for rare diseases (CC Report 11).

6.5.1.2 Age of patients

In their deliberations on the definitions of clinical need, Citizens Council members concluded that the age of the patient should be taken into account when appraising health technologies (CC Report 1). This is compatible with the views of the majority (22/29, 76%) of members in 2008 who felt that NICE should be more lenient in its considerations of the cost effectiveness of treatments for children (CC Report 11). However, both are inconsistent with the majority (72-76%) views of members in 2003 who, when deliberating specifically on whether age should be taken into account when NICE makes decisions, concluded that age as a determinant of social roles or how much chance people have had to experience life should not be a relevant consideration. The members in 2003 agreed that age would be a relevant consideration only as a determinant of clinical effectiveness or risk (CC Report 2).

6.5.1.3 Socio-economic characteristics of patients

Citizens Council members in 2002 concluded that social and economic factors should never be a factor in determining clinical need (CC Report 1), which is consistent with the majority view (16/29, 55%) of members in 2008 that patients who are socially disadvantaged should not be a reason for NICE to depart from its usual thresholds for determining cost effectiveness of treatments (CC Report 11). In contrast, a majority (15/26, 58%) of members in 2006 concluded that NICE should issue guidance that concentrates resources on trying to improve the health of the most disadvantaged, to narrow the gap in social inequality, even if that has only a modest impact on the health of the population as a whole (CC Report 7).

6.5.1.4 Impact of the disease and treatment on length of life

Citizens Council members in 2002 concluded that the effect of the disease on the length of life for the individual should be considered in determining clinical need (CC Report 1). Whilst a majority (24/29, 83%) of members in 2008 agreed that NICE should be more lenient in its considerations of the cost effectiveness of treatments that are life saving, only a minority (10/29, 34%) supported the view that NICE should be more lenient in its considerations of the cost effectiveness of treatments that are life extending (CC Report 11).

Table 6.1. Summary of consistency among NICE Citizen Council reports, Chapter 5 public survey results and NICE process documents, n (%) of themes reflected

	NICE CC Reports vs.			
	NICE CC Reports	**Public survey (Chapter 5)**	**NICE process documents**	**NICE 'special circumstances' (Rawlins et al., 2010)**
Consistent	33 (72%)	12 (48%)	16 (50%)	4 (31%)
Mixed	8 (17%)	13 (52%)	10 (31%)	9 (69%)
Inconsistent	5 (11%)	-	6 (19%)	-
Total reflected / tested	46	25	32	13
Not reflected / Not tested	42	63	51	73
Unclear if reflected / tested	2	2	7	4

Consistent = Majority views in CC Reports consistent with majority views in other CC reports or with representation of theme in other documents or with majority views in public survey
Mixed = Evidence of both consistency and conflict in majority views in CC Reports or with representation of theme in other documents or with majority views in public survey
Inconsistent = Majority views in CC Reports conflict with majority views in other CC reports or with representation of theme in other documents or with majority views in public survey

Table 6.2. Inconsistencies among themes in NICE Citizens Council Reports

CC Report	Theme	Inconsistency with / Mixed views*
1. Clinical need, November 2002 (n=30)	What is the number of patients affected? (if only a small number of people are affected by a particular condition, although their need may be severe, this should be less of a priority)	Inconsistent: CC Report 4, 11
	What is the effect of the disease on the length of life for the individual?	Inconsistent: CC Report 11
	What is the age of the patient? (age of a patient should be taken into account)	Mixed views: Consistent: CC Report 11 / Inconsistent: CC report 2
	Social and economic factors should never be a factor	Mixed views: Consistent: CC Report 11 / Inconsistent: CC Report 7
2. Age, November 2003 (n=29)	Age influencing social roles: (22 /29 disagreed relevant)	Inconsistent: CC Report 1, 11
	How much chance people have had to experience life due to their age? "Fair innings"? (21/29 disagreed relevant)	Inconsistent: CC Report 1, 11
4. Ultra-orphan drugs, November 2004 (n=27)	Should NHS consider paying premium prices for drugs to treat patients with very rare diseases? (16/27 [59%] thought yes with certain conditions)	Mixed views: Consistent: CC Report 11 / Inconsistent: CC Report 1
7. Health inequalities, June 2006 (n=26)	NICE should issue guidance that concentrates resources on trying to improve the health of the most disadvantaged members of society to narrow the gap between the least and most disadvantaged, even if this has only a modest impact on the health of the population as a whole? (15/26 [58%] agreed should)	Mixed views: Consistent: CC Report 5 / Inconsistent: CC Report 1, 11
11. Departing from the threshold, November 2008 (n=29)	Reason to depart from threshold- the patients are children 22/29 (76%)	Mixed views: Consistent: CC Report 1 / Inconsistent: CC Report 2
	Reason to depart from threshold- the illness is rare 20/29 (69%)	Mixed views: Consistent: CC Report 4 / Inconsistent: CC Report 1
	Reason to depart from threshold- the intervention will have a major impact on society at large 16 /29 (55%)	Mixed views: Consistent: CC Report 1 / Inconsistent: CC Report 4
	Reason to depart from threshold- the patients concerned are socially disadvantaged 13/29 (45%) [i.e. minority view]	Mixed views: Consistent: CC Report 1 / Inconsistent: CC Report 7
	Reason to depart from threshold- the treatment is life extending 10/29 (34%) [i.e. minority view]	Inconsistent: CC report 1
*Inconsistent = Majority views on themes in CC Reports conflict with majority views in other CC Reports; Mixed views = Evidence of consistency with majority views in some CC Reports and conflict in majority views in other CC Reports		

6.5.2 Consistency of views of NICE Citizens Council and the wider public

The public survey explored the views of 4,118 adults in Great Britain towards nine specific medicines prioritisation criteria that have been used or have been proposed to be used in the future value-based pricing system planned to commence in the UK from 2014 (Chapter 5). Seven of these criteria could be mapped to 27 themes extracted from CC Reports (Table 6.1, p139; Table 6.3, p142). Consistency in the views of Citizens Council members and the public was observed for three of these criteria: severity of disease; unmet need; and (medicines that have) wider societal benefits. The views of Citizens Council members were mixed towards the remaining four criteria explored in the survey (Table 6.3, p145). No strict inconsistencies were evident between the views of the public and the Citizens Council members.

6.5.3 Incorporation of NICE Citizens Council views in process documents and the 'special' circumstances

Of the 32 (36%) themes from CC Reports that feature in the NICE process documents, inconsistent or mixed views were apparent in 50% (Table 6.1, p139). These are dominated by the four broad areas of inconsistency among CC Reports discussed in 6.5.1. Further inconsistencies arise from the current *Guide to the methods of technology appraisal*, which specifies a reference case in which all QALYs gained are viewed equally, irrespective of disease characteristics or the benefactors (NICE 2008a). A notable further inconsistency arises between NICE's Social Value Judgement guide, which categorically rejects the 'Rule of rescue' (NICE 2008b), and CC Reports 1 and 6 that both support this as a relevant consideration in decision-making. NICE's *Supplementary guidance to appraisal committees on appraising life-extending, end-of-life treatments*, which permits more lenient consideration of the cost effectiveness of such treatments (under specific circumstances) (NICE 2009b) is also inconsistent with CC Report 11, which reports a majority of Citizens Council members in favour of prioritising treatments that are life saving, but only a minority in favour of prioritising treatments that are life extending (Table 6.4, p143).

Of the 13 CC Report themes reflected in the 'special circumstances' in which NICE may deviate from the usual thresholds for cost effectiveness, views of the Citizens Council were mixed for 9 (69%). These mixed views relate to the age and socio-economic characteristics of patients, and the priority given to life-extending, end-of-life treatments, as noted in 6.5.1.

Table 6.3. Comparison of NICE Citizens Council views with public survey results for common considerations

Criteria	Public survey results (n=4,118)	NICE CC Report consistency
Severity of disease	59.6% (majority) prioritised treatment of severe disease over treatment of moderate-severity disease	Consistent: CC Report 1, 9, 10, 11 (72%)
Unmet need	56.5% (majority) prioritised treatment of patients with no existing alternative treatments available.	Consistent: CC Report 1, 6, 11 (66%)
Significant innovation	63.1% (majority) prioritised medicines that work in new way and deliver considerable improvement in health (medicines that work in new way without considerable improvement in health not prioritised)	Not reflected in this context
Wider societal benefits	50.0% (majority) prioritised medicines for treatment of patients who are reliant on carers	Consistent: CC Report 1, 9, 11 (76%)
Disadvantaged populations	34.5% (minority) prioritised treatments for disadvantaged populations, most indifferent to socio-economics	Mixed views: Consistent: CC Report 1, 11 (45%) / Inconsistent: CC Report 7 (58%)
Children	37.5% (minority) prioritised children over adults, majority indifferent	Mixed views: Consistent: CC Report 2 (72%-76% view age *per se* irrelevant) / Inconsistent: CC Report 1 (age important), 11 (76% prioritise children)
EoL treatments	34.4% (minority) prioritised treatment of patients with life expectancy of 18 months compared with 60 months, majority indifferent	Mixed views: Consistent: CC Report 11 (34%) / Inconsistent: CC Report 1
Rare disease	15.5% (minority) prioritised rare disease over common disease, majority indifferent or favour treatment of common disease.	Mixed views: Consistent: CC Report 1 / Inconsistent: CC Report 4 (59% with conditions), 11 (69%)

EoL = End-of-life
Public survey percentages refer to percentage of 4,118 members of the public expressing this view
NICE CC Report percentages refer to the percentage of Citizen Council members expressing that view in each CC Report. NB In CC Reports where no percentage is presented, the view / recommendation is assumed to represent the majority view of Citizen Council members
Consistent = Majority views in CC Reports consistent with majority views ofpublic survey
Mixed = Majority views in some CC Reports consistent with majority views ofpublic survey, and majority views in other CC Reports inconsistent with majority views ofpublic survey
Inconsistent = Majority views in CC Reports conflict with majority views ofpublic survey

Table 6.4. Inconsistencies among themes in NICE Citizens Council Reports and NICE process guides

CC Report	Theme	Inconsistency with / mixed views
1. Clinical need, November 2002 (n=30)	Is condition potentially fatal? (saving lives, or the 'rule of rescue' is extremely important)	Inconsistent: SVJ Document
	How bad is the pain and how severe are the symptoms?	Mixed views: Possibly consistent: SVJ Document / Inconsistent/?: Methods of TA
	What is the age of the patient? (age of a patient should be taken into account)	Inconsistent: SVJ Document; Methods of TA
	Social and economic factors should never be a factor	Mixed views: Consistent: Methods of TA / Inconsistent: SVJ document; Methods for PH
4. Ultra-orphan drugs, November 2004 (n=27)	Should NHS consider paying premium prices for drugs to treat patients with very rare diseases? (16/27 [59%] thought yes with certain conditions)	Inconsistent: SVJ Document; Methods of TA
5. CC Report on Mandatory Public Health Measures, July 2005 (n=24)	Where does the balance lie between needs and benefits versus harm and inconvenience?	Mixed views: Consistent: Methods for PH / Uncertain?: SVJ Document
	When and how should the state intervene?	Mixed views: Consistent: SVJ Document / Inconsistent: Methods of TA / Uncertain ?: Methods for PH
6. Rule of rescue, January 2006 (n=27)	Should NICE reject the Rule of Rescue? (21/27 [78%] said not in certain exceptional cases, e.g.Is the intervention required to avoid immediate loss of life?)	Inconsistent: SVJ Document
7. Health inequalities, June 2006 (n=26)	Should NICE issue guidance that concentrates resources on trying to improve the health of the most disadvantaged members of society, thus narrowing the gap between the least and most disadvantaged, even if this has only a modest impact on the health of the population as a whole? (15/26 [58%] agreed should)	Inconsistent: Methods for PH; Methods of TA

CC Report	Theme	Inconsistency with / mixed views
11. Departing from the threshold, November 2008 (n=29)	Reason to depart from threshold- the illness is rare 20/29 (69%)	Inconsistent: SVJ Document
	Reason to depart from threshold- the treatment is life extending 10/29 (34%) [i.e. minority view]	Inconsistent: Supplementary EoL policy
	Reason to depart from threshold- the patients are children 22/29 (76%)	Mixed views: Inconsistent: SVJ Document
	Reason to depart from threshold- the illness under consideration is extremely severe 21/29 (72%)	Mixed views: Possibly consistent: SVJ Document / Inconsistent/?: Methods of TA
	Reason to depart from threshold- the patients concerned are socially disadvantaged 13/29 (45%) [i.e. minority view]	Mixed views: Consistent: Methods of TA / Inconsistent: SVJ Document

Methods of PH = Methods for the development of Public Health guidance, 3rd Ed, 2012
Methods of TA = Guide to the methods of technology appraisal, 2008
NICE PASLU = Process for advising on the feasibility of of implementing a patient access scheme (Interim), September 2009
Supplementary EoL policy = Supplementary advice to appraisal committees: Appriasing life-extending, end-of-life treatments, 2009
SVJ Document = Social Value Judgements: Principles for the development of NICE guidance, 2nd Ed, 2008

6.6 Discussion

6.6.1 Overview of findings

Views were reliable across CC Reports among 72% of common themes; strict inconsistencies (i.e. polar opposite views) were evident in only 10%. The areas of mixed or inconsistent views among the different CC Reports were in relation to the implied value judgements on rarity of the condition, the age of patients, the socio-economic characteristics of patients, and the impact of the disease and treatment on length of life.

There was inconsistency evident between in the views expressed in CC Reports 1 and 2, even though the Council membership was the same for both meetings; however, members' views were consistent in 10 of the 12 common themes in CC Reports 1 and 11 (see Appendix to Chapter 6), despite different membership and a lapse of six years. It is beyond the scope of this study to determine the reasons for this; some degree of inconsistency may be inevitable and expected, due to the context in which themes were deliberated, or the way in which questions are posed to the Citizens Council. However, it seems reasonable to conclude that, without some means against which to check the consistency of views obtained from a given Citizens Council, questions on the reliability of those views as a reflection of the public's views (Abelson et al., 2003; Buxton and Chambers, 2011) will remain.

Citizens Council views were consistent with those of the wider public for three of seven prioritisation criteria: severity of disease, patients with unmet needs, and wider societal benefits. For the remainder, views of the Citizens Council and the wider public were mixed, primarily due to the inconsistencies observed among the Citizens Council's views contained in the CC Reports. There was no prioritisation criterion for which the views of the Citizens Council were strictly at odds with those of the public. On this basis, there was a degree of alignment between the deliberation-based views of the Citizens Council members and the survey-based views of the wider public; however, it is not possible to determine which of the mixed views are 'correct' or most reliable.

The majority of themes identified in the CC Reports do not feature in the NICE process guides. Of those that do, evidence of strict inconsistency against the CC reports is apparent in 19% and mixed views in 31%, mainly in the same areas of inconsistency among the CC Reports (the age and socio-economic characteristics of patients, the rarity

of the condition, and the impact of the disease and treatment on length of life), but also in relation to the relevance of the 'Rule of rescue'.

Ultimately, it is NICE's Board that decides which of the views of its Citizens Council should be adopted, taking account of relevant legislation on discrimination, equality and human rights (NICE 2008b), in addition to the NHS Constitution, Principle 1 of which states: "*The NHS provides a comprehensive service, available to all irrespective of gender, race, disability, age, sexual orientation, religion or belief. It has a duty to each and every individual that it serves and must respect their human rights. At the same time, it has a wider social duty to promote equality through the services it provides and to pay particular attention to groups or sections of society where improvements in health and life expectancy are not keeping pace with the rest of the population*" (Dept Health 2012a).

These legal and constitutional requirements may explain the approach taken by NICE towards age and health inequalities in its Social Value Judgements guide (NICE 2008b), which are not consistent with some of the mixed views observed in CC Reports. However, they do not explain the rejection in the Social Values Judgement guide of the 'Rule of rescue' (supported in CC Reports 1 and 6) as a priority setting criterion, nor the adoption of *Supplementary guidance on appraising life-extending, end-of-life treatments* issued in 2009 (NICE 2009b) (which was not supported in CC Report 11 as a valid priority-setting criterion). Rarity of disease (supported as a valid priority-setting criterion in CC Reports 4 and 11) is referred to in the context of orphan medicines in NICE's Social Values Judgement guide, which states these medicines should be evaluated in the same way as any other treatment, and notes that ultra-orphan drugs were not expected to be appraised by NICE (NICE 2008b). From April 2013, however, NICE's remit will include appraisal of high cost medicines for rare diseases (NICE 2012e).

It is of note that CC Reports 11 to 15 were published after the current edition of the Social Value Judgements guide (NICE 2008b), and it is possible that this may be formally updated in light of these subsequent CC Reports; however, NICE is able to provide supplementary advice/guidance to its appraisal committees outside of a formal update of its key process guides, as demonstrated by its *Supplementary guidance on appraising life-extending, end-of-life treatments* issued January 2009 (NICE 2009b), and its clarification of differential discount rates to be applied to future health benefits, introduced in July 2011 (NICE 2011). The latter of these was deliberated upon by the Citizens Council four months after this clarification was introduction (CC Report 15). Some commentators have noted the potential for deliberative engagement programmes to be

heavily influenced by the organisation of which they are a part [Abelson et al., 2003; Buxton and Chambers, 2011], and despite independent facilitation of Citizens Council meetings, the publically available documentation used in this review do not provide evidence either way.

Collectively, 50% of the themes common to the CC Reports and NICE process documents are consistent, but the extent to which the inclusion of these themes in NICE process documents can be attributed directly to the views of the Citizens Council is not clear. Although the various NICE process documents are relatively explicit in terms of the social value judgements to which they subscribe, there is a lack of transparency surrounding the basis of selection and adoption by NICE's Board of specific social value judgements from the many views expressed across 15 CC Reports.

6.6.2 Comparison with other studies

The mixed views observed in CC Reports towards age/children and disadvantaged populations may be reflective of the mixed evidence from the empirical ethics literature on the validity of age and socioeconomic status as priority setting criteria (Sassi et al., 2001; Dolan et al., 2005); however, the methods employed in empirical studies often vary, complicating interpretation of the collective evidence. There are few studies in the empirical ethics literature exploring whether or not rarity of disease or treatments for patients at the end of life are valid priority-setting criteria. The public survey from Chapter 5, used here as a reference for the views of Citizens Council, is consistent with an earlier Norwegian study (Desser et al., 2010) in finding no evidence of a societal preference for treating rare diseases over common diseases using a similar survey method, and with a Canadian study in a convenience sample of students using a different (discrete choice experiment) approach (Metzakis et al., 2011).

The public survey in Chapter 5 found no public preference for NICE's end-of-life premium, a finding which has been confirmed in a subsequent study using a discrete choice experiment in adults in England and Wales (Shah et al., 2012). So whilst it remains unclear which of the CC Reports providing mixed views towards rare diseases and end-of-life treatments is most reliable, a degree of consistency in support of the public survey findings from Chapter 5 can be claimed, and for those views in the CC Reports that are consistent with this.

6.6.3 Strengths and Limitations

This is the first review to assess NICE's Citizens Council effectiveness beyond participation process and experience. In doing so, it contributes to the literature on deliberative public engagement processes generally (Lenaghan 1999; Abelson et al., 2003; Mitton et al., 2009). By comparing the deliberation-based views of the Citizens Council members contained in several reports against each other and against the survey-based public views in Chapter 5, it has been possible to make inferences on the extent to which NICE's Citizens Council reflects the views of the less informed public it represents.

There are, of course, some important caveats. This review is based entirely on the publically available CC Reports, NICE process documents, and the results of the public survey. It should be expected that each Citizens Council meeting will be context specific and without attending each meeting, it is possible that this context has been missed. Mapping and categorisation of themes as consistent or otherwise across several context-specific CC Reports has been necessary and is, to a degree, subjective and would benefit from further audit. No formal statistical analyses of consistency (reliability testing) have been attempted across CC Reports due to their context-specific nature and the fact that the populations of Citizens Council members contributing to each CC Report are often not independent. As some CC Reports provide several related themes, and there are instances of mixed views between CC Reports, it has also been necessary to count each related theme separately. Although this may appear to introduce an element of double-counting of themes and their consistency, on average, the proportion of themes demonstrating consistency or otherwise would remain similar.

There is no gold standard against which to judge the views of NICE's Citizens Council, and so it is not possible to state with confidence that its views are somehow legitimate towards one theme but not towards another. The comparison of its views against those derived from the public survey has been conducted *post hoc*, rather than in a controlled manner, which may limit its robustness. The fact that only the public survey from Chapter 5 has been used as a reference against which to compare the views of the Citizens Council may also be viewed as a limitation. However, the survey in Chapter 5 was designed specifically to elicit public views for multiple prioritisation criteria used by NICE, in a large demographically representative sample of the public, using a consistent choice-based method. This provides for a more internally valid reference than could be achieved from multiple study comparisons, using multiple population samples and varied methodological approaches. The public survey in Chapter 5 was also the first empirical

study of public preferences for NICE's end-of-life policy, and the first UK study to explore public views on rarity of disease as a priority-setting criterion. Given there are no empirical studies that conflict with its findings on these points, this is the most relevant study to use as the reference.

Finally, the current review stopped short of examining individual NICE recommendations for evidence of the use of Citizens Council-informed social value judgements in its decision-making. The publication that outlines the 'special circumstances' in which NICE may depart from the usual thresholds for costs effectiveness included selected examples of NICE recommendations that reflect those circumstances (Rawlins et al., 2010). However, a comprehensive review of the social value judgements underpinning individual NICE recommendations has recently been published [Shah et al., 2011a] and concluded that, in practice, NICE advisory bodies have been extremely reluctant to depart from the usual decision rules on cost effectiveness on grounds of equity, except in the special case of life-extending end of life treatments (which was not supported in CC Report 11, nor the public survey in Chapter 5, as a valid prioritisation criterion).

6.6.4 Conclusion

NICE has gone further than most HTA organisations to incorporate the views of the public into its decision-making processes, and for that it should be applauded. There has been little previous focus on the effectiveness of public engagement exercises, or on comparisons of different methods of public engagement. This review has found NICE's Citizens Council views were consistent with those of the survey-based wider public for three of seven prioritisation criteria, and mixed towards the remainder; there was no prioritisation criterion for which the views of the Citizens Council were strictly at odds with those of the surveyed public. However, without some means against which to check the consistency of views obtained from any given Citizens Council meeting, questions on the reliability of those views as representative of the public's views will remain.

The first four NICE Citizens Council meetings were estimated to have cost £470,000, with subsequent meetings estimated to cost £80,000 to 90,000 each (2005 prices) (Davis et al., 2005), yet evidence that the Citizens Council views have influenced NICE decision-making processes is lacking. Further research is warranted to explore the use of alternative or supplementary approaches to public engagement, which based on the findings of this review, could reasonably include relatively inexpensive internet-based surveys, which appear to offer a degree of consistency with the deliberation-based views

of NICE's Citizens Council. Greater transparency on the basis of adoption or rejection by NICE of its Citizens Councils' or public views is also required.

6.7 Authors' contributions

WGL (the candidate) conceived the study, extracted data, categorised and mapped themes, analysed data and drafted the manuscript. WGL and DAH critically revised the manuscript for intellectual content. WGL finalised the manuscript.

Chapter 7

Discussion and Conclusions

7.1 Summary

Health care providers have limited resources to meet ever increasing demands. The provision of health care, and use of health care services, involves complex decision-making by a wide range of stakeholders. Relevant criteria and processes are required to ensure that decision-making is appropriate. This thesis explores the preferences of different stakeholders involved in health care decision-making, particularly national-level decision-makers and the public, to determine the relevance of current and proposed criteria and processes for medicines reimbursement in the UK NHS. Seven research questions have been addressed, as summarised below.

