

Paper presented to the Health Economists Study Group Meeting,
University of Bangor, 29 June-1July 2011

Looking for the 'Values' to inform Value-Based Pricing: A Review of the Empirical Ethics (*Equity*) Evidence!

Colin Green, University of Exeter¹

1. INTRODUCTION

The debate concerning the weighting of health outcomes to reflect priority-setting objectives, and concerns over the use of health maximisation (QALY maximisation) as the relevant maximand in a health technology appraisal context, have contributed to a growing body of work that examines potential social considerations that may be used to weight outcomes, or to inform health decisions outside of economic evaluation. This type of research has been referred to as 'empirical ethics' (Culyer 2001, Richardson & McKie 2005), and is characterised as research to explore the equity principles, or social values², associated with specific decisions on the funding, or not, of health care interventions. Recent proposals on the introduction of a framework for value-based pricing (VBP), for use in the assessment of new medicines for the NHS, brings the consideration of the weighting of health outcomes, and the social values associated with health technology appraisal, into a more prominent policy setting. This paper presents findings from a review of the empirical ethics (equity) literature, together with reflections on how the current literature may or may not be of use to inform a future framework for value-based pricing (VBP), and other difficult priority-setting dilemmas.

The paper starts with a brief introduction to VBP, it presents a summary review of the empirical equity literature, and thereafter it presents some reflections on how health economics and specifically CEA may fit with a VBP framework - presented to the HESG audience as *work(thoughts)-in-progress* in order to encourage discussion.

2. VALUE-BASED PRICING

The rationale for, and current introductory detail on, the proposed introduction of VBP have been set out in a consultation document from the Dept of Health (2010). VBP has been introduced as "A new value-based approach to the pricing of branded medicines" (DH, December 2010). The DH refer to PPRS³ as the current system of pricing medicines, and that it has not promoted innovation or access in the way the Govt is looking for. The DH also refer to the NHS being in the position of having to pay high prices [for medicines] that are not always justified by the benefits of the new medicines. VBP is proposed as a means of providing a closer link between the price the NHS pays and the value that a medicine delivers. Section 4 of the consultation document sets out the outline of the VBP proposals, highlighting that the key principle of VBP is to ensure NHS funds are used to gain the greatest possible value for

¹ email:colin.green@pms.ac.uk

² Terminology varies, and social values are also referred to as distributive preferences (e.g. Culyer, 2001), distributive principles (e.g. Wilmot *et al*, 2004), social preferences (e.g. Bleichrodt *et al*, 2005), equity arguments (Nord *et al*, 1995), decision-making criteria, preferences over alternative distributions of health care (e.g. Sassi *et al*, 2001), public preferences (e.g. Dolan & Tsuchiya, 2005), societal values (e.g. Schwappach, 2002), impersonal preferences (Schwappach & Strasmann, 2006), other regarding preferences (Binmore, 2010), as well as other variations around these terms.

³ Pharmaceutical Price Regulation Scheme: PPRS controls the prices of branded medicines through regulating the profits that pharma companies are allowed to make on their sales to the NHS.

patients, suggesting that the Government would set a range of thresholds or maximum prices reflecting the different values that medicines offer.

In the proposals, the value of new products would be assessed and their benefits compared with the benefits that could be gained if the funds required were used to help patients elsewhere in the NHS. The consultation draws attention to the common use of CEA, and the use of a “common currency” for quantifying benefits in a consistent and comparable way across the full range of health-related conditions – referring to the QALY, currently used by NICE, as one (but not the only) option for this common currency measure of benefit. With reference to existing NICE methodology, and in making the case for a change from PPRS to VBP⁴, the DH acknowledges the important role NICE plays, and recognises the continued role NICE will play, but also states “... *the current approach may not adequately reflect all the components that contribute to a treatment's full impact on health and quality of life. And it may not always be apparent how important factors that patients and society value, such as impacts on carers, are taken into account. This could be reflected more systematically and transparently in assessments of new medicines.*” (p10) This puts explicit emphasis on social value considerations.