Chapter 2 addressed Research question 1:
How and to what extent do key considerations in health care decision-making differ between decision-makers (agents) and those affected by their decisions (principals)? In particular, are there differences in preferences between principals and agents towards benefits and harms of health care interventions, and towards health-related and process-related aspects of health care?

Based on a systematic literature review of conjoint analyses conducted specifically in two or more stakeholders involved in principal-agent type relationships, the preferences of decision-makers and those on whose behalf they act differ towards key health care decision-making factors more often than not. Process-related aspects of health care are important influences on both principals and agents; both appear willing to make trade-offs against health outcomes in order to secure non-health aspects of health care. However, both principals and agents were significantly more likely to rank health-related outcomes ahead of process-related outcomes.

Chapter 3 addressed Research question 2:
What does previous guidance of the All Wales Medicines Strategy Group (AWMSG) reveal about the factors that influence its reimbursement decisions for new medicines?

Univariate analyses indicate that the incremental cost effectiveness ratios (ICER) for medicines with positive AWMSG recommendations were significantly lower than for those with negative recommendations. In addition, there is some evidence that AWMSG's ultra-orphan policy (AWMSG 2011b), which permits more lenient consideration of the cost effectiveness of treatments for very rare diseases, has elevated the recommendations for ultra-orphan drugs to similar rates as non-orphan medicines. However, neither cost

effectiveness nor ultra-orphan status is observed to be a significant predictor of recommendations in multivariate logistic regression analyses. Results of these analyses suggest a positive influence on recommendations of the presence of probabilistic sensitivity analyses (PSA), used to explore combined uncertainty in parameter values, and a counter-intuitive significant negative influence of evidence from high quality randomised controlled trials. There are challenges in applying the usual hierarchies of evidence in the HTA process (Rawlins 2008), and in particular to the appraisal of high cost, specialist medicines close to market authorisation.

Chapter 4 addressed Research questions 3 and 4:
What are the stated preferences of AWMSG appraisal committee members towards specific new medicines adoption criteria? Are the stated preferences of AWMSG appraisal committee members externally valid?

Based on a discrete choice experiment, committee members were significantly influenced by the ICER and the QALY gains delivered by new medicines, whether uncertainty had been assessed using PSA and, depending on the choice model used, the impact of the underlying disease on patients. Committee members were willing to accept higher incremental cost effectiveness ratios and lower QALY gains for medicines that treat diseases impacting primarily upon survival rather than quality of life, and where uncertainty in the cost effectiveness estimates has been explored thoroughly using PSA. On average, appraisal committee members' stated preferences appear consistent with their actual decision-making behaviours. However, prediction of recommendations for individual medicines from parsimonious stated preference models is more limited, due to a multitude of different considerations inherent in HTA processes.

Chapter 5 addressed Research questions 5 and 6:
Do the current UK prioritisation criteria reflect the public's stated preferences for health care resource allocation? Do the proposed criteria for rewarding new medicines with higher prices under the future value-based pricing system in the UK reflect the public's stated preferences for health care resource allocation?

Based on a large empirical study of public preferences, all four criteria proposed for rewarding new medicines under the value-based pricing system (tackle severe diseases, address unmet needs, are innovative, and have wider societal benefits) (Dept Health 2010a) are supported by the public. However, three of five current NICE prioritisation

criteria (end-of-life premium, the prioritisation of children or disadvantaged populations), stated to be based predominantly on the views of its Citizens Council (Rawlins et al., 2010), are not supported by the wider members of the public, nor are the special funding status for treatments of rare diseases as permitted by AWMSG (AWMSG 2011b) and SMC (SMC 2010), nor the assumptions used to justify the introduction of the Cancer Drugs Fund in England (Dept Health 2010c). Preferences are sensitive to the health gains that may be realised and the number of patients who may be treated. Equity-efficiency trade-offs were made by respondents, which may be driven by genuine specific social or private value judgements, and/or more general, egalitarian principles of fairness.

Chapter 6 addressed Research question 7:
How effective is NICE's Citizens Council in reflecting the views of the public and influencing NICE's decision-making?

A comparison of NICE's Citizens Council Reports and the results of the public survey in Chapter 5 indicates Citizens Council members' views are consistent with those of the survey-based wider public for three of seven prioritisation criteria, and mixed towards the remainder; there is no prioritisation criterion for which the views of the Citizens Council are strictly at odds with those of the surveyed public. The mixed views between the Citizens Council and the public arise from the mixed views apparent among the views of the Citizens Council when discussing similar issues in different meetings. The majority of views, or themes, discussed by NICE's Citizens Council do not appear to feature among NICE's process documents, and of those that do, 50% are consistent with the Citizens Council's views. Evidence that Citizens Council views have influenced NICE decision-making processes is lacking.

7.2 Implications

The findings of the studies in this thesis have a number of important implications.

7.2.1 Implications for current reimbursement processes in the UK

- *Health care decision-making processes should involve those affected by the decisions*

This would seem non-contentious; the findings of the systematic literature review in Chapter 2 support the involvement of patients and wider stakeholders at all levels of health care decision-making. NICE, AWMSG and SMC all involve lay representatives among their appraisal committee's membership, and NICE has gone further than most HTA bodies in creating its Citizens Council with the aim of establishing the social values that should inform its decision-making processes (Rawlins 2009).

- *Both decision-makers and the public support the pursuit of efficiency in resource allocation*

Chapters 3 and 4 together indicate that national-level decision-makers (in Wales) have a preference for reimbursing medicines that they consider to be more cost effective and deliver greater health gains, confirming observations of other national decision-makers' preferences (Tappenden et al. 2007; Koopmanschap et al., 2010). In the context of the broad objectives of a tax-funded health service, this would be anticipated. Chapter 5 demonstrates that, all else being equal, the public prefers the NHS to fund treatments that deliver greater health gains, and an acceptance that trade-offs need to be made. It is important to note that the threshold for cost effectiveness should represent the marginal value of health to the NHS; however, there is no empirical basis for the £20,000 to £30,000 per QALY threshold range that is currently suggested in the UK (Rawlins and Culyer, 2004). Nonetheless, the stated preference study in Chapter 4 indicates that, on average, the current threshold range is adhered to by decision-makers in Wales, in line with operational guidelines (AWMSG 2011a).

- *QALYs alone are inadequate as the usual measure of effectiveness employed by national-level decision-makers in the UK. Non-health/process-related aspects of health care have value and should be considered more explicitly*

QALYs have long being criticised for their focus only on the health outcomes of interventions (Mooney 1989; Rawles 1989; ECHOUTCOME 2013). The Kennedy report *Appraising the value of innovation and other benefits: a short study for NICE*, 2009, recognised that QALYs may not capture all relevant benefits, but concluded that the ICER/QALY approach is the best method available and that NICE should make its considerations of other benefits more transparent (Kennedy 2009). Chapter 2 confirms the importance of non-health aspects of health care to both decision-makers (agents) and those people on whose behalf they act (principals). NICE's current *Guide to the methods of technology appraisal* (NICE 2008a), and the draft update to this (NICE 2012b), state that: "*... significant characteristics of healthcare technologies that have a value to people that is independent of any direct effect on health should be noted"* (NICE 2008a; NICE 2012b). However, it remains unclear how these characteristics are to be valued alongside QALYs when noted. Given non-health, process-related aspects of health care have value, more explicit consideration of these within decision-making is warranted.

- *Several current UK prioritisation criteria do not reflect the preferences of the public. Policies introduced on the basis of perceived –and not actual– societal values may lead to inappropriate resource allocation decisions*

To the extent that they are estimated using public valuations of health states, via the UK EQ-5D tariff (Kind et al., 1999) QALYs should already reflect public preferences for different health states. However, society does not value all QALYs equally (Sassi et al. 2001; Dolan 2005; Shah 2009; Bobinac 2012). Chapter 5, which explored the public's views towards current and proposed medicines prioritisation criteria in the UK, lends support to this notion. In many ways, prioritisation criteria represent attempts to address the inadequacies of the QALY-based approach to distributive decision-making; however, there is a lack of public support for three of five prioritisation criteria to have been used by NICE (end-of-life premium, the prioritisation of children or disadvantaged populations), the assumptions used by the UK Government to justify the Cancer Drug Fund (Dept Health 2010c); and for the special funding status for treatments of rare diseases as permitted by AWMSG and SMC (AWMSG 2011b; SMC 2010). The use of unsupported prioritisation criteria may lead to inappropriate resource allocation, with associated health and economic consequences (Collins and Latimer, 2013).

- *Some means against which to check the reliability of small deliberative group views, such as those of NICE's Citizens Council, and their ability to reflect community values, is warranted*

Chapter 6 reports a degree of alignment between the deliberation-based views of NICE's Citizens Council and the survey-based preferences of the public towards specific prioritisation criteria. In cases where views were not aligned, this reflected inconsistencies between the views expressed by the Citizens Council towards the same issues in different meetings (e.g. views on the special funding status for treatments of rare diseases). Without some means against which to check the consistency of views obtained from any given Citizens Council meeting (or any other deliberative public engagement exercise involving small numbers of participants, for that matter), and the extent to which views reflect community values, questions on their reliability and relevance to health care decision-making (Buxton and Chambers 2011) will remain.

- *The basis of NICE's adoption or rejection of its Citizens Council's views needs to be more explicit and transparent*

The three unsupported prioritisation criteria used by NICE were stated to be based predominantly on the views of its Citizens Council (Rawlins et al., 2010). The lack of public support for these raises questions not only around whether the views of NICE's Citizens Council are reflective of the public view, but also whether and how NICE has elected to act upon them. Although there are some instances where NICE's process documentation explicitly cites the views of its Citizens Council (NICE 2008b), in general, there is a lack of transparency around the basis of their adoption or rejection. There may, of course be legitimate reasons for NICE to reject a majority view of its Citizens Council (e.g. legislation and constitutional requirements), but without explicit reasoning across the wide range of issues NICE's Citizens Council has expressed views upon, this is not possible to judge. Evidence that Citizens Council views have influenced NICE decision-making processes is lacking.

- *Recourse to procedural justice does not avoid the need to ensure relevant social value judgements are used in decision-making*

NICE aspires to meet the conditions of Daniels and Sabin's 'accountability for reasonableness' framework, which aim to ensure fair and legitimate procedures for priority-setting (Daniels and Sabin, 1997). It is possible to point to individual features of the NICE (and other decision-makers') appraisal processes and note how these conditions have been met (Rawlins 2005; Daniels 2009). However, it is equally possible to identify areas in which NICE (and others) arguably fail (Schlander 2008). Several current criteria and processes for the reimbursement of medicines in the UK, including those of NICE and other decision-makers, are not supported by the public. In the case of NICE's end-of-life premium, this was not specifically supported by NICE's Citizens Council, but was nonetheless introduced by NICE whilst subscribed to the accountability for reasonableness framework. So irrespective of whether NICE ticks all of its boxes, accountability for reasonableness is not a panacea for distributive justice and resource allocation. Recourse to procedural justice may provide a framework for easier defence of decision-making (Daniels 2009), but it does not obviate the need for social value judgements (Sabik and Lie, 2008).

7.2.2 Implications for future reimbursement processes in the UK

In addition to the above, implications for future processes are suggested below.

- *The proposed value-based system for reimbursing medicines may better reflect the social value judgements of the public; however, it will not address all concerns of the current system. Questions remain on how the new system will operate*

The new system of value-based pricing, to be implemented in the UK from 2014, proposes to reward medicines with higher prices if they tackle severe diseases, address unmet needs, are innovative, or have wider societal benefits (Dept Health 2010a). All were supported as valid priority setting criteria in Chapter 5. These criteria relate to both health- and non-health-related benefits of health care, which Chapter 2 notes are important to both decision-makers and those on whose behalf they act. In theory, therefore, it is possible that the values to be reflected specifically in the new arrangements for reimbursing medicines will be more aligned with the social value judgements of the public, and so may lead to more appropriate decision-making than is the case currently.

However, despite its imminence, details of how the scheme will be operated are still lacking.

It is suggested that these criteria will be taken into account by quantitatively weighting the QALY gains of new medicines [Dept Health 2010a]. Work is underway to estimate the relevant weights using internet-based public surveys (as was the method in Chapter 5), but previous attempts to weight QALYs have floundered due to a range of methodological problems (Baker et al., 2010; Dolan et al., 2008). Non-health and wider benefits of medicines (i.e. those aspects that are missing from the current QALY-based analyses, and which are important to stakeholders – Chapter 2) will be factored into decision-making in terms of their impact on people other than the patient and will consider issues of paid and unpaid labour, formal and informal care consumption, personal consumption and public consumption, which are to be estimated with reference to patients' age, gender, quality of life and their International Classification of Disease (ICD) headline condition (Dept Health 2013). Whether or not these are the correct or only wider benefits of medicines the public would wish to see reflected remains unknown.

As the aim of value-based pricing is to determine an appropriate price, rather than a recommendation for use, an accurate estimate of societal willingness to pay for QALYs will be required. More aspects of the decision-making process will be internalised within economic models than is currently the case [Hughes 2011]. Chapter 4 demonstrates that the quantification of preferences within a parsimonious decision-making model can be predictive *on average* across multiple medicines, but the performance of such models towards any one individual medicine is limited due to their necessary simplification of inherently complex decision-making tasks. Transparency and explicit evidence of relevance for factors that are both internalised and external to models will be essential, for both decision-makers and submitting pharmaceutical companies. At the moment value-based pricing appears to raise more questions than it sets out to address.

- *The Cancer Drugs Fund in England, the Rare Disease Drug Fund in Scotland, AWMSGs ultra-orphan drugs policy and NICEs end-of-life premium should not continue when value-based pricing commences*

The common features of medicines that are prioritised by the current Cancer Drugs Fund, orphan/ultra-orphan drugs policies and NICE's end-of-life policy are that they address severe diseases and (possibly) unmet needs. Both severity of disease and unmet needs are supported as relevant prioritisation criteria by the public and are to be used to

specifically reward medicines with higher prices under value-based pricing. In the absence of any empirical evidence to support cancer, or rarity, or short life extension *per se* as relevant priority setting criteria, the correct weighting of QALYs for severity and unmet needs should mean these medicines require no further weighting under value based pricing. In terms of medicines for rare disease, some would argue that orphan drugs legislation already well rewards these. It is of note that the ongoing work to estimate the relevant QALY weights for use in the value-based pricing system also considers inclusion of an adjustment for end-of-life treatment (Dept Health 2012d), which was not specifically referred to as a criterion in the UK Government's consultation on value-based pricing (Dept Health 2010a).

7.2.3 Other implications

- *Efforts should be made to improve the relevance of priority setting criteria for all health care interventions and services, not just branded medicines*

The focus of this thesis is the relevance of criteria and processes involved in the reimbursement of medicines; however, the importance of non-health related aspects of health care, identified in Chapter 2, was based on stated preferences across a wide range of health care services, and there is no reason to suggest that the social value judgements of NICE's Citizens Council and the public, identified in Chapters 5 and 6, would not extend to non-pharmaceutical-based health care interventions or services. Whilst efforts are being made (such as they are) to address the inadequacies of the current approaches to priority setting around *new* branded medicines, in the form of value-based pricing, there are no equivalent efforts directed towards non-pharmaceutical-based health care interventions or services. It is possible to speculate on reasons for this, such as greater advocacy for new medicines from a large and politically influential pharmaceutical industry; however, it is not clear why new medicines should be a special case and other health care interventions and services, which are funded by the same tax-paying public, should not be evaluated on the same "fairer" basis.

- *Public consultation on complex health care policies can be achieved rapidly using internet-based approaches*

The UK Government's consultation exercises for value-based pricing and the Cancer Drugs Fund in England, involved a number of awareness-raising activities aimed at a range of organisations and invited stakeholders (Dept Health 2011; Dept Health 2010f).

Both consultations asked questions that invoke social value judgements in responses, but few efforts appear to have been made to engage the views of the public. This is reflected in the proportion of responses received from members of the public: 8/188 (4%) and 14/160 (9%) responses to the value-based pricing and Cancer Drugs Fund consultations, respectively (Dept Health 2011; Dept Health 2010f). Given the considerable efforts directed towards public engagement in other areas of health care decision-making and priority-setting (e.g. establishment of NICE's Citizens Council), the lack of public engagement on issues as pertinent as these is perplexing. Chapters 5 and 6 together demonstrate that, with careful design and appropriate piloting, it is possible to consult on complex health care issues with a large, socio-demographically representative sample of the public, rapidly and relatively inexpensively.

- *Different jurisdictions will require their own public engagement programmes*

The UK NHS is a social insurance scheme that provides health care that is largely free to all at the point of access. The extent to which the UK public preferences in this thesis would reflect the social values of populations in other jurisdictions, where alternative health care systems may also predominate, is unclear. However, the conclusion that health care decision-making processes should involve those affected by the decisions would seem relevant in other jurisdictions, given the broad, international coverage of the literature review in Chapter 2. The findings of Chapter 6, which provides one of only a few studies to have assessed the effectiveness of a deliberation-based approach, and to have compared this with an alternative approach, will provide useful insights for any decision-makers considering public engagement exercises.

7.3 Strengths and Limitations

This thesis has addressed important questions surrounding the reimbursement of medicines in the UK, using a range of research methods: a systematic literature review, two studies using secondary data collection and analysis, and two studies using primary data collection and analysis. In doing so it makes several unique contributions to existing knowledge and the literature:

- Chapter 2 provides the first systematic review to characterise the scale of the differences that may exist in the preferences of principals and agents towards key decision-making considerations, and an indication of their relative importance. This was possible by taking an alternative approach to extracting and coding data than was used previously.
- Chapter 3 provides the first revealed preference study of an alternative decision-making body to NICE in the UK.
- Chapter 4 presents the first stated preference study of an alternative decision-making body to NICE in the UK. Data from Chapter 3 contributed to its design and permitted the external validity of stated preferences to be tested, which has been done rarely in the health economics arena generally and never before for a HTA body. This, and the use of both fixed and flexible choice task formats, also provides useful insights for discrete choice experiment methodology.
- Chapter 5 provides the first comprehensive empirical analysis of public views on current and proposed prioritisation criteria that are central to UK medicines reimbursement policies. This was the first empirical analysis of public preferences towards cancer treatments, and it was also the first empirical analysis of UK public views towards the funding of treatments for rare diseases, and NICE's end-of-life policy. The findings have important policy implications: they have been used to support the Welsh Government's rejection of the introduction of a Cancer Drugs Fund in Wales (walesonline 2013); and they validate the social value judgements underpinning the proposed criteria for rewarding medicines under the future value-based pricing system in the UK, and have informed the discussions of the Department of Health's recent methods workshops for value-base pricing.
- Chapter 6 provides the first review to assess the effectiveness of NICE's Citizens Council beyond participation process and experience, which was made possible by the data from Chapter 5. Few studies to date have assessed the effectiveness of public engagement programmes or have compared different approaches to public engagement.

Relevant UK stakeholders' views have been ascertained and explored, and several UK-specific policy- and process-related implications of the findings have been highlighted, in line with the stated aims of the thesis.

Limitations of the thesis include its focus on the dominant existing procedures in the UK. Over the last 20 years, QALY-based cost effectiveness analyses have been confirmed as a key component of HTA and reimbursement decision-making. Many other jurisdictions do not specify the use of QALYs or an ICER threshold (Stafinski et al. 2011a). However, the different approaches to resource allocation reflect political, cultural and funding system differences, in addition to any views on whether cost utility analysis *per se* is appropriate.

In addition, alternative approaches to HTA that have been considered in the UK are not explored specifically. Multi-criteria decision analysis (MCDA) techniques involve quantitative weighting or scoring of specific aspects relevant to decision-making to determine a summary estimate of the overall value of products or services (Thakula 2011; Devlin and Sussex, 2011). These may be seen as a natural progression from current approaches and value based pricing. However, a balance needs to be struck between their improved transparency and consistency compared with the current deliberation-modified systems, and the inflexibility of parsimonious models to account for individual or local factors (Devlin and Sussex, 2011). This latter point is highlighted by the stated preference models in Chapter 4. Ultimately, MCDA techniques require correct identification of relevant criteria, and estimation of relevant weights; in this regard, the implications of current and future processes outlined in this thesis are likely to be applicable to MCDA. The approach to consider only the dominant UK (QALY)-based approach to determine cost effectiveness, and the future value-based pricing system, therefore adequately addresses the aims of the thesis.

These aims were to consider the relevance of the current and proposed criteria for medicines reimbursement. No attempt has been made to explore other potential criteria that the public may feel are relevant for making resource allocation decisions, and no attempt has been made to estimate preference weights for any criterion. Many of the findings of the thesis hinge on the reliability of the results of the public survey in Chapter 5, which is just one study among many in the wider empirical ethics literature. However, that public survey was designed specifically to ascertain preferences towards these very criteria in a large socio-demographically representative sample of the public in the UK, using a consistent choice-based method. This would provide for a more internally valid reference than could be achieved from comparisons across multiple studies, with varied methodological approaches that characterise the wider empirical ethics literature. The public survey in Chapter 5 was also the first empirical study of public preferences for cancer treatments *per se* and NICE's end-of-life policy, and the first UK study to explore

public views on rarity of disease as a priority-setting criterion. The findings on NICE's end-of-life policy have been confirmed in a subsequent UK public study (Shah et al., 2012), and there are no other relevant empirical studies to use as the reference for these other criteria.

The studies contained within the systematic review in Chapter 2 and the empirical methods employed in Chapters 4 and 5 relate to stated preferences of stakeholders towards hypothetical scenarios. These inevitably involve simplification of complex decision-making and may not reflect the choices that would be made in real life. However, for the discrete choice experiment of AWMSG members in Chapter 4, the use of revealed preference data from Chapter 3 alongside the hypothetical choices contributes to its face validity and, on average, actual decision-making behaviours were aligned with their stated preferences. For the public survey in Chapter 5, the choice tasks were constructed to reflect actual policy wording as closely as possible, but are inevitably abstract and hypothetical for members of the public, who are not usually involved in population-based health care decision making.

A range of methodological strengths and limitations specific to each of the five studies are detailed further within their respective Chapters.

7.4 Future research directions

The implications discussed in 7.2 outline unresolved issues and potential areas for future research efforts.

The importance of non-health, process-related attributes of health care to different stakeholders, and a need for explicit consideration of these in decision-making processes has been highlighted. Although it is proposed that value-based pricing will better reflect the value of new medicines, it is unclear whether this will capture all of the non-health aspects of health care that are relevant to patients and the wider public. The systematic review in Chapter 2 could provide be a starting point for exploring the types of non-health/process-related aspects of health care that may be relevant to patients and the public, but concerted efforts to establish these, and the weight they should be given in decision-making, is warranted.

Chapter 5 explored the public's preferences towards actual prioritisation criteria and policies via hypothetical choice tasks. Public preferences for some of these criteria (e.g. cancer, NICE's end-of-life policy, unmet needs, innovative medicines) have not previously been explored empirically. The findings on NICE's end-of-life policy have subsequently been confirmed in a large sample of the public using a more complex stated preference discrete choice experiment (Shah et al., 2012), which helps to validate the approach used in Chapter 5; however, it would be useful to also test preferences for the other criteria to determine their reliability. As the hypothetical choice tasks presented to the public in Chapter 5 reflect abstract scenarios to determine general principles, it would be useful to also assess preferences for the same criteria using more specific decisions, based on real-life case studies, to test the whether the general principles still hold.

Concerted efforts are required to assess different methods and levels of public engagement to determine their reliability and effectiveness (Abelson et al., 2003; Mitton et al., 2009). Chapter 6 highlights there is little evidence that NICE's Citizens Council has impacted upon NICE's decision-making processes based on review of public documentation. It would be useful to repeat the earlier ethnographic study (Davis et al., 2005), now that both the Citizens Council and NICE have built up more experience of the process. It would be particularly useful to ascertain the contribution that Citizens Council members feel they have made to NICE decision-making.