Under the new system of value-based pricing, the Government proposals would apply weightings to the benefits provided by new medicines, which would imply a range of price [CEA?] thresholds. The consultation states that these thresholds or maximum prices would be explicitly adjusted to reflect a broader range of relevant factors so they could be used to calculate the full value of a new product. The proposals on a price threshold structure are as follows:

- i) there would be a basic threshold, reflecting the benefits displaced elsewhere in the NHS when funds are allocated to new medicines;
- ii) there would be higher thresholds for medicines that tackle diseases where there is greater “burden of illness”: the more the medicine is focused on diseases with unmet need or which are particularly severe, the higher the threshold;
- iii) there would be higher thresholds for medicines that can demonstrate greater therapeutic innovation and improvements compared with other products;
- iv) there would be higher thresholds for medicines that can demonstrate wider societal benefits [e.g. reduced reliance on carers].

Two very helpful discussion papers have been released (Claxton et al 2011, Dixon et al 2011), which reflect on the nature of VBP as set out by the Dept of Health, and more specifically from the perspective of the health economist. These contributions have highlighted many important issues included / considered in the consultation document, and some important areas that have been omitted. Claxton et al, have drawn attention to the mechanics of cost-effectiveness analyses, the use of thresholds for cost-effectiveness ratios, and have stressed the importance of assessing the opportunity cost [in the weighting of other factors] in the analytical framework. Claxton et al, and Dixon et al, present commentary on each of the stated ‘other relevant factors’, and both draw attention to the importance of ‘uncertainty’ in the development of a framework for VBP (which is not specifically raised in the VBP proposals).

The above context for VBP, largely taken directly from the consultation document (see DH 2010, for more detail), brings the question of the social value of health interventions / health care to the fore – although it has been a constant challenge for the health economics community (and others) for many years. To date, identifying specific ‘other relevant factors’ in the context of priority-setting dilemmas in health care, and finding a basis upon which to weight relevant factors (as is needed for the threshold structure in VBP), has not been an easy task. Over 20 years ago the challenge of weighting benefits was prominent in the health economics literature (e.g. Williams 1988, Wagstaff 1991). Whilst some (e.g. Williams 1996, Nord *et al*

⁴ Although it appears PPRS or a replacement will still prevail, so the change is from PPRS to PPRS (replacement) plus VBP.

1999, Hauck *et al* 2003) have suggested that economic evaluation can incorporate a range of other social values (other than efficiency), in the form of a weighting of health outcomes, this approach has yet to be adopted (used explicitly) in a policy arena, and remains a theoretical suggestion (Rawlins & Culyer, 2004).

VBP is promising as a vehicle for moving the debate on the social value of health care into a more prominent and transparent setting. But it seems clear that VBP is currently set out as a 'process', and like other contributions on process (e.g. from Rawls 1971, from Daniels & Sabin 1998), any VBP framework will still have to face up to the difficulty of identifying, defining, and weighing up competing social values (in a practical context)⁵. The DH consultation on VBP does not offer insights on how this may be done. Can the current evidence base on the social value of health care, the empirical equity/ethics literature, help?

In the next section a review of the empirical equity literature is presented, drawing attention to the diverse and growing literature in this area, the gaps in the data available, and the emerging messages. Thereafter, there is some discussion on whether economic evaluation is the appropriate framework for inclusion of social values – equity considerations – in priority-setting decisions, as has been suggested for the VBP framework.

3. REVIEW OF EMPIRICAL ETHICS / EQUITY LITERATURE

There have been a number of earlier informative reviews of parts of this literature (e.g. Sassi *et al* 2001, Schwappach 2002, Dolan *et al* 2005, Shah 2009). These reviews have all adopted a specific perspective or context. The review presented here has taken a broad perspective, not limited to evidence around economic evaluation or QALY maximisation, and has sought to examine the literature presenting empirical findings from both attitudinal studies and experimental studies. The earlier reviews have been used to inform the current literature search strategy, and the identification of relevant empirical literature.

This section presents a review (in outline/general form) of the literature reporting empirical findings of relevance to the analysis/consideration of equity, fairness, the social values, and distributive preferences, associated with health care interventions (see appendix for search methods).

3.1 Research questions

The literature review sought to address the following questions:

Is the maximisation of health (health gain) a valid representation of the social value of health interventions?

Is the social value of a health intervention (a) dependent on factors other than health? (b) related to the characteristics of eligible patient groups? (c) related to the characteristics of the health intervention?

3.2 Literature search methods

This is a methodological review, and as such it is not best suited to traditional literature searching methods (e.g. such as those used for review of clinical trials). Therefore, the literature search has used a combination of methods. The search strategy (see appendix) was developed via an iterative process, with expert input from an information scientist. Searching was limited to English language materials.

Databases used for the search comprised MEDLINE, EMBASE, ECONLit, PSychLit, HMIC, SCI & SSCI. HMIC was used to capture contributions from the grey literature. Literature searches were undertaken from inception of database up to January 2007 (in the first instance), with an update search undertaken up to December 2010.

⁵ The process of VBP will also need to contend with the difficulties associated with the 'relative' and 'non-static' nature of fairness (i.e. this weighing up of competing social values) in the allocation of resources.

3.3 Literature search results

The literature search identified over 5,000 abstracts and titles (including update searching). Many of the references identified were excluded as they were clearly related to clinical ethics, ethics committees, patient end-of-life decisions (e.g. do not resuscitate decisions, decisions over ventilation removal), abortion in a policy context, HIV/AIDS related clinical issues, mental health and criminal justice. Following a detailed application of the inclusion and exclusion criteria (see appendix), often requiring case-by-case judgment, an evidence base of empirical studies was established as a basis for the empirical review. With additions from the update literature search an evidence base of 117 empirical studies was established (see Table 2).

Quality Assessment

There is no basis for formally addressing the quality of the studies included, and there is no previous literature in this area covering quality assessment. In one previous literature review (Dolan *et al*, 2005) this fact is acknowledged, and the authors have simply drawn attention to the characteristics of the study design. The current review has gone further and made a simple inference on the quality of the methods applied (as table below). This allows a simple categorical approach to be used (high, moderate, low) against the methods reported in the empirical studies. This judgment against the quality of the methods rests on the premise that typically the methods associated with a 'high' category are regarded as preferable to those for the 'medium' category, and the subsequent 'low' category (regarded as a weak assumption). These assumptions are regarded as crude, but relatively non-contentious [what is your view?]. The literature in this area is reported to be simple, undeveloped, and often of a methodological nature (Sassi *et al* 2001, Schwappach *et al* 2002, Dolan *et al* 2005), and this simple assessment of methods allows further consideration of these issues.

Table 1. Categories used to consider the quality of methods used in studies:

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

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

3.4 Findings: Narrative review of the literature

3.4.1 *Summary of the literature*

Table 2 presents a summary picture of the literature identified and reviewed to inform on the empirical evidence surrounding potential social values and priority setting criteria. The table reports summary detail on 117 studies. It gives a summary of the characteristics of the empirical studies, including the simple judgment made here against the quality of the methods applied in the studies (low, moderate, or high quality methods).

The literature is of an international nature (whilst accepting that the literature search was restricted to English language materials), with over 18 different countries represented in the literature identified. A large number of studies are from the UK (38), USA (18), and Scandinavian countries (23). Australia (9) and Canada (5) provide multiple studies, but in a number of instances the countries contributing to the literature have only 1 to 2 empirical studies published (in the literature identified). The vast majority of the studies (over 90%) have

been published since 1995. In 86 of the studies reported here, the study design is judged to be of an experimental nature (at least in part) i.e. the design includes the presentation of tasks to respondents which involve some form of choice – opportunity cost – between competing alternatives. In 53 studies the study design is judged (in some substantive part) to be 'attitudinal', with data on attitudes and opinions reported i.e. the study design involves the collection of responses around the attitudes of individuals to specific questions, issues, and/or tasks presented (e.g. categorical responses against level of agreement or disagreement with presented statements). In 23 studies there is a mix of experimental and attitudinal data, or a mix of empirical and methodological (theoretical and/or analytical) considerations.

In 20 studies (18%) sample sizes used have been 50 people or less, although in some instances these studies have been of a more detailed qualitative design, or have reported findings from more than one sample.