Of course, once value-based pricing has become a reality, this will present opportunities over time for further research on the implementation of preferences within what is intended to be a fairer price setting framework. Public views of whether NHS-funded medicines should be valued differently to other NHS-funded health care services, which will not be covered under value based pricing, would be a valid enquiry.

7.5 Conclusions

This thesis confirms the view that current processes for economic evaluation of medicines are inadequate. The importance of non-health, process-related aspects of healthcare has been demonstrated and should be explicitly considered in decision-making. The involvement of patients and public as stakeholders in medicines decision-making at all levels is supported. NICE has gone to great efforts to incorporate the view of the public in its decision-making processes via its Citizens Council; however, how NICE decides to act upon the views of its Citizens Council is not transparent, and evidence that it has influenced NICE decision-making processes is lacking. Efforts to demonstrate fair decision-making processes do not obviate the need for incorporation of relevant social value judgements in decision-making. Several of the current criteria used by decision-makers in the UK to prioritise the funding of some medicines do not reflect the social value judgements of the public.

Moves towards value-based pricing of medicines are supported; the public considers the proposed criteria for rewarding new medicines with higher prices are relevant. However, it is unclear how value-based pricing will operate, and it will not eliminate all deficiencies in current approaches; the need for explicit, transparent decision-making will be just as great, if not greater. No decision-making process is perfect, but some processes may be fairer than others. It remains to be determined if value-based pricing will indeed lead to fairer decision-making for all stakeholders.

References

Abelson J, Forest P-G, Eyles J, et al (2003). Deliberations about deliberative methods: issues in the design and evaluation of public participation processes. *Social Science and Medicine*; 57: 239-51.

ABPI (2009). *Understanding the 2009 PPRS*. London: Association of the British Pharmaceutical Industry. Available at: http://www.abpi.org.uk/our-work/library/industry/Documents/Understanding%20the%20pprs%202009%20final.pdf (accessed 17 February 2013).

Appleby J (2013). *Spending on health and social care over the next 50 years: why think long term?* London: The Kings Fund. Available at: http://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/Spending%20on%20health%20...%2050%20years%20low%20res%20for%20web.pdf (accessed 17 February 2013).

Arana JE, Leon CJ, Quevedo JL (2006). The effect of medical experience on the economic evaluation of health policies: a discrete choice experiment. *Social Science and Medicine*; 63: 512-24.

Arrow KJ (1950). A difficulty in the concept of social welfare. *The Journal of Political Economy*; 58: 328-46.

Arrow KJ (1963). Uncertainty and the welfare economics of medical care. *The American Economic Review*; 53: 941-73.

AWMSG (2007). *AWMSG criteria for appraising new medicines*. Cardiff: All Wales Medicines Strategy Group. Available at: http://www.wales.nhs.uk/sites3/Documents/371/AWMSG%20criteria%20for%20appraising%20medicines%20updated%20July08.pdf (accessed 30 April 2010).

AWMSG (2008). *Guidance notes to manufacturers on the completion of the therapeutic appraisal forms A & B*. Cardiff: All Wales Medicines Strategy Group. Available from URL: http://www.wales.nhs.uk/sites3/docopen.cfm?orgid=371&id=107877 (accessed 30 April 2010).

AWMSG (2011a). *Guidelines for appraising medicines*. Cardiff: All Wales Medicines Strategy Group. Available at: http://www.wales.nhs.uk/sites3/Documents/371/Guidelines%20for%20appraising%20medicines%20updated%20Sep%2011.pdf (accessed 15 February 2012).

AWMSG (2011b). *AWMSG policy relating to ultra-orphan medicines*. Cardiff: All Wales Medicines Strategy Group. Available at: http://www.wales.nhs.uk/sites3/Documents/371/AWMSG%20policy%20relating%20to%20Ultra-orphan%20Medicines%20Sep%2011.pdf (accessed 09 May 2012).

AWMSG (2011c). *AWMSG Policy on appraising life extending, end of life medicines*. Cardiff: All Wales Medicines Strategy Group. Available at: http://www.wales.nhs.uk/sites3/Documents/371/AWMSG%20Policy%20on%20appraising%20life-extending%2C%20end%20of%20life%20medicines.pdf (accessed 15 February 2012).

Baillie L, Hawe E, O'Neill P, et al (2011). *Health and the use of medicines in primary care in Wales*. Association of the British Pharmaceutical Industry/Office of Health Economics Consulting.

Baker R, Bateman I, Donaldson C, et al (2010). Weighting and valuing quality-adjusted life-years using stated preference methods: preliminary results from the Social Value of a QALY Project. *Health Technology Assessment*; 14: No. 27. Available at: http://www.hta.ac.uk/fullmono/mon1427.pdf (accessed 17 February 2013).

Bateman IJ, Carson RT, Day B, et al (2008). *Choice set awareness and ordering effects in discrete choice experiments*. Centre for Social and Economic Research on the Global Environment. Working Paper EDM 08-1. Available at: http://www.cserge.ac.uk/sites/default/files/edm_2008_01.pdf (accessed 01 March 2012).

Bevan A (1948). Speech to Royal College of Nursing conference, 2nd June. Available at: http://www.nhshistory.net/shorthistory.htm#_edn19 (accessed 08 February 2013).

Beauchamp TL, Childress JF (2001). *Principles of biomedical ethics*, 5th Ed. Oxford: Oxford University Press.

Bech M (2003). Politicians' and hospital managers' trade-offs in the choice of reimbursement scheme: a discrete choice experiment. *Health Policy*; 66:261-75.

Bederman SS, Mohamed NN, Kreder HJ, et al (2010). In the eye of the beholder – preferences of patients, family physicians, and surgeons for lumbar spinal surgery. *Spine*; 35: 108-15.

Bijlenga D, Birnie E, Mol BWJ, et al (2011). Obstetrical outcome valuations by patients, professionals, and laypersons: differences within and between groups using three valuation methods. *BMC Pregnancy and Childbirth*; 11: 93.

Bishop AJ, Marteau TM, Armstrong D, et al (2004). Women and health care professionals' preferences for Down's Syndrome screening tests: a conjoint analysis study. *British Journal of Obstetrics & Gynaecology*; 111:775-9.

Black WC (1990). The cost-effectiveness plane: a graphical representation of cost effectiveness. *Medical Decision Making*; 10(3): 212-5.

Bobinac A, van Excel NJA, Rutten FHF, et al (2012). Inquiry into the relationship between equity weights and the value of the QALY. *Value in Health*; 15: 1119-26.

Brereton L, Vassodevan V (2010). *The impact of the NHS market: an overview of the literature*. London: CIVITAS. Available at: http://www.civitas.org.uk/nhs/download/Civitas_LiteratureReview_NHS_market_Feb10.pdf (accessed 15 February 2013).

Breckenridge A, Woods K, Walley T (2010). Medicines regulation and Health Technology Assessment. *Clinical Pharmacology and Therapeutics*; 87; 152-4.

Breckenridge A (2011). Patients opinions and preferences in drug development and regulatory decision making. *Drug Discovery Today: Technologies*; 8 (1): e11-14.

Bridges JFP, Hauber AB, Marshall D, et al (2011). Conjoint analysis applications in health—a checklist: A report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. Value in Health; 14: 403-14.

Bridges JFP, Searle SC, Selck FW, et al (2012). Designing family-centered male circumcision services: a conjoint analysis approach. *Patient*; 5: 101-11.

Brouwer WBF, Culyer AJ, van Exel NJA, Rutten FFH (2008). Welfarism vs. extra-welfarism. *Journal of Health Economics*; 27 (2): 325-8.
Buxton MJ, Chambers JD. What values do the public want their health care systems to use in evaluation technologies? [Editorial] *European Journal of Health Economics* 2011; 12: 285-8.

Casson SG, Ruiz FJ, Miners A (2013). How long has NICE taken to produce Technology Appraisal guidance? A retrospective study to estimate predictors of time to guidance. *BMJ Open*; 3: e001870.

CC Report 1 / NICE Citizens Council Report 1. *Determining "Clinical Need"* (21-23 November 2002). Available at: http://www.nice.org.uk/media/065/B4/FINALNICEFirstMeeting_FINALReport.pdf (accessed 02 March 2013).

CC Report 2 / NICE Citizens Council Report 2. *NICE Citizens Council Report on Age* (8-10 May and 20-22 November 2003). Available at: http://www.nice.org.uk/media/065/9F/Citizenscouncil_report_age.pdf (accessed 02 March 2013).

CC Report 3 / NICE Citizens Council Report 3. *The attitude of the public to the use of patient information obtained from medical records by the National Confidential Enquiries* (13-15 May 2004). Available at: http://www.nice.org.uk/media/065/91/Citizens_Council_May_2004_Report.pdf (accessed 02 March 2013).

CC Report 4 / NICE Citizens Council Report 4. *Ultra orphan drugs* (18-19 November 2004). Available at: http://www.nice.org.uk/media/065/9F/Citizenscouncil_report_age.pdf (accessed 02 March 2013).

CC Report 5 / NICE Citizens Council Report 5.*Mandatory public health measures* (30 June - 2 July 2005). Available at: http://www.nice.org.uk/media/065/73/NICE_Citizens_Council_Report_Public_Health.pdf (accessed 02 March 2013).

CC Report 6 / NICE Citizens Council Report 6. *Rule of rescue* (28-28 January 2006). Available at: http://www.nice.org.uk/media/065/5D/Rule_of_rescue_report_final_0606.pdf (accessed 02 March 2013).

CC Report 7 / NICE Citizens Council Report 7. *Health inequalities* (8-10 June 2006). Available at: http://www.nice.org.uk/media/065/9F/Citizenscouncil_report_age.pdf (accessed 02 March 2013).

CC Report 8 / NICE Citizens Council Report 8. *"Only in research"* (25-27 January 2007). Available at: http://www.nice.org.uk/media/129/29/OIRReport300407.pdf (accessed 02 March 2013).

CC Report 9 / NICE Citizens Council Report 9. *Patient safety* (7-9 June 2007). Available at: http://www.nice.org.uk/media/A3E/37/CCPatientSafetyReportJune07v1.0.pdf (accessed 02 March 2013).

CC Report 10 / NICE Citizens Council Report 10. *Quality adjusted life years (QALYs) and the severity of illness* (31 January - 2 February 2008). Available at: http://www.nice.org.uk/media/2C3/31/CCReportOnQALYsAndSeverity.pdf (accessed 02 March 2013).

CC Report 11 / NICE Citizens Council Report 11. *Departing from the threshold* (27-29 November 2008). Available at: http://www.nice.org.uk/media/231/CB/NICECitizensCouncilDepartingThresholdFinal.pdf (accessed 02 March 2013).

CC Report 12 / NICE Citizens Council Report 12. *Health innovation and value* (28-30 May 2009). Available at: http://www.nice.org.uk/media/D18/73/CitizensCouncilReportInnovation.pdf (accessed 02 March 2013).

CC Report 13 / NICE Citizens Council Report 13. *Smoking and harm reduction* (15-17 October 2009). Available at: http://www.nice.org.uk/media/4EB/13/CCReportOnHarmReductionUpdated300410.pdf (accessed 02 March 2013).

CC Report 14 / NICE Citizens Council Report 14. *The use of incentives to improve health* (20-22 May 2010). Available at: http://www.nice.org.uk/media/9AF/56/CCReportIncentives.pdf (accessed 02 March 2013).

CC Report 15 / NICE Citizens Council Report 15. *How should NICE assess future costs and health benefits?* (3-4 November 2011). Available at: http://www.nice.org.uk/media/06B/B8/Citizens_Council_report_on_Discounting.pdf (accessed 02 March 2013).

Cheung SW, Aranda D, Driscoll CLW, et al (2010). Mapping clinical outcomes expectations to treatment decisions: an application to vestibular schwannoma management. *Otology & Neurotology*; 31:284-93.

Chim L, Kelly PJ, Salkeld G, et al (2010). Are cancer drugs less likely to be recommended for listing by the Pharmaceutical Benefits Advisory Committee in Australia. *Pharmacoeconomics*; 28: 463-75.

Claxton K, Longo R, Longworth L, et al (2009). *The value of innovation*. NICE Decision Support Unit, May 2009. Available at: http://www.nicedsu.org.uk/PDFs%20of%20reports/Value%20of%20Innovation%20-%20final%20DSU%20report.pdf (accessed 21 March 2013).

Claxton K, Martin S, Soares S, et al (2013). *Methods for the estimation of the NICE cost effectiveness threshold*. Centre for Health Economics Research Paper 81. Available at: http://www.york.ac.uk/media/che/documents/papers/researchpapers/CHERP81_Methods_estimation_NICE_costeffectiveness_threshold.pdf (accessed 23 February 2013).

Clement FM, Harris A, Li JJ, et al (2009). Using effectiveness and cost-effectiveness to make drug coverage decisions: a comparison of Britain, Australia and Canada. *JAMA*; 320: 1437-43.

Coast J (2004). Is economic evaluation in touch with society's health values? *BMJ* 329: 1233-6.

Coleman MP, Forman D, Bryant H, et al (2011). Cancer survival in Australia, Canada, Denmark, Norway, Sweden and the UK, 1995 – 2007 (The International Cancer

Benchmarking Partnership): an analysis of population-based cancer registry data. *Lancet* 377: 127-38.

Collins M, Latimer N (2013). NICE's end of life decision making scheme: impact on population health. *BMJ*; 346: f1363.

Coulter A, Collins A (2011). *Making shared decision-making a reality: No decision about me, without me*. London: The King's Fund; 2011. Available at: http://www.kingsfund.org.uk/sites/files/kf/Making-shared-decision-making-a-reality-paper-Angela-Coulter-Alf-Collins-July-2011_0.pdf (accessed 16 November 2012).

Cross JT, Garrison LP (2008). *Challenges and opportunities for improving benefit-risk assessment of pharmaceuticals from an economic perspective*. OHE Briefing; No. 43. London: Office of Health Economics.

Culyer AJ (1989). The normative economics of health care finance and provision. *Oxford Review of Economic Policy*; 5 (1): 34-58.

Dakin HA, Devlin NJ, Odeyemi IAO (2006). "Yes", "No" or "Yes, but"? Multinomial modelling of NICE decision-making. *Health Policy*; 77: 352-67.

Daniels N, Sabin J (1997). Limits to health care: fair procedures, democratic deliberation, and the legitimacy problem for insurers. *Philosophy and Public Affairs*; 26: 303-50.

Daniels N (2000). Accountability for reasonableness: Establishing a fair process for priority setting is easier than agreeing on principles. (Editorial) *BMJ*; 321:1300–1.

Daniels N (2009). Accountability for reasonableness and the Citizens Council. In: Littlejohns P, Rawlins M (eds). *Patients, the public and priorities in healthcare*. London: Radcliffe Publishing Ltd.

Davis C, Wetherell M, Barnett E, et al (2005). *Opening the box: evaluating the Citizens Council of NICE*. The Open University. Available at: http://www.nice.org.uk/media/A0D/35/Final_evaluation_document_-_as_published_(18-3-05).pdf (accessed 05 February 2013).

Davison SM, Kromm SK, Currie GR (2010). Patient and health professional preferences for organ allocation and procurement, end-of-life care and organization of care for patients

with chronic kidney disease using a discrete choice experiment. *Nephrology, Dialysis, Transplantation*; 25: 2334-41.

De Bekker-Grob EW, Essink-Bok ML, Meerding WJ, et al (2009). Preferences of GPs and patients for preventive osteoporosis drug treatment. *Pharmacoeconomics*; 27: 211-9.
De Bekker-Grob EW, Ryan M, Gerard K (2010). Discrete choice experiments in health economics: a review of the literature. *Health Economics*; DOI: 10.1002/hec.1697.

Dept Health (1998). *A first class service: Quality in the new NHS*. London: Department of Health. Available at: http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4006902 (accessed 18 February 2013).

Dept Health (2009). *NHS Chief Executive's Annual Report for 2008-09*. London: Department of Health. Available at: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_099700.pdf (accessed 15 July 2012).

Dept Health (2010a). *A new value-based approach to the pricing of branded medicines: a consultation*. London: Department of Health. Available at: http://www.dh.gov.uk/en/Consultations/Closedconsultations/DH_122760 (accessed 17 February 2013).

Dept Health (2010b). *Equity and excellence: liberating the NHS*. London: Department of Health. Available at: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/@ps/documents/digitalasset/dh_117794.pdf (accessed 21 February 2013).

Dept Health (2010c). *Impact assessment of a proposal for a Cancer Drugs Fund*. London: Department of Health. Available at: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_120930.pdf (accessed 09 May 2012).

Dept Health (2010d). *Value-based pricing impact assessment*. London: Department of Health. Available at: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_122823.pdf (accessed 09 May 2012).

Dept Health (2010e). *The Cancer Drugs Fund: a consultation*. London: Department of Health. Available at: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_120931.pdf (accessed 09 May 2012).
Dept Health (2010f). *The Cancer Drugs Fund – Government response to consultation*. London: Department of Health. Available at: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_125721.pdf (accessed accessed 09 May 2012).

Dept Health (2011). *A new value-based approach to the pricing of branded medicines – Government response to consultation*. London: Department of Health. Available at: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_128404.pdf (accessed 09 May 2012).

Dept Health (2012a). *The NHS Constitution for England*. London: Department of Health. Available at: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_132958.pdf (accessed 08 February 2013).

Dept Health (2012b). *The Pharmaceutical Price Regulation Scheme: 11th Report to Parliament*. London: Department of Health. Available at: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_132793.pdf (accessed 17 February 2013).

Dept Health (2012c). *Creating change: IHW one year on*. London: Department of Health. Available at: https://www.wp.dh.gov.uk/publications/files/2012/12/Creating-Change-IHW-One-Year-On-FINAL.pdf (accessed 23 February 2013).

Dept Health (2012d). *Technical workshop on QALY weightings: introductory briefing*; November 2012. London: Department of Health.

Dept Health (2013). *Technical workshop on wider societal benefits: briefing document*; February 2013. London: Department of Health.

Desser A, Gyrd-Hansen D, Olsen AJ, et al (2010). Societal views on orphan drugs: cross sectional survey of Norwegians aged 40 to 67. *BMJ*; 341: c4715.

Devlin N, Parkin P (2004). Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. *Health Economics*; 13: 437-52.

Devlin N, Sussex J (2011). *Incorporating Multiple Criteria in HTA: Methods and Processes*. London: Office of Health Economics.

DHSS (2007). *Managed introduction of new medicines into the NHS in Wales*. Welsh Health Circular (2007) 028; Cardiff: Department of Health and Social Services.

Dolan P (1999). Whose preferences count? *Medical Decision Making*; 19: 482-6.

Dolan P, Shaw R, Tsuchiya A, et al (2005). QALY maximisation and people's preferences: a review of the literature. *Health Economics*; 14: 197-208.

Dolan P, Tsuchiya A (2006). The elicitation of distributional judgements in the context of economic evaluation. In: Jones AM (editor) *The Elgar companion to health economics*. Cheltenham: Edward Elgar Publishing Ltd, UK. pp 382–91.

Dolan P, Edlin R, Tsuchiya A, et al (2008). *The relative societal value of health gains to different beneficiaries: Final report*. Available at: http://eprints.whiterose.ac.uk/10902/1/HEDS_DP_08-12.pdf (accessed 21 March 2013).

Drummond MF (1989). Output measurements for resource allocation decisions in health care. *Oxford Review of Economic Policy*; 5: 59-74.

Drummond MF, O'Brien B, Stoddart GL, et al (1997). *Methods for the economic evaluation of health care programmes,* 2nd Ed; Oxford University Press, Oxford.

Ebell M, Siwek J, Weiss BD, et al (2004). Strength of Recommendation Taxonomy (SORT): a patient-centred approach to grading evidence in the medical literature. American Family Physician; 69: 548-56.

ECHOUTCOME (2013). European Consortium in Healthcare Outcomes and Cost Benefit Research. *Special Conference – Health Outcomes in Europe: The QALY Controversy*; 25th January 2013, Brussels, Belgium. Available at: http://www.echoutcome.eu/index.php/en/publications.html (accessed 13 March 2013)

EEC Directive 65/65/EEC1965. *Council Directive 65/65/EEC of 26 January 1965 on the approximation of provisions laid down by law, regulation or administrative action relating to medicinal products*. Available at: http://www.echamp.eu/fileadmin/user_upload/Regulation/Directive_65-65-EEC_-_Consolidated_Version.pdf (accessed 21 March 2013).

Esplid I, Cairns J, Qvist V, et al (2006). Preferences over dental restorative materials among young patients and dental professionals. *European Journal of Oral Sciences*; 114: 15-21.

European Medicines Agency (2010). *Benefit-risk methodology project*. EMA/549682/2010. London: European Medicines Agency. Available at: http://www.ema.europa.eu/ema/index.jsp?curl=pages/special_topics/document_listing/document_listing_000314.jsp (accessed 21 March 2013).

European Parliament (1999). *Regulation (EC) No. 141/2000*. European Parliament and the Council of the European Union. Available at: http://ec.europa.eu/enterprise/pharmaceuticals/eudralex/vol-1/reg_2000_141/reg_2000_141_en.pdf (accessed 4 May 2010).

Faggioli G, Scalone L, Mantovani LG, et al (2011). Preferences of patients, their family care-givers and vascular surgeons in the choice of aortic aneurysms treatment options: the PREFER study. *European Journal of Vascular Surgery*; 42: 26-34.

Ferner RE, Hughes DA, Aronson JK (2010). NICE and new: appraising innovation. *BMJ*; 340: b5493.

Fiebig DG, Haas M, Hossain I, et al (2009). Decisions about Pap tests: what influences women and providers? *Social Science and Medicine*; 68: 1766-74.

Foot C, Harrison T (2011). *How to improve cancer survival: Explaining England's relatively poor rates*. London: The King's Fund. Available at: http://www.kingsfund.org.uk/publications/cancer_survival.html (accessed 09 May 2012).

Flynn KE, Smith MA, Vanness D (2006). A typology of preferences for participation in healthcare decision making. *Social Science and Medicine*; 63: 1158-69.

Fontain N, Hemila K (2000). *Regulation (EC) No 141/2000 of the European Parliament and of the Council of 16 December 1999 on Orphan Medicinal Products*. Official Journal of the European Communities. Available at: http://ec.europa.eu/health/files/eudralex/vol-1/reg_2000_141/reg_2000_141_en.pdf (accessed 09 May 2012).

Ford JA, Waugh N, Sharma P, et al (2012). NICE guidance: a comparative study of the introduction of the single technology appraisal process and comparison with guidance from Scottish Medicines Consortium. *BMJ Open*; 2 :e000671.

George B, Harris A, Mitchell A (2001). Cost effectiveness analysis and the consistency of decision-making: evidence from pharmaceutical reimbursement in Australia (1991 to 1996). *Pharmacoeconomics*; 19: 1103–9.

Gidman W, Elliot R, Payne K, et al (2007). A comparison of parents and pediatric anaesthesiologists' preferences for attributes of child daycase surgery: a discrete choice experiment. *Pediatric Anaesthesia*; 17: 1043-52.

Greenberg D, Cohen JT, Neumann PJ (2010). Are incremental benefits from new technology decreasing? An analysis of QALY gains over time. *Proceedings of the 13th Biennial Society for Medical Decision Making European Meeting*; Austria [abstract No.1].

Grégoire J-P, McNeil P, Skilton K, et al (2001). Inter-provincial variation in Government drug formularies. *Canadian Journal of Public Health*; 92: 307–13.

Gregorian RS, Gasik A, Kwong WJ, et al (2010). Importance of side effects in opioid treatment: a trade-off analysis with patients and physicians. *Journal of Pain*; 11: 1095-108.

Guttman R, Castle R, Fiebig DG (2009). *Use of discrete choice experiments in health economics: a review of the literature*. CHERE Working Paper 2009/2. Centre for Health Economics Research and Evaluation, University of Technology, Sydney.