The simple qualitative categorisation of studies, against the methods applied, provides some indication of the 'quality' of the methods used in the evidence base identified. In at least 72 of the 117 studies the sample used is a convenience sample. In at least 64 of the studies the design used a self-complete format (e.g. postal questionnaire, self-complete questionnaire). Whilst the consideration of the methods applied is simple, it indicates that in over 50% of the studies the methods used would indicate a low quality threshold (methods). In only 19-20 studies can the methods applied be used to indicate a high quality threshold (methods). It is important to note that in a number of studies using either 'low' or 'moderate' threshold methods there may be a high 'quality' level with respect to the rigour applied to the rationale for the study and the development and exploration of findings. Conversely, in those studies presenting with high quality methods (interviews in a random sample) this may not be a true reflection of the overall quality of the study, with such studies often using simple designs and failing to address issues that may be important in the collection of data and the interpretation of findings in a health policy context.

For example, the UK study by Bowling (1996) is recorded here as having high quality methods, yet its design is simple and the findings from this study are not easily transferable to health policy decisions. The study sought the views of adults on priority setting over health services, using a simple grouped ranking of 12 specific health care descriptions e.g. "treatments for children with life-threatening illness" or "a health screening and education programme which could prevent a large number of people needing life saving operations in the future (for example, screening for cancers)". A related question on how to allocate a given sum of money led to responses that were contradictory to the answers from attitudinal questions presented.

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

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

The review of the literature presented here is of a summary nature, and is set out to highlight a number of general areas in the literature that are commonly discussed and addressed, and to allow the diverse literature to be drawn together. The review does not seek to provide

summary detail on all of the studies reviewed (although a large number are used as examples), but does seek to present a summary of the key factors present in the empirical literature. In many cases studies are able to inform on more than one of the sections presented.

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

In health economics research, and economic evaluation in particular, it has become standard practice to seek to maximise health gain inside the budget constraint (Oliver 2004). In the conduct of economic evaluation (cost utility analysis) it is now broadly accepted [for presentation of CEA/CUA], especially in the UK, that this generally entails maximising QALYs with available resources.

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

The findings presented in the studies reviewed strongly suggest that the message from the literature is of a willingness to make some form of sacrifice with respect to efficiency in order to distribute resources according to other criteria i.e. respondents chose alternatives that are clearly not providing the greatest amount of health gain, and are not the most efficient use of resources (and in some cases it is clearly a very inefficient use of resources). These general observations are drawn from a wide range of studies, that vary in design, sample size, sample characteristics, and context (e.g. Nord 1993, 1995, Ubel 1996, Ubel *et al* 1998, Ubel *et al* 1999, Ubel & Lowenstein 1996, Ubel *et al* 1999, Cookson & Dolan 1999, Dolan & Cookson 2000, Abellan-Perpinan & Pinto-Prades 1999, Block *et al* 2001, Pernerger *et al* 2002, Lindholm 1998, Schwappach 2003, Lindholm *et al* 1997, Johannesson & Gerdtham 1996, Shmueli 1999, Charny *et al* 1989, Gyrd-Hansen 2004, Oddsson 2003, Choudry *et al* 1997, Beach *et al* 2003, Edwards *et al* 2003, Anderson & Lytkens 1999, Tsuchiya & Dolan 2007, Green 2009).

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

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

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

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

3.4.3.1 Characteristics of the groups receiving health care

Age

There is a big literature on age, and it is not possible to provide detail here. A large number of studies have considered age as a priority setting criterion (e.g. Bowling 1996, Johannesson & Johannesson 1996, Lewis & Charny 1989, Busschbach *et al* 1993, Charny *et al* 1989, Varekamp *et al* 1998, Williams 1997, Tsuchiya 1999, Browning & Thomas 2001, Tsuchiya *et al* 2003, Schwappach 2003, Johri *et al* 2005, Cropper *et al* 1994, Zweibel *et al* 1993, Kuder and Roeder 1995, Nord *et al* 1995). Whilst there are a number of studies which show a preference to treat the young over the old (i.e. age as important), the broader evidence base does not support the general use of age as a priority setting criterion. When faced with pairwise choices between old versus the young, respondents to studies do tend to select the younger groups ahead of the older groups, but this is not reflective of more general findings (from attitudinal studies, and more qualitative studies) indicating age is not a valid basis on which to discriminate. This finding is supported by the work of the NICE Citizens council. At the present time it is difficult to interpret the literature on age. It is clear that further empirical research is needed to disentangle age from other confounding factors. At the present time it is not possible to say how a given benefit should or could be weighted [allocated] across different age groups. It is also clear that in the UK NHS (e.g. NICE) there is no explicit basis upon which to use age as a priority setting criterion.