Hann GJ, Shapira SS (1996). *A catalogue and computer programme for the design and analysis of orthogonal symmetric and asymmetric fractional factorial experiments*. Report 66-C-165. General Electric Research and Development Centre, Schenectady, New York.

Harris AH, Hill SR, Chin G, et al (2008). The role of value for money in public insiurance coverage decisions for drugs in Australia: a retrospective analysis 1994-2004. *Medical Decision Making*; 28: 713-22.

Hawe E, Yuen P, Baillie L (2011). *OHE guide to health and healthcare statistics*. London: Office of Health Economics. Available at: http://www.ohe.org/publications/article/ohe-guide-to-uk-health-and-health-care-statistics-2.cfm (accessed 17 February 2013).

Haycox A (2008). Does 'NICE blight' exist, and if so, why?. *Pharmacoeconomics*; 26: 987-9.

Hendrix M, Pavlova M, Nieuwenhuijze MJ, et al (2010). Differences in preferences for obstetric care between nulliparae and their partners in the Netherlands: a discrete choice experiment. *Journal of Psychosomatic Obstetrics and Gynecology*; 31:243-51.

Hicks JR (1939). The foundation of welfare economics. *Economic Journal*; 49: 696-712.

Hitchings A, Baker E, Khong T (2012). Making medicines evergreen. *BMJ* ;345:e7941.

House of Commons Health Committee (2005). *The influence of the pharmaceutical industry*. Vol 1. Available at: http://www.publications.parliament.uk/pa/cm200405/cmselect/cmhealth/42/42.pdf (accessed 17 February 2013).

Hughes D, Ferner R (2010). New drugs for old: disinvestment and NICE. *BMJ*; 340: 690-92.

Hughes DA (2011). Value-based pricing: incentive for innovation or zero net benefit? (Editorial) *Pharmacoeconomics*; 29: 731-5.

Hughes D, Duerden M (2011). Ring-fencing a budget for cancer drugs: is it fair? Against. *Journal of the Royal College of Physicians of Edinburgh* 41: 226-28.

Institute for Fiscal Studies/Nuffield Trust (2012). *NHS and social care funding: the outlook to 2021/22*. Available at: http://www.nuffieldtrust.org.uk/sites/files/nuffield/publication/120704_nhs-social-care-funding-outlook-2021-22_0.pdf (accessed 15 July 2012).

Jadad AR, Moore RA, Carroll D, et al (1996). Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Controlled Clinical Trials*; 17: 1-12.

Johnson FR (2006). *Measuring conjoint stated preferences for pharmaceuticals: a brief introduction*. RTI Health Solutions, USA.

Johnson FR, Backhouse M (2006). Eliciting stated preferences for health technology adoption criteria using paired comparisons and recommendation judgements. *Value in Health*; 9: 303-11.

Johnson FR, Hauber B, Ozdemir S, et al (2010). Are gastroenterologists less tolerant of treatment risks than patients? Benefit-risk preferences in Crohns' disease management. *Journal of Managed Care Pharmacy*; 16: 616-26.

Johnson TP, Wislar JS (2012). Response rates and nonresponse errors in surveys. *JAMA* 307: 1805-06.

Kaldor N (1939). Welfare propositions of economic and interpersonal comparisons of utility. *Economic Journal*; 49: 549-52.

Kennedy I (2009). *Appraising the value of innovation and other benefits: a short study for NICE*. Available at: http://www.nice.org.uk/media/98F/5C/KennedyStudyFinalReport.pdf (accessed 17 February 2013).

Kind P, Hardman G, Macran S (1999). *UK population norms for EQ-5D*. Discussion Paper 172, University of York Centre for Health Economics. Available at: http://www.york.ac.uk/media/che/documents/papers/discussionpapers/CHE%20Discussion%20Paper%20172.pdf (accessed 09 May 2012).

Klein R, Maybin J (2012). *Thinking about rationing*. King's Fund. Available at: http://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/Thinking-about-rationing-the-kings-fund-may-2012.pdf (accessed 21 March 2013).

Koopmanschap MA, Stolk EA, Koolman X (2010). Dear policy maker: have you made up your mind? A discrete choice experiment among policy makers and other health professionals. *International Journal of Technology Assessment in Health Care*; 26: 198–204.

Lancaster K (1966). A new approach to consumer theory. *Journal of Political Economy*; 74: 132-57.

Lancsar E, Louviere J (2008). Conducting discrete choice experiments to inform healthcare decision making: a user's guide. *Pharmacoeconomics* 2008; 26: 661-77.

Lee A, Gin T, Lau AS, et al (2005). A comparison of patients' and health care professionsals' preferences for symptoms during immediate postoperative recovery and the management of postoperative nausea and vomiting. *Anasthesia, Analgesie, Reanimation*; 100: 87-93

Lenaghan J (1999). Involving the public in rationing decisions. The experience of citizens juries. *Health Policy*; 49: 45-61.

Lewis SM, Cullinane FN, Carlin JB, et al (2006). Women's and health professionals' preferences for prenatal testing for Down's syndrome in Australia. *The Australian New Zealand Journal of Obstetrics and Gynaecology*; 46: 205-11.

Light DW, Warburton R (2011). Demythologizing the high costs of pharmaceutical research. *B*ioSocieties; 6: 34-50.

Light DW, Lexchin JR (2012). Pharmaceutical research and development: what do we get for all that money? *BMJ*; 345: e4348.

Linley WG, Hughes DA (2010). Ultra-orphan drug status, but not innovativeness, was associated with increased rates of positive recommendations of new medicines by the All Wales Medicines Strategy Group. Poster presentation T03-7 at 7th Health Technology Assessment International Annual Meeting; 6-9th June 2010; Dublin, Ireland.

Longacre VA, Imaeda A, Garcia-Tsao G, et al (2008). A pilot project examining the predicted preferences of patients and physicians in the primary prophylaxis of variceal hemorrhage. *Hepatology*; 47:169-76.

Mantovani LG, Monzinin MS, Mannucci PW, et al (2005). Differences between patients', physicians' and pharmacists' preferences for treatment products in haemophilia: a discrete choice experiment. *Hemophilia*; 11: 589-97.

Marshall DA, Johnson FR, Kullin NA, et al (2009). How do physician assessments of patient preferences for colorectal cancer screening test differ from actual preferences? A comparison in Canada and the United States using a stated-choice survey. *Health Economics*; 18: 1420-39.

Mason A, Drummond M, Ramsel S, et al (2010). Comparison of anticancer drug coverage decisions in the United States and United Kingdom: does the evidence support the rhetoric? *Journal of Clinical Oncology*; 28: 3234-38.

McCabe C, Claxton K, Tsuchiya A (2005). Orphan drugs and the NHS: should we value rarity? *BMJ*; 331: 1016-19.

McCabe C, Claxton K, Culyer AJ (2008). The NICE cost-effectiveness threshold: what it is and what that means. *Pharmacoeconomics*; 26 (9):733-44.

McGuire A. Henderson J, Mooney G (1997). *The economics of health care*. Routledge, London.

McColl E, Jacoby A, Thomas L, et al (2001). Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients. *Health Technology Assessment*; 5 (31): 1-256.

Medicines Act 1968: Elizabeth II. *Chapter 67*, (1968) London: The Stationery Office. Available at: http://www.legislation.gov.uk/ukpga/1968/67/pdfs/ukpga_19680067_en.pdf (accessed 21 March 2013).

Mellor S, Green P (2002). A novel approach to modelling the prescribing decision, integrating physician and patient influences. International *Journal of Market Research*; 44: 449-61.

Menard S. *Applied logistic regression analysis*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-106. Thousand Oaks, CA: Sage.

Mentzakis E, Stefanowska P, Hurley J (2011). A discrete choice experiment investigating preferences for funding drugs used to treat orphan diseases: an exploratory study. *Health Economics, Policy and Law*; 6: 405-33.

Mestre-Ferandiz J, Sussex J, Towse A (2012). *The R&D costs of new medicines*. London: Office of Health Economics.

Mitton C, Smith N, Peacock S, et al (2009). Public participation in health care priority setting: A scoping review. *Health Policy*; 91: 219-28.

Mohamed AF, Johnson FR, Hauber AB, et al (2012). Do patients and physicians have similar preferences for chronic hepatitis B treatment outcomes in Turkey? Flora; 17: 29-38.

Moher D, Liberati A, Tetzlaff J, et al (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine* 2009; 6(7): e1000097.

Montgomery AA, Fahey T (2001). How do patients' treatment preferences compare with those of clinicians? *Quality in Health Care*; 10 (Suppl I):i39-i43.

Mooney G (1989). QALYs: are they enough? A health economist's perspective. *Journal of Medical Ethics*; 15: 148-152.

Mooney G, Ryan M. Agency in health care: getting beyond first principles. *Journal of Health Economics* 1993; 12: 125-135.

Morris S, Devlin N, Parkin D (2007). *Economic analysis in health care*. John Wiley & Sons Ltd, Chichester, UK.

Morton RL, Snelling P, Webster AC, et al (2012). Dialysis modality preference of patients with CKD and family caregivers: a discrete choice experiment. *American Journal of Kidney Disease*; 60:102-11.

Mühlbacher AC, Nübling M (2011). Analysis of physicians' perspectives versus patients' preferences: direct assessment and discrete choice experiments in the therapy of multiple myeloma. *European Journal of Health Economics*; 12:193–203 (and also: Mühlbacher AC, Lincke H-J, Nübling M (2008). Evaluating patients' preferences for multiple myeloma therapy, a Discrete-Choice-Experiment. *Psychosocial Medicine*; 5, Doc10.)

Mulley A, Trimble C, Elwyn G (2012). *Patients' preferences matter: stop the silent misdiagnosis*. London: The Kings Fund. Available at: http://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/patients-preferences-matter-may-2012.pdf (accessed 03 January 2013).

National Prescribing Centre (2008). *What you need to know about prescribing, the 'drugs bill' and medicines management: a guide for NHS managers*. Available at: http://www.npc.nhs.uk/resources/nhs_guide_for_managers.pdf (accessed 17 February 2013).

NCGHSS (2010). National Commissioning Group for Highly Specialised Services / National Specialised Commissioning Group website: http://www.specialisedservices.nhs.uk (accessed 30 April 2010).

Neuman E, Neuman S (2007). *Agency in health-care: are medical care-givers perfect agents?* Forschungsinstitut zur Zukunft der Arbeit (Institute for the Study of Labor) Discussion Paper No. 2727. (and also published as Neuman E, Neuman S (2009). Agency in health-care: are medical care-givers perfect agents? *Applied Economics Letters*; 16: 1355-60).

NHS Information Centre (2011). *Hospital prescribing, England: 2011*. Available at: https://catalogue.ic.nhs.uk/publications/prescribing/primary/hosp-pres-eng-2011/hosp-pres-eng-2011-rep.pdf (accessed 17 February 2013).

NHS Information Centre (2012). *Use of NICE appraised medicines in the NHS in England – 2010 and 2011*, Experimental statistics. Available at: https://catalogue.ic.nhs.uk/publications/prescribing/primary/use-nice-app-med-nhs-exp-stat-eng-10-11/use-nice-app-med-nhs-exp-stat-eng-10-11-rep.pdf (accessed 21 February 2013).

NCGHSS (2011). *National specialised commissioning priorities 2011/12*. Available at: http://www.specialisedservices.nhs.uk/library/21/National_Specialised_Commissioning_Priorities_201112.pdf (accessed 09 May 2012).

NICE (2004). *Guide to the methods of technology appraisal*. London: National Institute for Clinical Excellence. Available at: http://www.nice.org.uk/niceMedia/pdf/TAP_Methods.pdf (accessed 23 February 2013).

NICE (2008a). *Guide to the methods of technology appraisal*. London: National Institute for Health and Clinical Excellence; June 2008. Available at: http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf (accessed 14 January 2013).

NICE (2008b). *Social Value Judgements: principles for the development of NICE guidance*. London: National Institute for Health and Clinical Excellence. Available at: http://www.nice.org.uk/media/C18/30/SVJ2PUBLICATION2008.pdf (accessed 23 February 2013).

NICE (2008c). Report on NICE Citizens Council meeting: *Departing from the threshold*. London: National Institute for Health and Clinical Excellence. Available at: http://www.nice.org.uk/media/231/CB/NICECitizensCouncilDepartingThresholdFinal.pdf (accessed 09 May 2012).

NICE (2009a*). Guide to the single technology appraisal process*. London: National Institute for Health and Clinical Excellence. Available at: http://www.nice.org.uk/media/913/06/Guide_to_the_STA-proof_6-26-10-09.pdf (accessed 25 June 2010).

NICE (2009b). *Supplementary advice to the Appraisal Committees: Appraising life-extending, end of life medicines*. London: National Institute for Health and Clinical Excellence. Available at: http://www.nice.org.uk/media/E4A/79/SupplementaryAdviceTACEoL.pdf (accesed 09 May 2012).

NICE (2009c). *Process for advising on the feasibility of implementing a patient access scheme*. London: National Institute for Health and Clinical Excellence. Available at: http://www.nice.org.uk/media/FCC/CC/PASLUProcessGuide.pdf (accessed 02 March 2013).

NICE (2011). *Clarification to section 5.6 of the Guide to Methods of Technology Appraisals*; July 2011. London: National Institute for Health and Clinical Excellence. Available at: http://www.nice.org.uk/media/955/4F/Clarification_to_section_5.6_of_the_Guide_to_Methods_of_Technology_Appraisals.pdf (accessed 02 March 2013).

NICE (2012a). *Good Practice Guidance: Developing and updating local formularies*. London: National Institute for Health and Clinical Excellence. Available at: http://www.nice.org.uk/media/94A/F8/GPG1Guidance.pdf (accessed 17 January 2013).

NICE (2012b). *Review of the guide to the methods of technology appraisal: draft methods guide*. London: National Institute for Health and Clinical Excellence. Available at: http://www.nice.org.uk/media/CB1/43/GuideToMethodsOfTechnologyAppraisal2012.pdf (accessed 14 January 2013).

NICE (2012c). *Methods for the development of NICE public health guidance* (third edition). London: National Institute for Health and Clinical Excellence. Available at: http://publications.nice.org.uk/methods-for-the-development-of-nice-public-health-guidance-third-edition-pmg4 (accessed 02 March 2013).

NICE (2012d).*Positively equal: a guide to addressing equality issues in developing NICE clinical guidelines* (second edition). London: National Institute for Health and Clinical Excellence. Available at: http://www.nice.org.uk/media/F3B/7B/PositivelyEqual.pdf (accessed 02 March 2013).

NICE (2012e). NICE to assess high cost drugs for rare conditions (media release). London: National Institute for Health and Clinical Excellence. Available at: http://www.nice.org.uk/newsroom/news/NICEToAssessHighCostDrugsForRareConditions.jsp (accessed 21 March 2013).

NRS (2010). *National Readership Survey 2010 lifestyle data*. Available at: http://www.nrs.co.uk/lifestyle.html (accessed 09 May 2012).

Nunes V, Neilson J, O'Flynn N, et al (2009). *Clinical Guidelines and Evidence Review for Medicines Adherence: involving patients in decisions about prescribed medicines and supporting adherence.* London: National Collaborating Centre for Primary Care and Royal College of General Practitioners. Available at: http://www.nice.org.uk/nicemedia/pdf/CG76FullGuideline.pdf (accessed 01 March 2013).

OHE (2007). *Report of the OHE Commission on NHS Outcomes, Performance and Productivity*. London: Office of Health Economics. Available at: http://www.ohe.org/publications/article/nhs-outcomes-performance-and-productivity-42.cfm (accessed 03 January 2013).

OFT (2007). *The Pharmaceutical Price Regulation Scheme: an OFT market study*. London: Office of Fair Trading. Available at: http://www.oft.gov.uk/shared_oft/reports/comp_policy/oft885.pdf (accessed 17 February 2013).

ONS (2009). *General Lifestyle Survey 2009*. London: Office for National Statistics Available at: http://www.ons.gov.uk/ons/rel/ghs/general-lifestyle-survey/2009-report/index.html (accessed 09 May 2012).

ONS (2010). *nomis official labour force statistics: headline indicators 2011/12*. London: Office for National Statistics Available at: http://www.nomisweb.co.uk/tml/templates/data/LFS%20headline%20indicators.xls (accessed 09 May 2012).

ONS (2011). *Mid-2010 population estimates for adults aged 16 years and over*. London: Office for National Statistics Available at: http://www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-uk--england-and-wales--scotland-and-northern-ireland/mid-2010-population-estimates/rft---mid-2010-population-estimates.zip (accessed 09 May 2012).

Opuni M, Bishai D, Gray GE, et al (2010). Preferences for characteristics of antiretroviral therapy provision in Johannesburg, South Africa: results of a conjoint analysis. *AIDS Behaviour*; 14:807-15.

Orme B (2010). *Getting started with conjoint analysis: strategies for product design and pricing research*. (2nd Ed), Madison, Wis.: Research Publishers LLC.

Oudhoff JP, Timmermans DRM, Kmol DL, et al (2007). Prioritising patients on surgical waiting lists: a conjoint analysis study on the priority judgements of patients, surgeons, occupational physicians and general practitioners. *Social Science and Medicine*; 64: 1863-75.

Payne K, Fargher EA, Roberts SA, et al (2011). Valuing pharmacogenetic testing services: a comparison of patients' and health professionals' preferences. *Value in Health*; 14: 121-34.

PausJenssen AM, Singer PA, Detsky AS (2003). Ontario's formulary committee: how recommendations are made. *Pharmacoeconomics*;21:285-94.

Pedersen LB, Kjaer T, Kragstrup J, et al (2012). Do general practitioners know patients' preferences? An empirical study on the agency relationship at an aggregate level using a discrete choice experiment. *Value in Health*; 15: 514-23.

Peng C-YJ, Lee KL, Ingersoll GM. An introduction to logistic regression analysis and reporting. *Journal of Educational Research* 2002; 96: 3-14.

Pieterse AH, Stiggelbout AM, Baas-Thijssen MCM (2007). Benefit from preoperative radiotherapy in rectal cancer treatment: disease-free patients' and oncologists' preferences. *British Journal of Cancer*; 97: 717-24.

Porzsolt F, Clouth J, Deutschmann M, et al (2010). Preferences of diabetes patients and physicians: a feasibility study to identify the key indicators for appraisal of health care values. *Health and Quality of Life Outcomes*; 8: 125.

Puntmann I, Schmake N, Melander A, et al (2010). EVITA: a tool for the early EValuation of pharmaceutical Innovations with regard to Therapeutic Advantage. *BMC Clinical Pharmacology* ;10: 5.

Rawles J (1989). Castigating QALYs. *Journal of Medical Ethics*; 15: 143-7.

Rawlins MD, Culyer AJ (2004). National Institute for Clinical Excellence and its value Judgments. *BMJ*; 329:224-7.

Rawlins M (2005). Pharmacopolitics and deliberative democracy. *Clinical Medicine*; 5: 471-5.

Rawlins M (2008). De Testimono: on the evidence for decisions about the use of therapeutic interventions. *Lancet*; 371: 2152-61.

Rawlins M (2009). Background to NICE's Citizens Council. In: Littlejohns P, Rawlins M (eds). *Patients, the public and priorities in healthcare*. London: Radcliffe Publishing Ltd.

Rawlins M, Barnett D, Stevens A (2010). Pharmacoeconomics: NICE's approach to decision-making. *British Journal of Clinical Pharmacology* 70: 346-9.

Relman AS, Angell M (2002). America's other drug problem: how the drug industry distorts medicine and politics. *The New Republic*, December 16: 27-41.

Richards M (2010). *Extent and causes of international variations in drug usage. A report for the Secretary of State for Health by Professor Sir Mike Richards CBE*. Available at: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/@ps/documents/digitalasset/dh_117977.pdf (accessed 17 February 2013).

Richards M (2008). *Improving access to medicines for NHS patients: a report for the Secretary of State for Health by Professor Sir Mike Richards CBE*. Available at: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_089952.pdf (accessed 09 May 2012).

Roberts F (1948). Where are we going? Medicine in a planned economy. *BMJ*; 1: 485-9.

Rowe G, Frewer LJ (2005). A typology of public engagement mechanisms. *Science, Technology and Human Values*; 30: 251-90.

Ryan M, Scott DA, Reeves C, et al (2001). Eliciting public preferences for healthcare: a systematic review. *Health Technology Assessment*; 5: No. 5. Available at: http://www.hta.ac.uk/fullmono/mon505.pdf (accessed 12 December 2012).

Ryan M, Gerard K (2003). Using discrete choice experiments to value health care programmes: current practice and future research reflections. *Applied Health Economics and Health Policy*; 2: 55-64.

Ryan M, Gerard K, Amaya-Amaya M, editors (2008). *Using discrete choice experiments to value health and health care*. Dordrecht, The Netherlands: Springer; 2008.

Sabik LM, Lie RK (2008). Principles versus procedures in making health care coverage decisions: addressing inevitable conflicts. *Theoretical Medicine and Bioethics*; 29:73-85.

Sampietro-Colom L, Espallargues M, Rodríguez E, et al (2008). Wide social participation in prioritizing patients on waiting lists for joint replacement: a conjoint analysis. *Medical Decision Making*; 28: 554-66.

Sassi F, Archard L, Le Grand J (2001). Equity and the economic evaluation of healthcare. *Health Technology Assessment*; 5: 1-138.

Scalone L, Mantovani LG, Borghetti F, et al (2009). Patients', physicians', and pharmacists' preferences towards coagulation factor concentrates to treat haemophilia with inhibitors: results from the COHIBA Study. *Haemophilia*; 15: 473-86.

Schlander M (2008). The use of cost-effectiveness by the National Institute for Health and Clinical Excellence (NICE): no(t yet an) exemplar of a deliberative process. *Journal of Medical Ethics*; 34: 534-9.

Schmitz JM, Rhoades H, Grabowski J (1994). A menu of potential reinforcers in a methadone maintenance program. *Journal of Substance Abuse Treatment*; 11: 425-31.

Scott A, Vick S. Patients, doctors and contracts: an application of principal-agent theory to the doctor-patient relationship. *Scottish Journal of Political Economy* 1999; 46: 111-34.

Sen A (1985). *Commodities and capabilities*. Oxford: Elsevier Science Publishers.

ShafeyM, Lupichuck SM, Do T, et al (2011). Preferences of patients and physicians concerning treatment options for relapsed follicular lymphoma: a discrete choice experiment. *Bone Marrow Transplantation*; 46: 962-9.

Shah KK (2009). Severity of illness and priority setting in healthcare: a review of the literature. *Health Policy* 93: 77-84.

Shah, KK, Cookson R, Culyer AJ, et al (2011a). *NICE's social value judgements about equity in health and health care*. Centre for Health Economics Research Paper 70. York: University of York. Available at: http://www.york.ac.uk/media/che/documents/papers/researchpapers/CHERP70_Nice's_social_value_judgements_about_equity_in_health.pdf (accessed 21 March 2013).

Shah KK, Praet C, Devlin NJ, et al (2011b). *Is the aim of the health care system to maximise QALYs? An investigation of what else matters in the NHS*. Research Paper 11/03. London: Office of Health Economics. Available at: http://www.ohe.org/publications/article/is-the-aim-of-the-health-care-system-to-maximise-qalys-4.cfm (accessed 03 January 2013).

Shah K, Tsuchiya A Risa Hole A et al (2012). *Valuing health at the end of life: a stated preference discrete choice experiment*. NICE Decision Support Unit Report. Available at: http://www.nicedsu.org.uk/DSU%20End%20of%20Life%20full%20report%20-%20version%203%20_Dec%202012_.pdf (accessed 21 March 2013).

SMC (2010). *SMC modifiers used in appraising new medicines*. Glasgow: Scottish Medicines Consortium. Available at: http://www.scottishmedicines.org.uk/About_SMC/Policy_Statements/SMC_Modifiers_used_in_Appraising_New_Medicines (accessed 09 May 2012).

SMC (2012). *Working with SMC – A Guide for Manufacturers*. Glasgow: Scottish Medicines Consortium. Available at: http://www.scottishmedicines.org/Submission_Process/Submission_Guidance_and_Templates_for_Industry/Templates-Guidance-for-Submission/Templates-Guidance-for-Submission (accessed 21 March 2013).

Soinin EJ, Kukkonen J, Myllykangas M, et al (2012). Contingent valuation of eight new treatments: what is the clinician's and politician's willingness to pay? *The Open Complementary Medicine Journal*; 4: 1-11.

Stacey D, Bennett CL, Barry MJ, et al. (2012). *Decision aids for people facing health treatment or screening decisions* (Review). The Cochrane Library, Issue 5. Available at: http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD001431.pub3/pdf (accessed 01 March 2013).

Stafinski T, Menon D, Philippon DJ, et al (2011a). Health technology funding decision-making processes around the world: the same, yet different. *Pharmacoeconomics*; 29: 475-95.

Stafinski T, Menon D, Davis C, et al (2011b). Role of centralised review processes for making reimbursement decisions on new health technologies in Europe. *ClinicoEconomics and Outcomes Research*; 3: 117-86.

Stanek EJ, Oates MB, McGhan WF, et al (2000). Preferences for treatment outcomes in patients with heart failure: symptoms versus survival. *Journal of Cardiac Failure*; 6 : 225-32.

Stiggelbout AM and de Vogel-Voogt E (2008). Health state utilities: a framework for studying the gap between the imagined and the real. *Value in Health* 11: 76-87.

Stinnett AA, Mullahy J (1998). Net health benefits: a new framework for the analysis of uncertainty in cost effectiveness analysis. *Medical Decision Making*; 19: S68-80.

Sung L, Alibhai SM, Ethier M-C, et al (2012). Discrete choice experiment produced estimates of acceptable risks of therapeutic options in cancer patients with febrile neutropenia. *Journal of Clinical Epidemiology*; 65: 627-34.

Sussex J, Towse A, Devlin N (2011). *Operationalising value-based pricing of medicines: a taxonomy of approaches*. Research Paper 11/04. London: Office of Health Economics. Available at: http://www.ohe.org/publications/article/operationalising-value-based-pricing-of-medicines-1.cfm (accessed 03 January 2013).

Taylor RS, Elston J (2009). The use of surrogate outcomes in model-based cost-effectiveness analyses: a survey of UK Health Technology Assessment reports. *Health Technology Assessment*; 13 (8).

Tappenden P, Brazier J, Ratcliffe J, et al (2007). A stated preference binary choice experiment to explore NICE decision making. *Pharmacoeconomics*; 25: 685-93.

Thakula P. *Multicriteria decision analysis for health technology assessment*. NICE Decision Support Unit; School of Health and Related Research, University of Sheffield, UK. Available at: http://www.nicedsu.org.uk/MCDA%20for%20HTA%20DSU.pdf (accessed 17 February 2013).

Thrumurthy SG, Morris JJA, Mughal MM, et al (2011). Discrete-choice preference comparison between patients and doctors for the surgical management of oesophagogastric cancer. *British Journal of Surgery*; 98: 1124-31.

Tilson L, O'Leary A, Usher C, et al (2010). Pharmacoeconomic evaluation in Ireland: a review of the process. *Pharmacoeconomics*; 28: 307-22.

Tsuchiya A, Williams A (2001). Welfare economics and economic evaluation. In: McGuire A, Drummond MF (eds). *Economic Evaluation in Health Care: Merging Theory with Practice*; Oxford: Oxford University Press.

van Empel IWH, Dancet EAF, Koolman XHE, et al (2011). Physicians underestimate the importance of patient-centredness to patients: a discrete choice experiment in fertility care. *Human Reproduction*; 26: 584-93.

Vermeulen H, Ubbink DT, de Zwart F, et al (2007). Preferences of patients, doctors, and nurses regarding wound dressing characteristics: a conjoint analysis. *Wound Repair and Regeneration*; 15: 302-7.

walesonline (2013).Critics of proposed cancer drugs fund for Wales 'ignorant or lying'; 15 February 2013. Available at: http://www.walesonline.co.uk/news/health-news/2013/02/15/opponents-of-cancer-drugs-fund-for-wales-lying-or-ignorant-91466-32816628/ (accessed 21 March 2013).

Walker A, Booth C, Brown A, et al (2009). How much good do new medicines do? Basic *Clinical Pharmacology and Toxicology*;105 (Suppl 1): O29.

Watson V, Ryan M, Watson E (2009). Valuing experience factors in the provision of Chlamydia screening: an application to women attending the family planning clinic. *Value in Health*; 12: 621-3.

Weale A (1998). Rationing health care [Editorial]. *BMJ*; 316: 410.

Weinstein MC, Stason WB (1977). Foundations of cost-effectiveness analysis for health and medical practices. *New England Journal of Medicine*; 296:716-21.

Webb DJ (2011). Value-based pricing: NICE work? *Lancet* 377: 1552-3.

Wellman G, Pathak D, Shommer J (2003). Drug attributes and patient out-of-pocket cost impact on preference: conjoint analysis of physicians, pharmacists, and consumers. *Journal of Pharmaceutical Marketing & Management*; 15: 47-61.

Whitehead M (1994). Equity issues in the NHS: who cares about equity in the NHS? *BMJ* 308: 1284-7.

Whitty JA, Scuffham PA, Rundle-Thiele SR (2011). Public and decision maker stated preferences for pharmaceutical subsidy: a pilot study. *Applied Health Economics & Health Policy*; 9: 73-9.

Whitty JA. (2013). An international survey of the public engagement practices of Health Technology Assessment organizations. *Value in Health*; 16: 155-63.

WHO (2008). *The global burden of disease: 2004 update*. Geneva: World Health Organisation. Available from URL: http://www.who.int/healthinfo/global_burden_disease/2004_report_update/en/index.html (accessed 30 April 2010).

Willink H (1944). White Paper: *A National Health Service*. Cmd 6501; February 1944.

Youngkong S, Baltussen R, Tantivess S, et al (2010). Criteria for priority setting of HIV/AIDS interventions in Thailand: a discrete choice experiment. *BMC Health Services Research*; 10: 197.

Appendix to Chapter 1

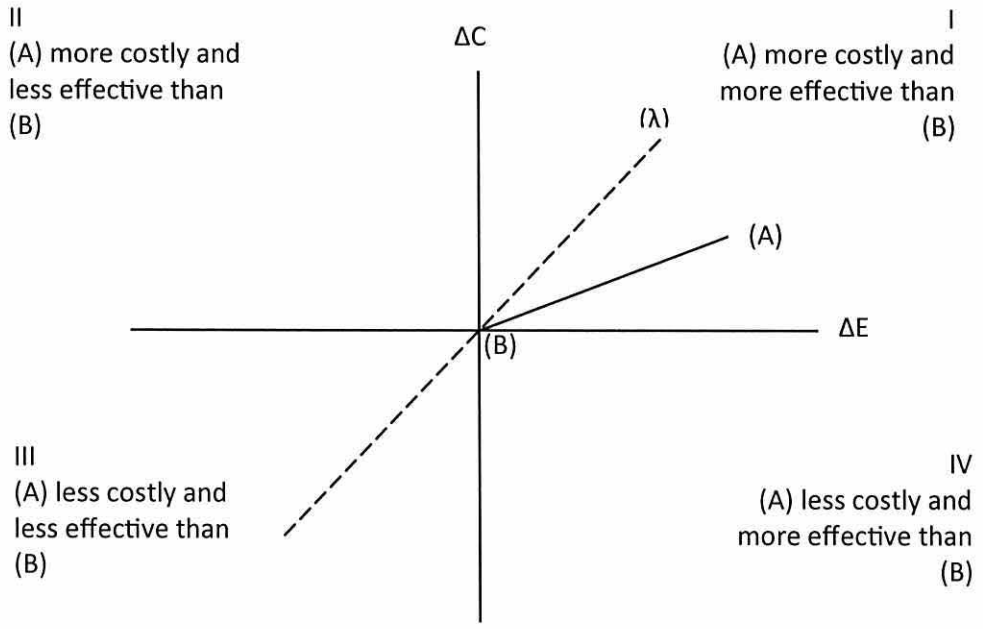
Decision rules in cost effectiveness analysis

If a decision-maker wished to determine the cost effectiveness of a new medicine, medicine A, relative to the next best alternative use of the available health resources, medicine B, it would be necessary to determine all relevant costs associated with the use of A (C_A) and B (C_B) and also the benefits of A (E_A) and B (E_B). The additional (incremental) cost of providing an additional (incremental) unit of effectiveness is termed the incremental cost effectiveness ratio (ICER) and is given by:

$$ICER = \frac{C_A - C_B}{E_A - E_B} = \frac{\Delta C}{\Delta E}$$

A plot of the incremental costs and incremental benefits is called the cost effectiveness plane (Black 1990), as shown in Figure 1 below (adapted from Drummond et al., 1997; Morris et al., 2007).

Figure 1.1A. Cost effectiveness plane



If medicine A is less costly and more effective than medicine B (quadrant IV of the cost effectiveness plane) then medicine A is always viewed to be an acceptable use of health care resources compared to medicine B. If medicine A is more costly and less effective than medicine B (quadrant II of the cost effectiveness plane) then medicine A is never an

acceptable use of health care resources compared with medicine B. However, when medicine A is both more costly and more effective than B (quadrant I), or less costly and less effective than B (quadrant III), then a threshold value of the ICER (λ), representing the decision-maker's maximum willingness to pay for an additional unit of effectiveness, is required to determine whether medicine A would be a cost effective use of resources. If the ICER for medicine A compared to medicine B is less than this threshold value, then medicine A would be considered to be cost effective, or if the ICER is greater, then medicine A would not be cost effective compared to medicine B (Morris et al., 2007).

This approach is relatively intuitive when considering the ICER as a single point estimate; however, the ICER actually has a range of statistically plausible values that are determined by the error terms and distribution of the various parameters that are used in its estimation. Consequently, this range of values of the ICER may be contained within one quadrant of the cost effectiveness plane, but straddle the threshold for cost effectiveness (λ), or it may straddle several quadrants. It would be useful, therefore, to understand the range of values within which the ICER could fall, such as the confidence interval around the point estimate; however, as ICERs are a ratio (of costs:effectiveness), and ratios do not have a simple probability distribution, estimation of the confidence interval can be problematic.

One approach to dealing with this problem is to convert the ICER into a single monetary value (or single health benefit value) similar to that obtained in a CBA, using the decision-maker's threshold value for cost effectiveness (i.e. their willingness to pay for an additional unit of effectiveness). This single monetary value is called the incremental net monetary benefit (NMB) (or the single health benefit value is called the incremental net health benefit, NHB) (Stinnett and Mullahy, 1998) and is estimated as:

$$NMB = \lambda \times \Delta E - \Delta C$$

$$NHB = \Delta E - \frac{\Delta C}{\lambda}$$

A positive NMB or NHB would indicate that the new medicine is cost effective (i.e. delivers positive new monetary or net health benefits).

Appendix to Chapter 2

Table 2.1A PRISMA Checklist (Moher et al., 2009) for the systematic review of health-related conjoint analyses

Section	Item #	Checklist item	Reported on page #	Notes (if applicable)
Title				
Title	1	Identify the report as a systematic review, meta-analysis, or both.	34	Title of Chapter 2: Systematic review of health-related conjoint analyses
Abstract				
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	35-36	Structured abstract provided. Systematic review not registered. No systematic review registration number.
Introduction				
Rationale	3	Describe the rationale for the review in the context of what is already known.	37-38	-
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	38	As not a review of clinical data, PICOS less clear, but objectives clearly stated and elements of PICOS are included: P: Two or more respondent groups who may reasonably be viewed to be involved in a principle-agent type relationship. I: Conjoint analyses of any health-related intervention, service, or technology included. Preferences of principals and agents towards any health-related intervention, service, or technology included. C: Comparisons are the preferences of principals and agents towards any health-related intervention, service, or technology included. O: Relative importance of different attribute types. S: Conjoint analyses (including DCEs, ACA, ranking and rating conjoint analyses)
Methods				
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	38	Systematic review not registered. No prior published protocol.
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	39	Only conjoint analyses considered. PICOS framework difficult to apply. See item #4.
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	38-39	-

Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	38-39	-
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	39	Studies for inclusion screened independently by both authors.
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	40	Proforma developed for data extraction. Data extracted by first author. (Further independent verification of data would be beneficial, as outlined in the Discussion of Limitations section)
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	40	Proforma for data extraction. Relative importance of attributes obtained or estimated. Attributes assigned to one of five categories.
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	39	First author compared studies against ISPOR good practice checklist for conjoint analyses.
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	40	Relative importance of attributes as reported or as inferred from marginal rates of substitution or contribution to overall utility estimates.
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency ($e.g., I^2$) for each meta-analysis.	40	Quantitative synthesis of data not feasible due to differences in scale and error variance. Descriptive statistics employed to explore implied rank order of importance of different attribute types.
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	n/a	Not specifically reported. Systematic literature search methods employed. Given the wide range of health care interventions and services eligible for inclusion, risk of publication bias not assessable. Selective reporting within studies unlikely given the nature of the empirical exercise.
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	n/a	No additional analyses planned or conducted, beyond the different principal-agent pairings.
Results				
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	41	Flow diagram provided (Figure 2.1).
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	53-59	Table of included studies provided (Table 2.3)
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	40, 203-206	Assessment of conjoint analysis studies against ISPOR good practice checklist (Table 2.2A)
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	53-59	Table of included studies (Table 2.3, pages 53-59) provides rank order of importance of attributes for each study.

Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	45-47	Not meta-analysis – descriptive stats.
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	n/a	-
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	45-47	-
Discussion				
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	47-49	General findings and implications for different stakeholders discussed.
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	50-52	-
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	50-52	-
Funding				
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	60	-

Table 2.2A ISPOR checklist for good practice in conducting conjoint analyses (Bridges et al., 2011)

Checklist item	Considerations
#1. Was a well-defined research question stated and is conjoint analysis an appropriate method for answering it?	Were a well-defined research question and a testable hypothesis articulated? Was the study perspective described, and was the study placed in a particular decision-making or policy context? What is the rationale for using conjoint analysis to answer the research question?
#2. Was the choice of attributes and levels supported by evidence?	Was attribute identification supported by evidence (literature reviews, focus groups, or other scientific methods)? Was attribute selection justified and consistent with theory? Was level selection for each attribute justified by the evidence and consistent with the study perspective and hypothesis?
#3. Was the construction of tasks appropriate?	Was the number of attributes in each conjoint task justified (that is, full or partial profile)? Was the number of profiles in each conjoint task justified? Was (should) an opt-out or a status-quo alternative (be) included?
#4. Was the choice of experimental design justified and evaluated?	Was the choice of experimental design justified? Were alternative experimental designs considered? Were the properties of the experimental design evaluated? Was the number of conjoint tasks included in the data-collection instrument appropriate?
#5. Were preferences elicited appropriately, given the research question?	Was there sufficient motivation and explanation of conjoint tasks? Was an appropriate elicitation format (that is, rating, ranking, or choice) used? Did (should) the elicitation format allow for indifference? In addition to preference elicitation, did the conjoint tasks include other qualifying questions (for example, strength of preference, confidence in response, and other methods)?
#6. Was the data collection instrument designed appropriately?	Was appropriate respondent information collected (such as sociodemographic, attitudinal, health history or status, and treatment experience)? Were the attributes and levels defined, and was any contextual information provided? Was the level of burden of the data-collection instrument appropriate? Were respondents encouraged and motivated?
#7. Was the data-collection plan appropriate?	Was the sampling strategy justified (for example, sample size, stratification, and recruitment)? Was the mode of administration justified and appropriate (for example, face-to-face, pen-and-paper, web-based)? Were ethical considerations addressed (for example, recruitment, information and/or consent, compensation)?
#8. Were statistical analyses and model estimations appropriate?	Were respondent characteristics examined and tested? Was the quality of the responses examined (for example, rationality, validity, reliability)? Was model estimation conducted appropriately? Were issues of clustering and subgroups handled appropriately?

#9. Were the results and conclusions valid?	Did study results reflect testable hypotheses and account for statistical uncertainty? Were study conclusions supported by the evidence and compared with existing findings in the literature? Were study limitations and generalizability adequately discussed?
#10. Was the study presentation clear, concise, and complete?	Was study importance and research context adequately motivated? Were the study data-collection instrument and methods described? Were the study implications clearly stated and understandable to a wide audience?

Table 2.3A Assessment of conjoint analysis studies against the ISPOR good practice checklist

Study Ref. Author, Year	#1	# 2	# 3	#4	#5	#6	#7	#8	#9	#10
1. Arana, 2006	Y	Y	Y	U	Y	Y	Y	Y	Y	Y
2. Bederman, 2010	Y	Y	Y	Y	U	U	U	Y	Y	Y
3. Bijlenga, 2011	Y	Y	U	N	U	Y	Y	Y	Y	Y
4. Bishop, 2004	Y	Y	Y	U	Y	Y	Y	Y	Y	Y
5. Bridges, 2012	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
6. Cheung, 2012	Y	Y	U	Y	U	U	U	Y	Y	Y
7. Davison, 2010	Y	U	U	Y	Y	U	Y	U	U	U
8. De Bekker-Grob, 2009	Y	Y	Y	Y	U	Y	Y	Y	Y	Y
9. Espelid, 2006	Y	Y	U	Y	U	Y	Y	U	U	Y
10. Faggioli, 2011	Y	Y	N	U	U	Y	Y	U	Y	Y
11. Fiebig, 2009	Y	Y	Y	Y	U	Y	Y	U	Y	Y
12. Gregorian, 2010	Y	Y	Y	Y	Y	U	Y	Y	Y	Y
13. Hendrix, 2010	Y	Y	Y	U	U	Y	Y	U	Y	Y
14. Johnson, 2010	Y	Y	U	Y	U	Y	Y	Y	Y	Y
15. Lee, 2005	Y	N	Y	U	U	Y	Y	U	Y	Y
16. Lewis, 2006	Y	Y	Y	Y	U	Y	Y	Y	Y	Y
17. Longacre, 2008	Y	U	U	U	Y	U	Y	U	Y	Y
18. Mantovani, 2005	Y	Y	Y	Y	U	Y	Y	Y	U	Y
19. Marshall, 2009	Y	U	Y	Y	U	Y	Y	Y	Y	Y
20. Mellor, 2002	Y	N	N	N	N	N	N	U	N	U
21. Mohamed, 2012	Y	Y	Y	Y	U	Y	Y	U	Y	Y
22. Morton, 2012	Y	Y	Y	U	Y	Y	Y	Y	Y	Y
23. Mühlbacher, 2011	Y	Y	Y	Y	Y	U	Y	Y	U	Y
24. Neuman, 2007	Y	Y	Y	U	U	U	U	U	Y	Y
25. Opuni, 2010	Y	Y	Y	U	Y	Y	Y	U	Y	Y
26. Oudhoff, 2007	Y	Y	Y	U	U	U	U	Y	Y	Y

Study Ref. Author, Year	#1	# 2	# 3	#4	#5	#6	#7	#8	#9	#10
27. Payne, 2011	Y	Y	Y	Y	U	U	Y	Y	Y	Y
28. Pedersen, 2012	Y	Y	Y	U	Y	Y	Y	U	Y	Y
29. Pieterse, 2007	Y	U	N	U	U	U	U	Y	Y	Y
30. Porzolt, 2010	Y	Y	U	U	U	U	U	U	Y	Y
31. Sampietro-Colom, 2008	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
32. Scalone, 2009	Y	Y	Y	Y	U	Y	Y	Y	Y	Y
33. Schmitz ,1994	Y	U	U	U	U	U	U	U	U	Y
34. Shafey, 2011	Y	Y	Y	Y	U	U	Y	Y	Y	Y
35. Stenek, 2000	Y	U	Y	Y	Y	Y	Y	Y	Y	Y
36. Sung, 2012	Y	Y	U	Y	U	Y	Y	U	Y	Y
37. Thrumurthy, 2011	Y	Y	U	U	Y	Y	Y	Y	Y	Y
38. van Empel, 2011	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
39. Vermeulen; 2007	Y	Y	Y	U	U	U	Y	Y	Y	Y
40. Wellman, 2003	Y	Y	Y	U	Y	U	Y	U	Y	Y
41. Youngkong, 2010	Y	Y	Y	U	Y	Y	Y	Y	Y	Y
42. Bech, 2003	Y	Y	Y	U	Y	U	Y	Y	Y	Y
43. Gidman, 2007	Y	Y	Y	U	U	U	Y	U	Y	Y
44. Soinin, 2012	Y	U	N	U	U	U	Y	U	U	U
45. Whitty, 2011	Y	Y	Y	Y	U	U	Y	U	Y	Y

Y=Yes – adequately reported; N= No – inadequately reported; U=Some elements considered but unclear all relevant factors considered

Appendix to Chapter 3

Table 3.1ASubmissions and recommendation of All Wales Medicines Strategy Group, 2007-2009 (n=60)

Submission for medicine	Appraised indication[a]	Recommendation[b]	
		Preliminary NMG	Final AWMSG
Ranolazine (Ranexa®)	Stable angina pectoris	Not recommended	Not recommended
Darunavir (Prezista®)	HIV-1 infection in treatment-naive patients	Routine	Routine
Degarelix (Firmagon®)	Advanced hormone-dependent prostate cancer	Not recommended	Not recommended
Quetiapine (Seroquel XL®)	Schizophrenia in adults	Routine	Routine
Filgrastim (Ratiograstim®)	Treatment of neutropenia and mobilisation of peripheral blood progenitor cells	Routine	Routine
Mecasermin (Increlex®)[c,d]	Growth failure in children and adolescents with insulin-like growth factor-I deficiency	Routine	Routine
Methoxy polyethylene glycol-epoetin beta (Mircera®)	Symptomatic anaemia associated with chronic kidney disease	Routine	Routine
Paricalcitol (Zemplar®)	Secondary hyperparathyroidism associated with chronic renal insufficiency	Not recommended	Not recommended
Ropinerol (Requip XL®)	Ideopathic parkinsons disease in patients already taking immediate-release ropinirole	Not recommended	Routine
Etravirine (Intelence®)	HIV-1 infection in treatment-experienced patients	Not recommended	Routine
Maraviroc (Celsentri®)	HIV-1 infection in treatment-experienced adults with CCR-5-tropic HIV-1	Routine	Routine
Anidulafungin (Ecalta®)	Invasive candidiasis in non-neutropenic adults	Routine	Routine
Nelarabine (Atriance®)[c,d]	T-cell lymphoblastic leukaemia and lymphoma	Restricted	Restricted
Thalidomide (Thalidomide Pharmion®)[c]	First-line treatment of multiple myeloma	Routine	Routine
Efavirenz / emtricitabine / tenofovir disoproxil (as fumarate) (Atripla®)	HIV-1 infection in adults with virologic suppression on current combination antiretroviral therapy for more than three months	Routine	Routine
Micafungin (Micamine®)	Invasive candidiasis	Not recommended	Not recommended
Bivalirudin (Angiox®)	Acute coronary syndromes planned for urgent or early intervention	Restricted	Restricted
Aripiprazole (Abilify®)		Not recommended	Routine
Tacrolimus prolonged-release (Advagraf®)[e]	Prevention of kidney and liver transplant rejection	Not recommended	Not recommended
Aliskiren (Rasilez®)	Essential hypertension	Not recommended	Not recommended
Eculizumab (Soliris®)[c,d]	Paroxysmal nocturnal haemoglobinuria	Restricted	Restricted
Alemtuzumab (MabCampath®)	B-cell chronic lymphocytic leukaemia when fludarabine combination chemotherapy is not appropriate	Restricted	Restricted
Ambrisentan (Volibris®)[c]	Pulmonary arterial hypertension	Restricted	Restricted
Atazanavir (Reyataz®)	Treatment-experienced HIV-1 infected adults	Routine	Routine