Social role

There is limited empirical evidence directly available on this issue. The evidence available suggests that public preferences are opposed to discriminating on the basis of productivity losses, and between individuals or groups on the basis of social role (Schwappach 2002).

Severity

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

Empirical studies have asked respondents to choose between patient groups who will gain equally from treatment, but differ in the severity of their current (starting) health condition. They have also asked respondents to choose between groups where health gains are different across groups that are at differing levels of pre-treatment and post-treatment disease severity. In such studies many respondents want to give priority to the severely ill patient groups, even when they have much lower health gains than other patients in the choices offered (Nord 1993, Nord *et al* 1995, Ubel *et al* 1998, Ubel 1999, Oddsson 2003, Green 2009).

Other studies, often exploring a number of factors other than severity of health, have found support for severity of health as an influential factor in setting health care priorities (e.g. Ubel &

Lowenstein 1995, 1996a, 1996b, Cookson & Dolan 1999, Shmueli 1999, Gyrd-Hansen 2004, Edwards *et al* 2003, Oliver 2004, Wiseman 2005,).

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

Health related behaviour / Lifestyle

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

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

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

Social (socio-economic) status

Distinct from social role, the issue of socio-economic status, or social class, has been considered in a number of studies. This has been in the form of preferences for the allocation of health care (resources or outcomes) for the richer versus the poorer groups in society (Dolan *et al* 1999), or against health profiles against different categories of socio-economic status (Mooney *et al* 1995, Wiseman 2005), or levels of education (Dolan *et al*, 1999). Mooney *et al* (1995), and Wiseman (2005) report a strong preference to favour the lower social class groupings, whilst Anand and Wailoo (2000), and Block *et al* (2001) do not find a preference according to social class groupings. Charny *et al* (1989) do report preferences by occupation type, but there was no clear message, with some respondents preferring high social status occupations, and others preferring lower social status occupations (e.g. lorry driver, versus a teacher). However, Charny *et al* report that the majority of respondents did not want to discriminate on the basis of occupation. Dolan *et al* (1999) report 23% of respondents suggested a lower priority for the richer groups, and that 8% of respondents were in favour of giving higher priority for the less educated groups in society. Emmelin *et al* (1999) report a small number of respondents (12%) indicate a preference to give priority to those groups regarded as socially disadvantaged.

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

difficult in such studies to separate out considerations around social status specifically and the issue of inequalities.

Gender & Race (Ethnicity)

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

Prior health care consumption

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

3.4.3.2 Characteristics of the health intervention / health effect

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

Start-point/pre-treatment health

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

Health 'end-point' after treatment

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

important to be aware that the issue can appear in a number of different conceptual presentations.

Number of people treated

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

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

Health improvement/gain

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

Nature of health intervention/improvement

The nature of the health intervention, or health improvement, (i.e. life-saving versus life-enhancing), and the direction of the health improvement (i.e. preventative health care versus treatment of current health conditions) have received some attention in the empirical literature (e.g. Nord 1993, Nord *et al*, 1995, Ubel *et al* 2000). Whilst life-saving interventions are considered in a number of the empirical studies identified, it is rare that the effect of the intervention (i.e. saving lives) is isolated within the general scenarios presented, and rarely is any inference drawn directly on the nature of the intervention. Many of the studies examining severity of health, or level of health gain, have some form of 'noise' from the nature of interventions (e.g. Johri *et al*, 2005), and the literature is currently too crude and simple for conclusions to be drawn surrounding the many factors that may be at play when respondents provide preference, or attitudinal, data.

In terms of 'direction of change' (e.g. prevent vs. cure) studies by Ubel *et al* (1998), Shmueli (1999), and Schwappach (2002) have considered this issue. The study by Ubel and colleagues is the most frequently cited in this area. Ubel *et al* (1998) examined preferences over preventative versus curative health care, finding no strong preference either way. This study

does report a greater number of people preferring to fund preventive versus curative care, but the differences were not large, and the strongest preference was to give equal importance to the different forms of health care. A number of different choice scenarios were used in the survey and most had other issues present in the scenarios given (e.g. different levels of health improvement between options, different levels of severity and disability). In one of the choices given between prevent and cure the levels of health improvement were the same, and 37% of respondents preferred the prevent option, 21% preferred the cure option, whilst the remainder preferred equal importance. It is important to note that, as with many studies, a convenience sample was used, a self-complete format was used, and the questions were framed in a very specific way using two patient groups who were described as residents in a nursing home. It is therefore difficult to establish the generalisability of the findings, and their policy relevance.