Submission for medicine	Appraised indication[a]	Recommendation[b]	
		Preliminary NMG	Final AWMSG
Atazanavir (Reyataz®)	Treatment-naive HIV-1 infected adults	Routine	Routine
Stiripentol (Diacomit®)[c]	Severe myoclonic epilepsy in infancy in conjunction with clobazam and valproate	Not recommended	Not recommended
Rufinamide (Inovelon®)[c]	Seizures associated with Lennox-Gastaut syndrome	Routine	Restricted
Raltegravir (Isentress®)	Treatment-experienced HIV-1 infected adults with evidence of HIV-1 replication despite ongoing antiretroviral treatment	Routine	Routine
Icatibant (Firazyr®)[c]	Acute attacks of hereditary angioedema in adults (with C1-esterase-inhibitor deficiency)	Not recommended	Not recommended
Abacavir/Lamivudine (Kivexa®)	Treatment-naive HIV-1 infected adults and adolescents	Routine	Routine
Teriparatide (Forsteo®)	Osteoporosis in men	Not recommended	Not recommended
Pegfilgrastim (Neulasta®)	Reduction in duration of neutropenia and febrile neutropenia in chemotherapy recipients	Restricted	Restricted
Trabectedin (Yondelis®)[c,d]	Advance soft tissue sarcoma after failure of anthracyclines and ifosfamide	Not recommended	Not recommended
Tenofovir DF (Viread®)	Chronic hepatitis B	Routine	Routine
Deferasirox (Exjade®)[c]	Chronic iron overload due to blood transfusions	Routine	Routine
Buprenorphine/Naloxone (Suboxone®)	Substitution treatment for opioid dependency	Restricted	Restricted
Lenolidomide (Revlimid®)[c]	Multiple myeloma	Not recommended	Not recommended
Docetaxel (Taxotere®)	Locally advanced squamous cell carcinoma of the head and neck	Restricted	Restricted
Ziconotide (Prialt®)[c]	Intrathecal treatment of severe chronic pain	Not recommended	Not recommended
Fondaparinux (Arixtra®)	Unstable angina or non- ST-segment elevation myocardial infarction	Routine	Routine
Fondaparinux (Arixtra®)	ST-segment elevation myocardial infarction	Routine	Routine
Epoetin Delta (Dynepo®)	Anaemia in chronic renal failure	Routine	Routine
Tacrolimus prolonged-release (Advagraf®)	Prevention of kidney and liver transplant rejection	Not recommended	Not recommended
Topotecan (Hycamtin®)	Carcinoma of the cervix recurrent after radiotherapy and for patients with Stage IVB disease, in combination with cisplatin	Restricted	Restricted
Topotecan (Hycamtin®)	Relapsed small cell lung cancer	Not recommended	Routine
Dasatinib (Sprycel®)[c]	Chronic phase chronic myeloid leukaemia and accelerated phase where there is resistance or intolerance to prior therapy	Not recommended	Restricted
Dasatinib (Sprycel®)[c,d]	Philadelphia chromosome positive acute lymphoblastic leukaemia and lymphoid blast chronic myeloid leukaemia with resistance or intolerance to prior therapy.	Restricted	Not recommended
Idursulfase (Elaprase®)[c,d]	Long-term treatment of patients with Hunter syndrome (Mucopolysaccharidosis II)	Restricted	Not recommended
Agalsidase alfa (Replagal®)[c,d]	Long-term enzyme replacement therapy in patients with confirmed diagnosis of Fabry disease	Restricted	Routine

Submission for medicine	Appraised indication[a]	Recommendation[b]	
		Preliminary NMG	Final AWMSG
Vinorelbine (Navelbine Oral®)	Relapsing advanced breast cancer stage III and IV following anthracycline regimen	Routine	Routine
Co-carledopa intestinal gel (Duodopa®)[c]	Advanced levodopa-responsive Parkinson's disease	Not recommended	Not recommended
Darunavir (Prezista®)	Human immunodeficiency virus (HIV-1) infection in highly pre-treated adults who have failed more than one regimen containing a protease inhibitor	Routine	Routine
Sunitinib (Sutent®)	Advanced and/or metastatic renal cell carcinoma	Not recommended	Not recommended
Tipranavir (Aptivus®)	Human immunodeficiency virus (HIV-1) infection, only for the treatment of highly pre-treated adult patients who have failed multiple protease inhibitors	Routine	Routine
Clofarabine (Evoltra®)[c,d]	Acute lymphoblastic leukaemic in patients aged ≤ 21 yrs who have relapsed or are refractory after receiving at least two prior regimens and where there is no other treatment option anticipated to result in a durable response	Restricted	Restricted
Dexrazoxane (Savene®)[c]	Anthracycline extravasation	Not recommended	Not recommended
Emtricitabine (Emtriva®)	Human Immunodeficiency Virus (HIV-1) infected treatment-naïve adults	Routine	Routine
Emtricitabine/Tenofovir (Truvada®)	Human Immunodeficiency Virus (HIV-1) infected treatment-naïve adults	Routine	Routine
Parathyroid hormone (Preotact®)	Postmenopausal osteoporosis	Not recommended	Restricted
Sorafenib (Nexavar®)[c]	Advanced renal cell carcinoma in patients who have failed prior interferon-α or interleukin-2 based therapy or are considered unsuitable for such therapy	Not recommended	Not recommended

AWMSG = All Wales Medicines Strategy Group; NMG = New Medicines Group; [a] Summary of appraised indication – see AWMSG website for full details (http://www.wales.nhs.uk/awmsg); [b] Only the final AWMSG recommendation is relevant to NHS Wales; [c] EU Orphan drug status; [d] Ultra-orphan status as defined by AWMSG; [e] Resubmission for same indication

Appendix to Chapter 4

Instructions and example choice task provided to AWMSG committee members completing the discrete choice experiment questionnaire

Introduction

The aim of this questionnaire is to help understand the factors that influence the recommendations of the All Wales Medicines Strategy Group (AWMSG) and its subgroup, the New Medicines Group (NMG), in relation to new medicines in Wales. The results of this questionnaire will contribute to a Bangor University funded PhD project that is being undertaken by Warren Linley (supervised by Prof. Dyfrig Hughes), who also provides health economic support to Welsh Medicines Partnership (WMP) in relation to the AWMSG new medicines appraisal programme.

You have been selected to participate as an existing or past voting member of either the AWMSG or the NMG. The entire questionnaire will take around 30-35 minutes to complete.

Thank you for taking the time to complete this questionnaire. Your support is critical to the success of this project.

Instructions

The questionnaire consists of two parts:

Part 1 presents you with a discrete choice experiment, which is a specific method of estimating your preferences for specific attributes of new medicines and characteristics of the conditions they aim to treat. You are presented with 28 choices between hypothetical "New Medicine A" and "New Medicine B". You are asked to respond to each and every choice. Detailed instructions are provided.

Part 2 explores your opinions on a range of potential factors that may or may not be of relevance to your decisionmaking.
You are requested to rate the importance of these factors, and are provided with an opportunity to list any other factors that you consider are important. You are also asked to state your level of agreement with a range of statements exploring priorities for NHS resources.

There are no right or wrong answers to any of the questions or scenarios that are posed; we are interested only in your own opinion and views when acting in your capacity as a

voting member of either AWMSG or NMG. All responses are anonymous. The only respondent-specific information we will collect is your status as a voting member of AWMSG or NMG.

Part 1 Discrete choice experiment Instructions

You are presented with a choice between recommending hypothetical New Medicine A or hypothetical New Medicine B for use in NHS Wales, based on each of their stated profiles. Note that New Medicine A and New Medicine B are indicated for the treatment of different diseases or conditions. Consequently, the main health impact of the disease and annual number of patients to be treated differ between the two new medicines profiles.

Also note that New Medicine A and New Medicine B are each compared against hypothetical current standards of care for their licensed indications, which may include treatment with older, established medicine(s), or may involve best supportive care where no other treatment options are currently available.

The profile of each hypothetical new medicine consists of the following five attributes and characteristics:

i) Main disease impact *before* treatment:

The main health burden due to the underlying disease or condition (before treatment with either new medicine) may be to reduce survival (Survival) or may be to reduce health-related quality of life (Quality of Life) of patients, compared with age-matched people without the condition.

ii) Annual number of patients to be treated:

The number of patients anticipated to be treated with the new medicine each year, if recommended. Note, this is not necessarily the number of patients afflicted by the disease or condition; it is the number of patients anticipated to receive treatment with the new medicine each year. There are three possible levels: 40 patients, 500 patients, or 1,000 patients to be treated per year with the new medicine if recommended.

iii) QALYs gained per treated patient:

The average number of quality-adjusted life-years gained per patient treated with the new medicine versus the current standard of care for that disease or condition. There are three possible levels:

0.1 QALYs, 0.8 QALYs, or 1.6 QALYs.

iv) Incremental cost per QALY gained:
The cost effectiveness of the new medicine presented as the incremental cost per quality-adjusted life-year (QALY) gained for the new medicine versus the current standard of care for that disease or condition. There are three possible levels:
£4,000 per QALY gained, £18,000 per QALY gained, or £40,000 per QALY gained.

v) Uncertainty in cost effectiveness is thoroughly explored:
This indicates whether or not (Yes or No) the degree of uncertainty in cost effectiveness estimates has been explored by assessing the combined uncertainty arising from several data sources (known as probabilistic sensitivity analysis). This permits an estimate of the probability of the medicine being cost effective.

An Example of a choice task is given below.

		New Medicine A		New Medicine B	
Main impact of disease *before* treatment		Survival		Quality of life	
Annual number of patients to be treated		40		500	
QALYs gained per treated patient		1.6		0.1	
Incremental cost per QALY gained		£40,000		£4,000	
Uncertainty in cost effectiveness is thoroughly explored		No		Yes	
Primary question	Which medicine would you prefer to recommend for approval (please tick **one** box, ✓)				
Secondary question	Given the choice, would you recommend approval of (please tick **one** box, ✓):	A ONLY	B ONLY	Both A & B	Neither

For each of the 28 Choice tasks you are asked a primary question, which is to select which one of New Medicine A and New Medicine B you would prefer to recommend.
You are also asked a secondary question, which is to indicate whether or not you would recommend New Medicine A only, New Medicine B only, Both A & B, or Neither. Note that if you select Neither, this would imply that both new medicines would be unavailable to patients via NHS Wales, and treatment available to patients would be the current standard of care for that disease or condition, which may include older, established medicine(s), or best supportive care where no other treatment options are available.
PLEASE NOTE: Some of the New Medicine profiles in each Choice task may appear very similar to those in other Choice tasks, and the tasks may appear to be repetitive. However, the whole of this discrete choice experiment has been designed to have specific

statistical properties. Please now complete all 28 Choice tasks ... and consider all Choice tasks as independent of each other.

Table 4.1A. Responses to each choice task, n (%)

	Forced choice model		Flexible choice model			
Choice task	A	B	A	B	Both	Neither
1	21 (51)	20 (49)	15 (37)	4 (10)	16 (39)	6 (15)
2	17 (41)	24 (59)	8 (20)	14 (34)	3 (7)	16 (39)
3	2 (5)	39 (95)	0 (0)	26 (63)	5 (12)	10 (24)
4	32 (78)	9 (22)	25 (61)	2 (5)	2 (5)	12 (29)
5	2 (5)	39 (95)	1 (2)	35 (85)	2 (5)	3 (7)
6	40 (98)	1 (2)	31 (76)	0 (0)	7 (17)	3 (7)
7	32 (78)	9 (22)	15 (37)	0 (0)	24 (59)	2 (5)
8	15 (37)	26 (63)	2 (5)	15 (37)	12 (29)	12 (29)
9	37 (90)	4 (10)	24 (59)	1 (2)	2 (5)	14 (34)
10*	41 (100)	0 (0)	39 (95)	1 (2)	1 (2)	0 (0)
11	30 (73)	11 (27)	21 (51)	1 (2)	12 (29)	7 (17)
12	12 (29)	29 (71)	1 (2)	20 (49)	11 (27)	9 (22)
13	14 (34)	27 (66)	0 (0)	13 (32)	25 (61)	3 (7)
14	39 (95)	2 (5)	31 (76)	1 (2)	7 (17)	2 (5)
15	2 (5)	39 (95)	0 (0)	26 (63)	2 (5)	13 (32)
16	31 (76)	10 (24)	4 (10)	5 (12)	12 (29)	20 (49)
17*	0 (0)	41 (100)	1 (2)	38 (93)	1 (2)	1 (2)
18	40 (98)	1 (2)	29 (71)	1 (2)	9 (22)	2 (5)
19	3 (7)	38 (93)	0 (0)	37 (90)	2 (5)	2 (5)
20	13 (32)	28 (68)	4 (10)	6 (14)	11 (27)	20 (49)
21	30 (73)	11 (27)	10 (24)	7 (17)	17 (41)	7 (17)
22	29 (71)	12 (29)	17 (41)	2 (5)	7 (17)	15 (37)
23	7 (17)	34 (83)	0 (0)	22 (54)	3 (7)	16 (39)
24	37 (90)	4 (10)	21 (51)	2 (5)	17 (41)	1 (2)
25	13 (32)	28 (68)	4 (10)	8 (20)	4 (10)	25 (60)
26	40 (98)	1 (2)	30 (73)	2 (5)	8 (20)	1 (2)
27	40 (98)	1 (2)	36 (88)	0 (0)	1 (2)	4 (10)
28	15 (37)	26 (63)	3 (7)	16 (39)	19 (46)	3 (7)
Mean Average (%)	55.2	44.8	32.4	26.6	21.1	19.9
* Dominant choice tasks used to demonstrate internal validity (rational trading behaviours)						

Table 4.2A. Stated preference and revealed preference data for period 2007-2009

Drug name	Attributes					AWMSG	Forced choice model		Flexible choice model	
	ICER (x£1000)	QALYg	UNCERTAINTY	No_PTS (x100)	IMPACT	recommendation	Predicted probability	Same as AWMSG*	Predicted probability	Same as AWMSG*
Ranolazine	16.09	0.33	1	13.13	1	Negative	0.7668	No	0.5974	No
Darunavir	0.00†	0.16	1	0.3	1	Positive	0.9115	Yes	0.8074	Yes
Degarelix	11.65	0.02	0	3.6	1	Negative	0.3597	Yes	0.3538	Yes
Mecasermin	47.54	3.78	1	0.03	0	Positive	0.8147	Yes	0.6949	Yes
Paricalcitol	10.35	0.24	0	1.45	0	Negative	0.2080	Yes	0.3873	Yes
Etravirine	28.30	0.94	1	0.46	1	Positive	0.7216	Yes	0.5342	Yes
Maraviroc	22.04	1.92	1	0.11	1	Positive	0.9269	Yes	0.7801	Yes
Nelarabine	71.39	0.28	0	0.03	1	Positive	0.0066	No	0.0133	No
Thalidomide	16.94	0.91	1	2.1	1	Positive	0.8602	Yes	0.7009	Yes
Efavirenz / Emtricitabine / Tenofovir	0.00†	0.13	1	4.08	1	Positive	0.9077	Yes	0.7994	Yes
Bivalirudin	5.32	0.04	1	5.53	1	Positive	0.8535	Yes	0.7228	Yes
Aripiprazole	23.22	0.02	1	6.41	0	Positive	0.3133	No	0.4110	No
Alemtuzumab	18.87	0.37	0	0.03	1	Positive	0.3187	No	0.3115	No
Ambrisentan	0.00†	0.15	1	0.13	1	Positive	0.9110	Yes	0.8069	Yes
Atazanavir (Treatment experienced)	6.21	0.09	1	1.35	1	Positive	0.8535	Yes	0.7247	Yes
Atazanavir (Treatment naive)	0.00†	0.24	1	1.38	1	Positive	0.9183	Yes	0.8153	Yes
Rufinamide	17.80	0.14	1	0.35	0	Positive	0.4505	No	0.5326	Yes
Raltegravir	16.47	2.18	0	0.29	1	Positive	0.8067	Yes	0.6622	Yes
Abacavir / Lamivudine	0.00†	0.00	1	0.24	1	Positive	0.8968	Yes	0.7894	Yes
Teriparatide	29.13	0.11	1	0.27	0	Negative	0.2452	Yes	0.3471	Yes
Deferasirox	0.00†	0.18	1	0.71	0	Positive	0.7781	Yes	0.7904	Yes

Drug name	Attributes					AWMSG	Forced choice model		Flexible choice model	
	ICER (x£1000)	QALYg	UNCERTAINTY	No_PTS (x100)	IMPACT	recommendation	Predicted probability	Same as AWMSG*	Predicted probability	Same as AWMSG*
Buprenorphine / Naloxone	13.43	0.02	0	1.6	0	Positive	0.1388	No	0.3053	No
Lenolidomide	28.59	1.66	1	0.76	1	Negative	0.8488	No	0.6542	No
Docetaxel	1.84	2.08	1	0.18	1	Positive	0.9615	Yes	0.9376	Yes
Ziconotide	11.10	1.62	1	0.09	0	Negative	0.8782	No	0.8389	No
Tacrolimus XR	3.81	0.50	0	0.42	1	Negative	0.64	No	0.5716	No
Topotecan (Cervical Cancer)	23.92	0.17	1	0.3	1	Positive	0.6114	Yes	0.4600	No
Topotecan (Small cell lung cancer)	23.64	0.24	0	0.08	1	Positive	0.2174	No	0.2313	No
Dasatinib (CML)	39.15	1.26	1	0.88	1	Positive	0.6108	Yes	0.4149	No
Dasatinib (Ph+ ALL)	65.85	0.70	0	0.24	1	Negative	0.0160	Yes	0.0254	Yes
Idursulfase	564.58	5.96	0	0.02	1	Negative	0.0000	Yes	0.0000	Yes
Agalsidase alfa	252.95	3.51	0	0.3	1	Positive	0.0000	No	0.0000	No
Co-carledopa intest gel	84.20	0.88	0	0.21	0	Negative	0.0015	Yes	0.0079	Yes
Darunavir	15.51	1.38	1	0.15	1	Positive	0.9216	Yes	0.7868	Yes
Sunitinib	29.40	0.69	1	0.29	1	Negative	0.6445	No	0.4721	Yes
Tipranavir	30.52	0.70	0	0.14	1	Positive	0.2112	No	0.2107	No
Emtricitabine	18.00	0.50	1	0.05	1	Positive	0.7839	Yes	0.6228	Yes
Emtricitabine / Tenofovir	18.00	0.50	1	0.82	1	Positive	0.7832	Yes	0.6213	Yes
Parathyroid hormone	43.18	0.07	0	5.06	0	Positive	0.0156	No	0.0589	No

Drug name	Attributes					AWMSG	Forced choice model		Flexible choice model	
	ICER (x£1000)	QALYg	UNCERTAINTY	No_PTS (x100)	IMPACT	recommendation	Predicted probability	Same as AWMSG*	Predicted probability	Same as AWMSG*
Summary data										
Mean values – Positive AWMSG Recommendation	26.94	0.78	71% PSA used	1.18	79% Survival Impact	71.8% Positive recommendation	Probability Positive recommendation 0.6252		Probability Positive recommendation 0.5552	
Mean values – Negative AWMSG Recommendation	77.70	1.15	45% PSA used	1.86	64% Survival Impact	28.2% Negative recommendation	Probability Negative recommendation 0.4190		Probability Negative recommendation 0.3869	
Correctly classified by model							64.10%		64.10%	
Sensitivity of model							67.86%		64.29%	
Specificity of model							54.55%		63.64%	
Area under ROC curve							0.6120		0.6396	

UNCERTAINTY: 1= Thoroughly explored with PSA, 0= Not;
IMPACT: 1=Main impact of disease before tretament is on survival, 0=Quality of life
* Recommendation using a predicted probability threshold of 50%
†ICER value assumed zero for medicines that dominate their modelled comparator in AWMSG submission

Appendix to Chapter 5

Public Survey Questionnaire

Introduction: We are interested in your views on how NHS spending should be prioritised. In the following questions you will be presented with several imaginary scenarios. Please read these carefully and indicate your preferred way for the NHS to spend money. Please try to answer every question. There are no right or wrong answers; we are simply interested in your views.

Scenario X

(Same for both Cohorts)

Imagine two diseases - Disease A and Disease B. They affect the same age groups and are equally common. The only difference between the two diseases is that, without treatment:
Disease A — causes **severe health problems** that affect patients' well-being considerably
Disease B — causes **moderate health problems** that have less effect on patients' well-being.

Medicine A (for treatment of Disease A) and Medicine B (for the treatment of Disease B) both improve the health and well-being of patients by the same amount, and they cost the same. As the NHS has a fixed amount of money, and there are no extra funds available. Treatment of patients using either Medicine A or Medicine B may mean that other treatments or services for other patients have to be reduced. **(Text appeared in all questions)**

If the NHS was able to pay for treatment for a maximum of:

- **100 patients with severe health problems** due to Disease A using Medicine A, or
- **100 patients with moderate health problems** due to Disease B using Medicine B, or
- **some combination** of the two,

how would you prefer NHS money to be spent? Please indicate using the scale below.

	All money spent on Medicine A to treat patients with **severe health problems**					Money divided equally					All money spent on Medicine B to treat patients with **moderate health problems**
	100 patients with severe health problems to be treated with Medicine A	90 patients with severe health problems	80 patients with severe health problems	70 patients with severe health problems	60 patients with severe health problems	50 patients with severe health problems	40 patients with severe health problems	30 patients with severe health problems	20 patients with severe health problems	10 patients with severe health problems	0 patients with severe health problems to be treated with Medicine A
	0 patients with moderate health problems to be treated with Medicine B	10 patients with moderate health problems	20 patients with moderate health problems	30 patients with moderate health problems	40 patients with moderate health problems	50 patients with moderate health problems	60 patients with moderate health problems	70 patients with moderate health problems	80 patients with moderate health problems	90 patients with moderate health problems	100 patients with moderate health problems to be treated with Medicine B
Tick one box	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Cohort 1 only:
Now imagine that Medicine A and Medicine B still cost the same, but the **improvement in health and well-being from the two medicines differ**:

Medicine A — will **improve health a little** in patients with Disease A (who have **severe health problems before treatment**),
Medicine B — will **improve health considerably** in patients with Disease B (who have **moderate health problems before treatment**).

If the NHS was able to pay for treatment for a maximum of:
- **100 patients** with Disease A using Medicine A (patients with **severe health problems gain a little improvement in health**), or
- **100 patients** with Disease B using Medicine B (patients with **moderate health problems gain a considerable improvement in health**), or
- **some combination** of the two,

how would you prefer NHS money to be spent? Please indicate using the scale below...

[Scale as above]

Cohort 2 only:
Now imagine that Medicine A and Medicine B still both improve the health and well-being of patients by the same amount, but **the costs of the two medicines now differ**. The NHS is now able to pay for treatment for a maximum of:
- **50 patients with severe health problems** due to Disease A using Medicine A, or
- **100 patients with moderate health problems** due to Disease B using Medicine B, or
- **some combination** of the two.

How would you prefer NHS money to be spent? Please indicate using the [*revised*] scale below.....

Other Scenarios using same format (presented in random order to respondents):

Children vs. Adults

Imagine two diseases - Disease A and Disease B. They are equally common and are equally serious in how they affect patients' health and well being. The only difference between the two diseases is that:

Disease A — typically affects **children**

Disease B — typically affects **adults**.

Common vs. Rare diseases

Imagine two diseases - Disease A and Disease B. They affect the same age groups and are equally serious in how they affect patients' health and well-being. The only difference between the two diseases is that:

Disease A — is **common** (e.g. affects 500,000 patients in the UK)

Disease B — is **rare** (e.g. affects 1000 patients in the UK).

No other treatment options vs. Several other treatment options

Imagine two diseases – Disease A, which may be treated with Medicine A, and Disease B, which may be treated with Medicine B. The two diseases affect the same age groups, are equally

common and are equally serious in how they affect patients' health and well being. The only difference between the two diseases is that:

Disease A — there are **several other treatments available from the NHS**, which improve patient's health and well being by the same amount as Medicine A

Disease B — there are **no other treatments available apart from Medicine B**.

Cancer vs. Non-cancer disease

Imagine two diseases - Disease A and Disease B. They are both potentially fatal, affect the same age groups and are equally common. The number of useful medicines available to treat each disease is the same. The only difference between the two diseases is that:

Disease A — is a type of **cancer**

Disease B — is some other **non-cancer type of disease**.

Reduced life expectancy of 18 months vs. Longer life expectancy of 60 months

Imagine two diseases – Disease A and Disease B. They are both fatal, affect the same age groups and are equally common. The only difference between the two diseases is that, without treatment:

Disease A — patients would die within 18 months (one and a half years)

Disease B — patients would die within 60 months (five years).

Medicine A (for treatment of Disease A) and Medicine B (for the treatment of Disease B) both increase length of life by the same amount of **six** months, they improve patients' well-being by the same amount, and they cost the same.

Disadvantaged populations vs. Non-disadvantaged populations

Imagine two diseases – Disease A and Disease B. They affect the same age groups, are equally common and are equally serious in how they affect patients' health and well-being. The only difference between the two diseases is that:

Disease A — typically affects **disadvantaged populations, e.g. those from low income families**

Disease B — typically affects **patients from populations that are not disadvantaged**.

Medicine that works in a new way vs. Medicine that works in similar way to other available medicines

Imagine a disease that is serious for which two new medicines – Medicine A and Medicine B – have been developed. The only difference between Medicine A and Medicine B is the way in which they work:

Medicine A — works in a **similar way to other medicines** that are available for the treatment of this disease,

Medicine B — works in a **new, different way**.

Other than that, **Medicine A and Medicine B both improve patients' health and well-being by the same amount, and they cost the same**.

Disease causing reliance upon carers vs. Disease not causing reliance upon carers

Imagine two diseases – Disease A and Disease B. They affect the same age groups, are equally common, and are equally serious. The difference between the two diseases is that:

Disease A — typically **patients have to rely on carers (e.g. family members) for their day-to-day needs**

Disease B — typically **patients do not have to rely on carers**.

Medicine A (for treatment of Disease A) and Medicine B (for treatment of Disease B) cost the same. Neither medicine will cure patients, but both medicines improve patients' own health and well-being by the same amount. **Patients with Disease A treated with Medicine A will also become less reliant on carers** for their day-to-day needs.

Figure 5.1A. Preferences of respondents in Cohorts 1 (n=2033) and 2 (n=2085) under assumption of equal health

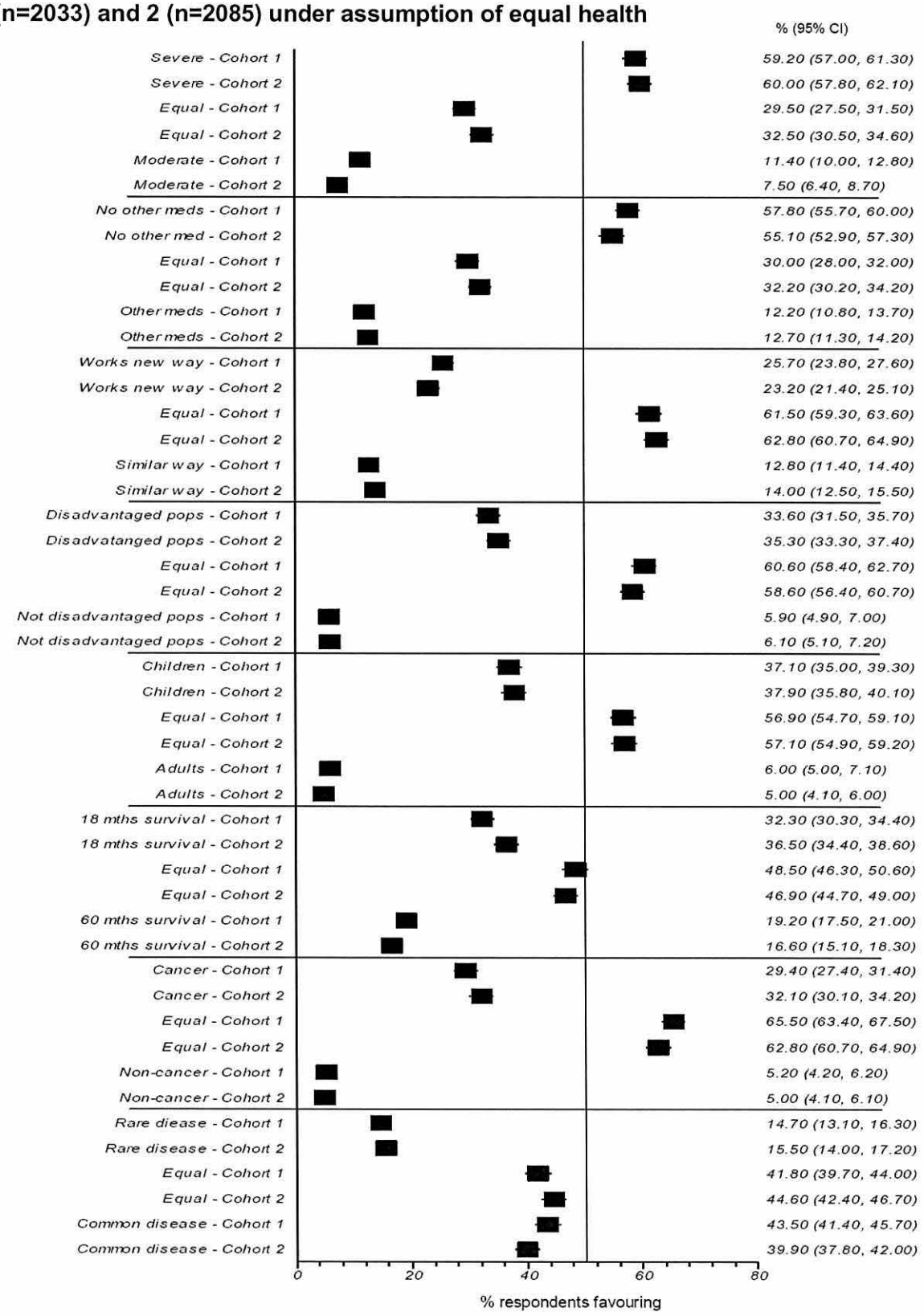
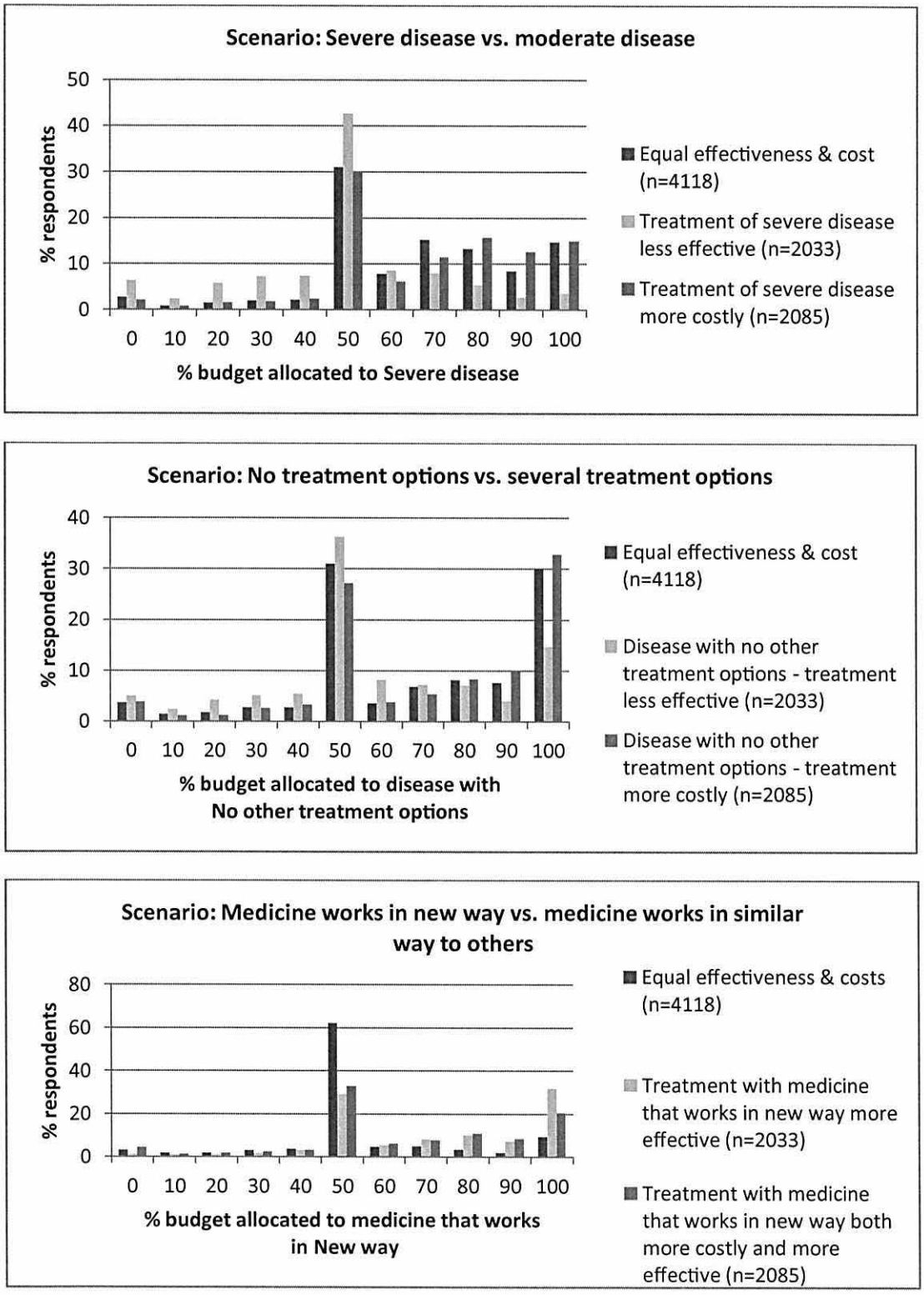
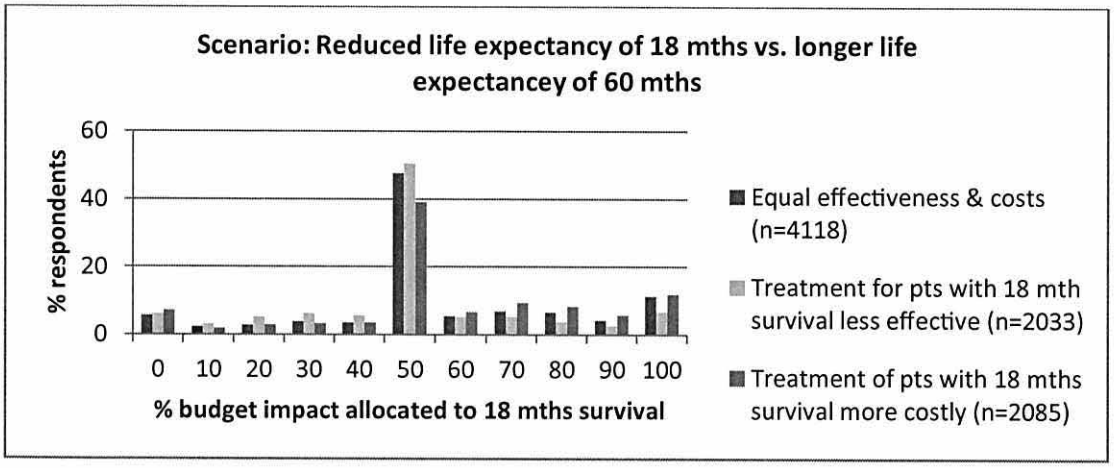
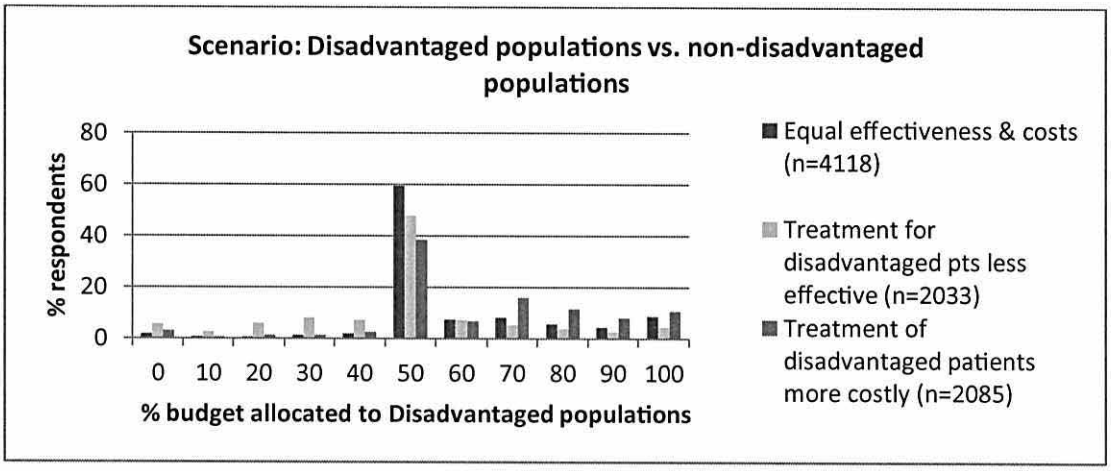
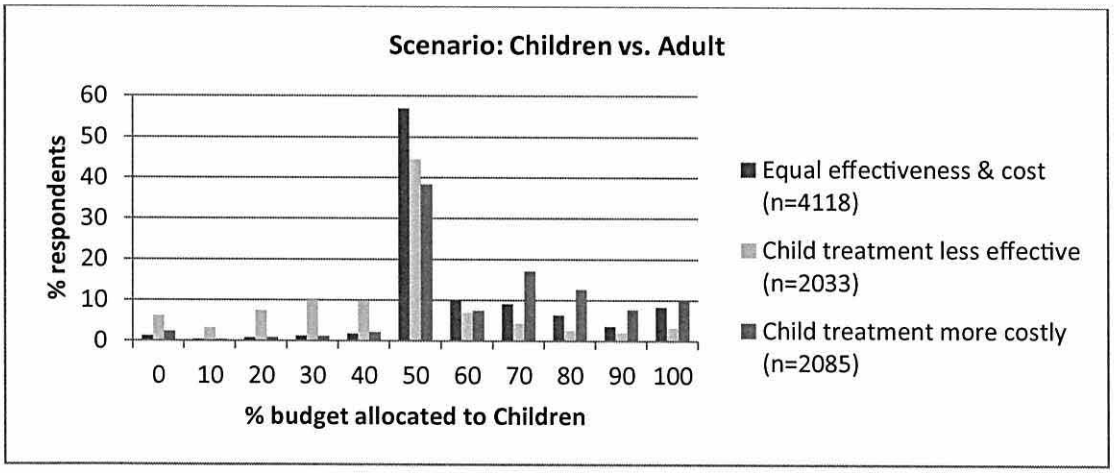


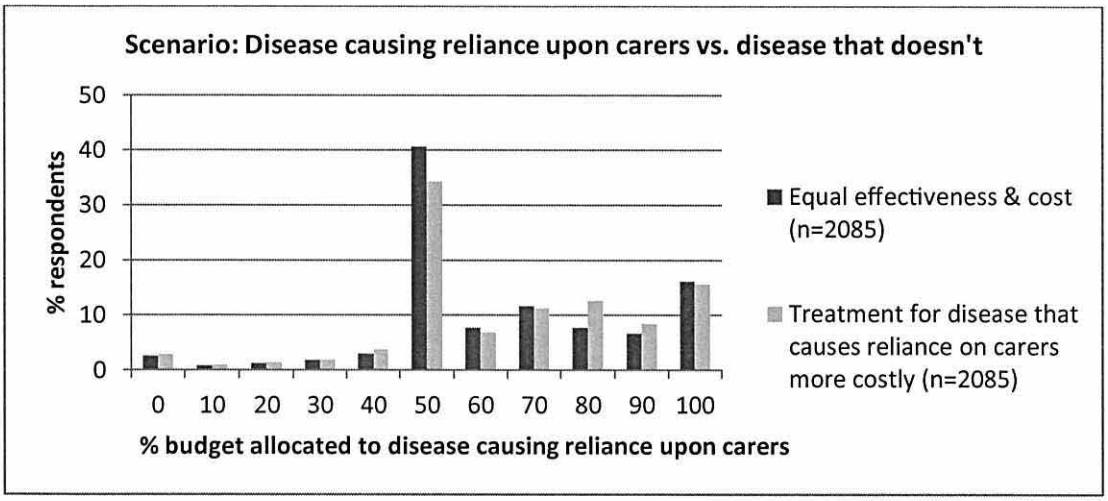
Figure 5.2A. Budget allocation preferences

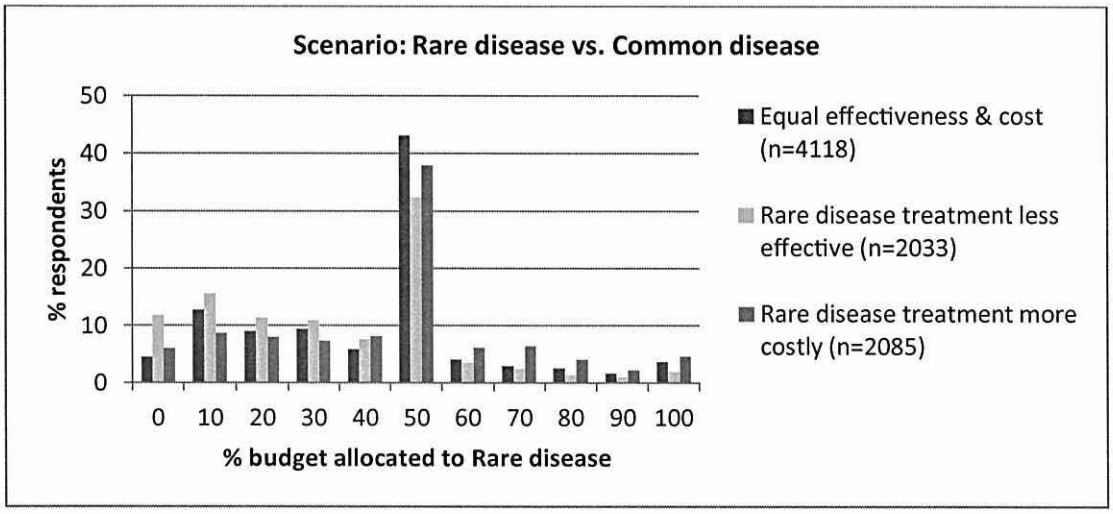


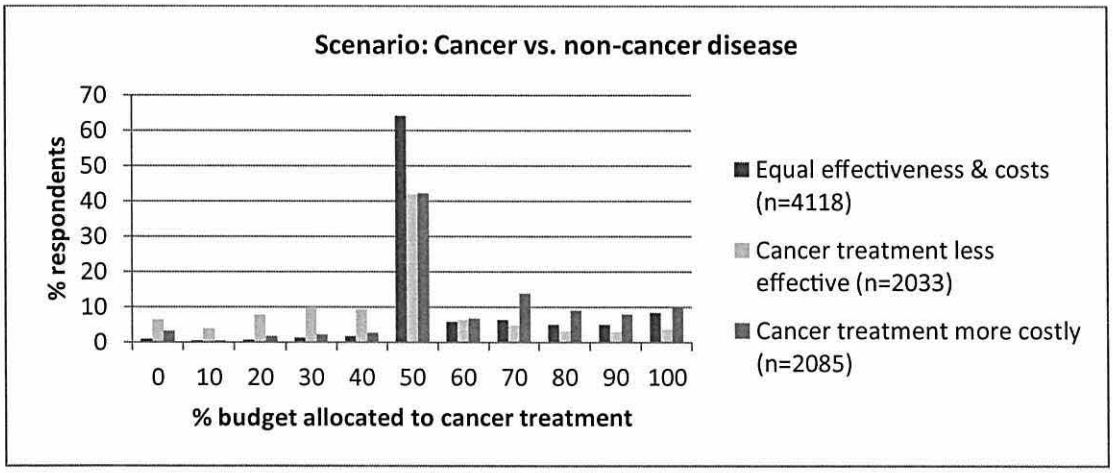

Scenario: Reduced life expectancy of 18 mths vs. longer life expectancey of 60 mths
% respondents
60
40
20
0
0 10 20 30 40 50 60 70 80 90 100
% budget impact allocated to 18 mths survival
Equal effectiveness & costs (n=4118)
Treatment for pts with 18 mth survival less effective (n=2033)
Treatment of pts with 18 mths survival more costly (n=2085)


Scenario: Disadvantaged populations vs. non-disadvantaged populations
% respondents
80
60
40
20
0
0 10 20 30 40 50 60 70 80 90 100
% budget allocated to Disadvantaged populations
Equal effectiveness & costs (n=4118)
Treatment for disadvantaged pts less effective (n=2033)
Treatment of disadvantaged patients more costly (n=2085)


Scenario: Children vs. Adult
% respondents
60
50
40
30
20
10
0
0 10 20 30 40 50 60 70 80 90 100
% budget allocated to Children
Equal effectiveness & cost (n=4118)
Child treatment less effective (n=2033)
Child treatment more costly (n=2085)


Scenario: Disease causing reliance upon carers vs. disease that doesn't
% respondents
50
40
30
20
10
0
0 10 20 30 40 50 60 70 80 90 100
% budget allocated to disease causing reliance upon carers
Equal effectiveness & cost (n=2085)
Treatment for disease that causes reliance on carers more costly (n=2085)


Scenario: Rare disease vs. Common disease
% respondents
50
40
30
20
10
0
0 10 20 30 40 50 60 70 80 90 100
% budget allocated to Rare disease
Equal effectiveness & cost (n=4118)
Rare disease treatment less effective (n=2033)
Rare disease treatment more costly (n=2085)


Scenario: Cancer vs. non-cancer disease
% respondents
70
60
50
40
30
20
10
0
0 10 20 30 40 50 60 70 80 90 100
% budget allocated to cancer treatment
Equal effectiveness & costs (n=4118)
Cancer treatment less effective (n=2033)
Cancer treatment more costly (n=2085)

Appendix for Chapter 6

Table 6.1A Summary of mapping of NICE Citizens Council Reports to other documents

CC Report	Theme – views of CC members	Consistent among NICE CC reports	Consistent with latest NICE SVJ document 2nd Ed, 2008, or other NICE process documents?	Consistent with NICE 'special' circumstances? (Rawlins et al., 2010)	Consistent with Public Survey (n=4,118)? (Chapter 5)
1. Clinical need, November 2002 (n=30)	**How bad is the pain and how severe are the symptoms?**	**Yes:** CC Report 4, 9, 10, 11	**Possibly:** SVJ Document **No/?:** Methods of TA, 2008	**Yes:** CC Report 1, 4, 9, 10, 11.	**Yes:** 59.6% prioritise the treatment of severe disease over treatment of moderate disease
	Is it potentially fatal? (saving lives, or the 'rule of rescue' is extremely important.)	**Yes:** CC Report 4, 6, 8, 11	**No:** SVJ Document	**Unclear:** [Refers to life extension – not life saving]	**Unclear:** Only 34.4% prioritise treatment of patients with life expectancy of 18 months compared with 60 months, 47.6% indifferent and remainder prioritise those with life expectancy 60 months. [Refers to life extension – not life saving]
	Is the disease contagious?	**Yes:** CC Report 11	Not reflected in this context	Not reflected	Not tested
	Are alternative treatments available?	**Yes:** CC report 6 CC report 11	**Yes:** NICE PASLU, 2009	Not reflected	**Yes:** 56.5% prioritise treatment of patients with no existing alternative treatments available.
	What is the long-term effect of the condition on the individual? (chronic conditions should be seen as more important)	**Yes:** CC Report 11	Not reflected	Not reflected	Not tested
	What are the chances of good clinical outcome? (priority should be given to treatments that have the most positive effect)	**Yes:** CC Report 2, 5	**Unclear:** SVJ Document	Not reflected	**Yes:** There was, in all cases, a statistically significant shift in preferences towards the populations that gained a considerable improvement in health and away from the population that gained a little improvement in health.

CC Report	Theme – views of CC members	Consistent among NICE CC reports	Consistent with latest NICE SVJ document 2nd Ed, 2008, or other NICE process documents?	Consistent with NICE 'special' circumstances? (Rawlins et al., 2010)	Consistent with Public Survey (n=4,118)? (Chapter 5)
	What is the number of patients affected? (if only a small number of people are affected by a particular condition, although their need may be severe, this should be less of a priority)	**No:** CC Report 4, 11	**Yes:** SVJ Document	Not reflected	**Yes:** CC Report 1 **No:** CC Report 4, 11 15.1% prioritised treatment of rare disease, 41.7% prioritised common disease and remainder were indifferent, suggesting that rare disease may be less of a priority
	What is the effect of the disease on the quality of life for the individual patient? (the effect of the disease or condition on the whole of a patient's life should be considered, including ...their families)	**Yes:** CC Report 6 CC Report 11	Not reflected	Not reflected beyond use of quality-adjusted life years (QALYs) as a common metric for treatment effectiveness.	**Yes:** 50.0% prioritised treatments for disease which causes patients to be reliant upon carers (e.g. family members) for day-to-day needs, and reduces that reliance on carers, 40.6% were indifferent and remainder prioritised treatments that were for patients without reliance on informal carers.
	What is the effect of the disease on the length of life for the individual?	**No:** CC report 11	**Yes:** EoL policy, 2009	**Yes:** CC report 1 **No:** CC report 11	**Yes:** CC Report 11 **No:** CC Report 1 34.4% prioritised patients with life expectancy 18 months, 47.6% indifferent, and remainder prioritised patients with life expectancy 60 months
	What are the psychological effects of the condition?	**Yes:** CC Report 6	**Unclear:** SVJ Document	Not reflected	Not tested
	What is the level of disability and/or independence of the individual?	**Yes:** CC Report 9 on Patient safety	Not reflected	Not reflected	**Yes:** 50.0% prioritised treatments for disease which causes patients to be reliant upon

CC Report	Theme – views of CC members	Consistent among NICE CC reports	Consistent with latest NICE SVJ document 2nd Ed, 2008, or other NICE process documents?	Consistent with NICE 'special' circumstances? (Rawlins et al., 2010)	Consistent with Public Survey (n=4,118)? (Chapter 5)
					carers (e.g. family members) for day-to-day needs, and reduces that reliance on carers, 40.6% were indifferent and remainder prioritised treatments that were for patients without reliance on informal carers.
	Is the condition time limited? (Higher priority should be given to chronic disease)	**Yes:** CC Report 11	Not reflected	Not reflected	Not tested
	Are there fluctuations in the individual's condition?	Not reflected	Not reflected	Not reflected	Not tested
	Is the disease or the condition cosmetic? (relevant if cosmetic conditions have an adverse effect on the patient's mental health and emotional well-being)	Not reflected	**Unclear:** SVJ Document	Not reflected	Not tested
	What are the side effects [of disease] encountered by the patient?	Not reflected beyond severity	Not reflected beyond severity	Not reflected	Not tested
	Is there any stigma related to the condition?	Not reflected	**Yes:** SVJ Document	Not reflected	Not tested
	What are the resources available, such as cost and equipment? (important to recognise that resources are limited, and that in some cases individual choice should sometimes be limited in the interests of the overall population)	Not reflected	**Yes:** SVJ Document	Not reflected beyond use of cost effectiveness.	Not tested
	What values does the patient have?	**Yes:** CC Report 5	**Yes:** SVJ document	Not reflected	Not tested
	What is the patient's ability to make an informed decision?	Not reflected	Not reflected	Not reflected	Not tested
	What is the age of the patient? (age of a patient should be taken into account)	**Yes:** CC Report 11 **No:**	**Yes:** SVJ Document;; Methods of TA; Positively Equal vs. CC	**Yes:** CC Reports 1 and 11. **No:**	**Yes:** CC Report 2 **No:**

CC Report	Theme – views of CC members	Consistent among NICE CC reports	Consistent with latest NICE SVJ document 2nd Ed, 2008, or other NICE process documents?	Consistent with NICE 'special' circumstances? (Rawlins et al., 2010)	Consistent with Public Survey (n=4,118)? (Chapter 5)
		CC report 2	Report 2 **No:** SVJ Document; Methods of TA; Positively Equal vs.CC Report 1,11	CC Report 2	CC Report 1 and 11 37.5% prioritised treatments for children over adults, 57.0% were indifferent, remainder favoured adults
	How fit is the patient to undergo treatment?	Not reflected in this context	Not reflected in this context	Not reflected in this context	Not tested in this context
	What are the patient's other conditions?	Not reflected	Note reflected	Not reflected	Not tested
	How able is the patient to self-manage their condition?	Not reflected in this context	Not reflected in this context	Not reflected in this context	Not tested in this context
	What is the family history, and are there any genetic/hereditary issues for the patient?	Not reflected	Not reflected	Not reflected	Not tested
	Social and economic factors should never be a factor	**Yes:** CC Report 11 **No:** CC Report 7	**Yes:** Methods of TA **No:** SVJ document; Methods for PH guidance **?:** Positively Equal	**Yes:** CC Report 1, 7 **No:** CC Report 11	**Yes:** CC Report 1, 11 **No:** CC Report 7 34.5% prioritised treatments for disadvantaged populations, 59.5% were indifferent and the remainder prioritised treatments for those not disadvantaged
	'Self-induced' diseases or conditions should not be a factor at all	Not reflected	**Yes:** SVJ Document	Not reflected	Not tested
	How loud the 'voice' of the patient is should not be a factor	**Yes:** CC Report 11	Not reflected	**Unclear:** Patients and their advocates...can explain where symptomatology of their condition is poorly reflected in clinical trials and health-related quality of life measure	Not tested
	In the Council's opinion different weight should be given to the views of different	Not reflected	Not reflected	Not reflected	Not tested

CC Report	Theme – views of CC members	Consistent among NICE CC reports	Consistent with latest NICE SVJ document 2nd Ed, 2008, or other NICE process documents?	Consistent with NICE 'special' circumstances? (Rawlins et al., 2010)	Consistent with Public Survey (n=4,118)? (Chapter 5)
	stakeholders in deciding clinical need				
2. Age, November 2003 (n=29)	**Age as indicator of risk:** (22/29 agreed relevant)	Not reflected in this context	**Yes:** SVJ Document	Not reflected in this context	Not tested in this context
	Age as a determinant of clinical effectiveness: (25/29 agreed relevant)	Not reflected in this context	**Yes:** SVJ Document	Not reflected in this context	Not tested in this context
	Age influencing social roles: (22 /29 disagreed relevant)	**No:** CC Report 1, 11	**Yes:** SVJ Document; Methods of TA; Positively Equal	**Yes:** CC Reports 1 and 11 **No:** CC Report 2	**Yes:** CC Report 2 **No:** CC Report 1, 11 37.5% prioritised treatments for children over adults, 57.0% were indifferent, remainder favoured adults
	How much chance people have had to experience life due to their age? "Fair innings"? (21/29 disagreed relevant)	**No:** CC Report 1, 11	**Yes:** SVJ Document; Methods of TA; Positively Equal	**Yes:** CC Reports 1 and 11 **No:** CC Report 2	**Yes:** CC Report 2 **No:** CC Report 1, 11 37.5% prioritised treatments for children over adults, 57.0% were indifferent, remainder favoured adults
3. Confidential enquiries, May 2004	N/A	Not reflected	Not reflected	Not reflected	Not tested
4. Ultra-orphan drugs, November 2004 (n=27)	**Should NHS consider paying premium prices for drugs to treat patients with very rare diseases?** 16/27 thought yes with certain conditions: • **The degree of severity of the**	**Yes:** CC Report 11 **No:** CC report 1	**No:** SVJ Document; Methods of TA	Not reflected	**Yes:** CC Report 1 **No:** CC Report 4, 11 15.5% prioritised rare disease

CC Report	Theme – views of CC members	Consistent among NICE CC reports	Consistent with latest NICE SVJ document 2nd Ed, 2008, or other NICE process documents?	Consistent with NICE 'special' circumstances? (Rawlins et al., 2010)	Consistent with Public Survey (n=4,118)? (Chapter 5)
	disease				over common disease, 45.2% indifferent, and remainder favour treatment of common disease. No evidence of a preference for treating rare diseases over common diseases
	• If the treatment will provide health gain, rather than just stabilisation of the condition	Not reflected	Not reflected	Not reflected	Not tested
	• If the disease or condition is life-threatening	**Yes:** CC Report 1, 6, 8, 11	**No:** SVJ Document	**Unclear:** [Refers to life extension – not life saving]	**Unclear:** Only 34.4% prioritise treatment of patients with life expectancy of 18 months compared with 60 months, 47.6% indifferent and remainder prioritise those with life expectancy 60 months. [Refers to life extension – not life saving]
5. CC Report on Mandatory Public Health Measures, July 2005 (n=24)	**Who has responsibility for the public's health, individuals or the state?**	Not reflected	**Yes:** SVJ Document	Not reflected	Not tested
	Where does the balance lie between needs and benefits versus harm and inconvenience?	Not reflected	**Yes:** Methods for PH **?:** SVJ Document	Not reflected	Not tested
	When and how should the state intervene?	**Yes:** CC Report 7	**Yes:** SVJ Document **No:** Methods of TA **?:** Methods for PH	Not reflected	Not tested
	How should mandatory interventions be introduced and monitored?	Not reflected	**Yes:** SVJ Document	Not reflected	Not tested
	Openness, trust and public	Not reflected	**Yes:**	Not reflected	Not tested

CC Report	Theme – views of CC members	Consistent among NICE CC reports	Consistent with latest NICE SVJ document 2nd Ed, 2008, or other NICE process documents?	Consistent with NICE 'special' circumstances? (Rawlins et al., 2010)	Consistent with Public Survey (n=4,118)? (Chapter 5)
	involvement		SVJ Document		
6. Rule of rescue, January 2006 (n=27)	**Should NICE reject the Rule of Rescue?** 21/27 said not in certain exceptional cases: **Is the intervention required to avoid immediate loss of life?**	**Yes:** CC Report 1	**No:** SVJ Document	Not reflected	Not tested
	Is there a good chance of an increased life expectancy?	**Yes:** CC Report 1, 2, 5	**Unclear:** SVJ Document:	Not reflected	Not tested
	Will it result in a significant improvement in quality of life? Are the treatment's side effects very severe and do they outweigh the good the treatment would do?	**Yes:** CC Report 1	Not reflected beyond use of QALYs	Not reflected	Not tested
	What will be the consequences should the treatment not be received?	Not reflected beyond severity	Not reflected beyond severity	Not reflected	Not tested
	What are the alternative treatments and how do they compare?	**Yes:** CC Report 1, 11	**Yes:** NICE PASLU, 2009.	Not reflected	**Yes:** CC Report 1, 11 56.5% prioritise treatment of patients with no existing alternative treatments available.
	Are future medical gains probable because of the research engendered by the treatment?	**Unclear:** CC Report 8	**Unclear:** SVJ document	Not reflected	Not tested
	Are the costs prohibitive to the NHS? To what extent does it increase the burden of costs on the NHS and society at large?	Not reflected beyond broad balance of costs and benefits	Not reflected beyond broad balance of costs and benefits	Not reflected beyond broad balance of costs and benefits	Not tested in this context
	To what extent is cost effectiveness demonstrable?	Not reflected beyond broad	Not reflected beyond broad balance of costs	Not reflected beyond broad balance of costs and benefits	Not tested in this context

CC Report	Theme – views of CC members	Consistent among NICE CC reports	Consistent with latest NICE SVJ document 2nd Ed, 2008, or other NICE process documents?	Consistent with NICE 'special' circumstances? (Rawlins et al., 2010)	Consistent with Public Survey (n=4,118)? (Chapter 5)
		balance of costs and benefits	and benefits		
	Are there good grounds for believing it would set a precedent for other patient groups lobbying for less cost effective treatments?	Not reflected	Not reflected	Not reflected	Not tested
	Will it avert danger to public health e.g. threat of an epidemic?	**Yes:** CC Report 1	Not reflected in this context	Not reflected	Not tested
	Will people feel society's worth is diminished if it appears to be acting inhumanely by ignoring the Rule of Rescue?	Not reflected	Not reflected	Not reflected	Not tested
7. Health inequalities, June 2006 (n=26)	**Should NICE issue guidance that concentrates resources on trying to improve the health of the most disadvantaged members of society, thus narrowing the gap between the least and most disadvantaged, even if this has only a modest impact on the health of the population as a whole?** (15/26 agreed should)	**Yes:** CC Report 5 **No:** CC Report 1, 11	**Yes:** Positively Equal **No:** Methods for PH; Methods of TA	**Yes:** CC Report 5 and 7. **No:** CC Report 11	**Yes:** CC Report 1, 11 **No:** CC Report 5, 7 34.5% prioritised treatments for disadvantaged populations over non-disadvantaged populations, 59.5% were indifferent and the remainder prioritised non-disadvantaged populations. No evidence of a preference for prioritising disadvantaged populations
8. CC Report on Only in Research, January 2007 (n=27)	**What circumstances should NICE consider to make an "only in research" (OIR) recommendation?** 15 other circumstances listed in addition to those below	Not reflected	Not reflected	Not reflected	Not tested
	NICE may wish to consider how OIR could be used as means of encouraging	**Unclear:** CC report 6	**Unclear:**	Not reflected	Not tested

CC Report	Theme – views of CC members	Consistent among NICE CC reports	Consistent with latest NICE SVJ document 2nd Ed, 2008, or other NICE process documents?	Consistent with NICE 'special' circumstances? (Rawlins et al., 2010)	Consistent with Public Survey (n=4,118)? (Chapter 5)
	innovation		SVJ Document		
	With treatments for life-threatening conditions where there is no other remedy available, NICE should consider granting it the "benefit of the doubt" with an OIR decision rather than a "no". (25/27, 93%)	**Yes:** CC Report 1, 4, 6, 11	Not reflected	Not reflected	Not tested
9. CC Report on Patient safety, June 2007 (n=22)	**Is it appropriate when developing "patient safety solutions" that NICE take the costs, as well as the benefits, into account?** Majority agreed it is. Case for moving a cost effectiveness threshold include: **The severity to an individual of any likely injury or harm resulting from the error**	**Yes:** CC Report 1, 4, 10, 11	**Yes:** SVJ Document	**Yes:** CC Report 1, 10, 11	**Yes:** CC Report 1, 10, 11 59.6% prioritise the treatment of severe disease over treatment of moderate disease
	The wider cost to society of coping with the aftermath cost to those left caring or bereaved, cost of litigation	**Yes:** CC Report 1	Not reflected	Not reflected	**Yes:** CC Report 1 50.0% prioritised treatments for disease which causes patients to be reliant upon carers (e.g. family members) for day-to-day needs, and reduces that reliance on carers, 40.6% were indifferent, remainder prioritised treatments for patients without reliance on informal carers.
	The extent to which the error is unique to the medical environment	**Yes:** CC Report 11	Not reflected	Not reflected	Not tested
	Failure to address the safety	Not reflected	Not reflected	Not reflected	Not tested

CC Report	Theme – views of CC members	Consistent among NICE CC reports	Consistent with latest NICE SVJ document 2nd Ed, 2008, or other NICE process documents?	Consistent with NICE 'special' circumstances? (Rawlins et al., 2010)	Consistent with Public Survey (n=4,118)? (Chapter 5)
	issue in question could have a severely damaging effect on public confidence in the NHS				
10. QALYS and severity of disease, February 2008 (n=26)	**Should NICE take into account severity of disease when making decisions?**	**Yes:** CC Report 1, 4, 9, 11	**Possibly:** SVJ Document **No/?:** Methods of TA, 2008	**Yes:** CC Report 1, 4, 9, 11	**Yes:** CC Report 1, 4, 9, 11 59.6% prioritise the treatment of severe disease over treatment of moderate disease
11. Departing from the threshold, November 2008 (n=29)	**Possible circumstances in which NICE should depart from the established threshold were:** **- the treatment in question is life-saving 24/29 (83%)**	**Yes:** CC Report 1, 6	Not reflected beyond broad consideration of costs and benefits, and consideration of clinical need	Not reflected [Refers to life extension – not life saving]	Not reflected 34.4% prioritise treatment of patients with life expectancy of 18 months compared with 60 months, 47.6% indifferent and remainder prioritise those with life expectancy 60 months. [Refers to life extension – not life saving]
	- the illness is a result of NHS negligence 23/29 (79%)	**Yes:** CC Report 9	Not reflected	Not reflected	Not tested
	- the intervention would prevent more harm in the future 23/29 (79%)	**Yes:** CC Report 1	Not reflected	Not reflected	Not tested
	- the patients are children 22/29 (76%)	**Yes:** CC Report 1 **No:** CC Report 2	**Yes:** SVJ Document vs. CC Report 2 **No:** SVJ Document; Methods of TA; Positively Equal vs. CC Reports 1 and 11	**Yes:** CC Reports 1 and 11. **No:** CC Report 2	**Yes:** CC Report 2 **No:** CC Report 1 and 11 37.5% prioritised treatments for children over adults, 57.0% were indifferent, remainder favoured adults
	- the intervention will have a major impact on the patient's family 22/29 (76%)	**Yes:** CC Report 1	Not reflected	Not reflected	**Yes:** CC Report 1 and 11 50.0% prioritised treatments for disease which causes

CC Report	Theme – views of CC members	Consistent among NICE CC reports	Consistent with latest NICE SVJ document 2nd Ed, 2008, or other NICE process documents?	Consistent with NICE 'special' circumstances? (Rawlins et al., 2010)	Consistent with Public Survey (n=4,118)? (Chapter 5)
					patients to be reliant upon carers (e.g. family members) for day-to-day needs, and reduces that reliance on carers, 40.6% were indifferent and remainder prioritised treatments that were for patients without reliance on informal carers.
	- the illness under consideration is extremely severe 21/29 (72%)	**Yes:** CC Report 1, 4, 9, 10	**Possibly:** SVJ Document **No/?:** Methods of TA, 2008	**Yes:** CC Report 1, 4, 9, 10	**Yes:** CC Report 1, 4, 9, 10 59.6% prioritise the treatment of severe disease over treatment of moderate disease
	- the intervention will encourage more scientific and technical innovation 21 / 29 (72%)	**Yes:** CC Report 6	**Unclear:** SVJ Document	Not reflected in this context	Not tested in this context
	- the illness is rare 20/29 (69%)	**Yes:** CC Report 4 **No:** CC Report 1	**No:** SVJ Document	Not reflected	**Yes:** CC Report 1 **No:** CC Reports 4 and 11 15.5% prioritised rare disease over common disease, 45.2% indifferent, and remainder favour treatment of common disease. No evidence of a preference for treating rare diseases over common diseases
	- there are no alternative therapies available 19 /29 (66%)	**Yes:** CC Report 1, 6	Not reflected beyond broad balance of costs and clinical need	Not reflected	**Yes:** CC Report 1, 6 56.5% prioritise treatment of patients with no existing alternative treatments available.
	- the intervention will have a major impact on society at	**Yes:** CC Report 1	Not reflected	Not reflected	**Yes:** CC Report 1, 11

CC Report	Theme – views of CC members	Consistent among NICE CC reports	Consistent with latest NICE SVJ document 2nd Ed, 2008, or other NICE process documents?	Consistent with NICE 'special' circumstances? (Rawlins et al., 2010)	Consistent with Public Survey (n=4,118)? (Chapter 5)
	large 16 /29 (55%)	**No:** CC Report 4			**No:** CC Report 4
	- the patients concerned are socially disadvantaged 13/29 (45%) [i.e. minority view]	**Yes:** CC Report 1 **No:** CC Report 7	**Yes:** Methods of TA; Positively equal vs. CC Report 1, 11 SVJ Document vs. CC Report 7 **No:** SVJ Document vs. CC Report 1, 11 SVJ Document vs. Methods of TA Methods of TA vs. CC Report 7	**Yes:** CC Report 7 **No:** CC Report 1, 11	**Yes:** CC Report 1, 11 **No:** CC Report 7 34.5% prioritised treatments for disadvantaged populations, 59.5% were indifferent and the remainder prioritised treatments for those not disadvantaged
	- the treatment is life extending 10/29 (34%) [i.e. minority view]	**No:** CC Report 1	**No:** Supplementary EoL policy	**Yes:** CC Report 1 **No:** CC Report 11 [Consistent with Supplementary advice to appraisal committees on appraising life-extending, end of life treatments, 2009.]	**Yes:** CC Report 11 **No:** CC Report 1 34.4% prioritise treatment of patients with life expectancy of 18 months compared with 60 months, 47.6% indifferent and remainder prioritise those with life expectancy 60 months. [Not consistent with Supplementary advice to appraisal committees on appraising life-extending, end of life treatments, 2009.]
	- the condition being tackled is time-limited 9/29 (31%) [i.e. minority view]	**Yes:** CC Report 1	Not reflected	Not reflected	Not tested
	- the illness is a result of	Not reflected	Not reflected	Not reflected	Not tested

CC Report	Theme – views of CC members	Consistent among NICE CC reports	Consistent with latest NICE SVJ document 2nd Ed, 2008, or other NICE process documents?	Consistent with NICE 'special' circumstances? (Rawlins et al., 2010)	Consistent with Public Survey (n=4,118)? (Chapter 5)
	corporate negligence 2/29 (7%) [i.e. minority view]	(beyond NHS negligence above)			
	- the stakeholders happen to be highly persuasive 0/29 (0%) [i.e. minority view]	**Yes:** CC Report 1	Not reflected	**Unclear:** Patients and their advocates...can explain where symptomatology of their condition is poorly reflected in clinical trials and health-related quality of life measure	Not tested
12. Innovation, May 2009 (n=28)	-	Not reflected in this context	Not reflected in this context	Not reflected in this context	Not tested in this context
13. Smoking and harm reduction, October 2009 (n=28)	-	Not reflected in this context	Not reflected in this context	Not reflected in this context	Not tested
14. Incentives for behaviour change, May 2012 (n=32)	-	Not reflected	Not reflected	Not reflected	Not tested
15. Discounting, November 2011 (n=28)	-	Not reflected	Not reflected	Not reflected	Not tested

Methods of PH = Methods for the development of Public Health guidance, 3rd Ed, 2012
Methods of TA = Guide to the methods of technology appraisal, 2008
NICE PASLU = Process for advising on the feasibility of implementing a patient access scheme (Interim), September 2009
Supplementary EoL policy = Supplementary advice to appraisal committees: Appriasing life-extending, end-of-life treatments, 2009
Positively equal: a guide to addressing equality issues in developing NICE clinical guidelines , 20
SVJ Document = Social Value Judgements: Principles for the development of NICE guidance, 2nd Ed, 2008