The study by Shumueli (1999) included many contextual factors, including a notion of preference over prolonging life compared to the prevention of severe and permanent disability. The preferences reported were dependent on the relative outcomes (life-expectancy) for the patient groups described, and it is not clear what specific preference there was over prevention versus treatment. Schwappach (2002a) considers the social value of avoiding a decline in health, comparing preferences for 'upward movements' on the health scale versus avoidance of 'downward movements' in health. Results indicate a preference towards curative (health improving) health care rather than preventative (avoiding decline), with the majority of respondents (69%) preferring the allocation of resources to already diseased patients. But respondents also indicated that the number of people involved was an important factor in the decision making process, as they were prepared to trade-off a preference for curative versus protective health care when the number of patients was different i.e. when a larger number of patients could be saved from decline. This study was predominantly a methods study, in a convenience sample (Germany) which was not representative of the wider population. It used the person trade-off approach for elicitation of preferences, and the design of the study raises some concerns in terms of policy relevance.

In the identified literature - no studies have been identified that have considered 'end of life' in an empirical equity context (*please advise if you know of any*)⁶.

In the identified literature - no studies have been identified that have considered the relative importance of different disease areas in an empirical equity context (*please advise if you know of any*).

3.4.3.3 Other general distributive preferences

Health inequalities

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

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

⁶ One of the HESG papers – Baker *et al* – addresses this topic (and may also address the related point below too).

also evidence that respondents are not prepared to trade-off health gains to address inequalities. Dolan & Robinson (2001) report findings from two empirical studies, with one suggesting inequality aversion, and another that respondents were inequality neutral.

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

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

The review of the published literature has also been supplemented by an outline review of the guidance from the NICE technology appraisal programme, together with consideration of the reports from the NICE Citizens Council, and policy documents informing on the NICE appraisal process (e.g. NICE 2005, 2008, Culyer & Rawlins 2004, Bryan *et al* 2006, Williams & Bryan 2007). This outline review consisted of a general appraisal of the literature in this area, including published NICE guidance, but it did not consist of a formal systematic review of NICE guidance. The reports from the Citizens Council have assisted NICE in its development of guidelines on social value judgments for NICE and its advisory bodies (NICE, December 2005, 2008). However, the reports from the Citizens Council vary in the level of prescriptive recommendations they make, often airing many different views, and 'scoping' the potentially relevant issues (e.g. clinical need), rather than drawing any strong views which may directly help NICE (e.g. report on 'rule of rescue').

Space constraints do not allow coverage of the NICE policy position in this paper. Whilst providing a very useful context for the acknowledgement of social value judgments in HTA, and the efforts of NICE to be explicit need to be applauded, the content of the SVJ document does not add greatly to the debate here – other than making clear that a range of factors are not appropriate criteria for the setting of priorities (in HTA recommendations) unless there is a clear clinical / effectiveness basis for doing so. Please consult the SVJ report.

NICE (2009) has issued supplementary advice (in addition to general guidance on methods, and the social value judgements report) to its appraisal committees on 'life-extending end of life treatments', for indications affecting small numbers of patients with incurable illnesses, and where no alternative treatment with comparable benefits is available through the NHS. In such situations NICE appraisal committees are able to consider giving greater weight to QALYs achieved by such treatments.

3.5 Empirical Equity/Social values: Summary of observations from the literature review

If a health policy maker wanted to use the literature to inform on weightings for health outcomes (e.g. QALYs), the evidence base is not able to inform on such weightings. The literature can provide some helpful insights on what 'other relevant factors [criteria]' may be in a priority setting context for health care. However, the evidence is conflicting in many cases, has used simple methods, and is subject to various forms of bias. The evidence base available is not (I would argue) robust enough to make policy suggestions. However, in some cases the evidence base, although wide and varied, is able to put across a clear, albeit general, policy message. This is the case for arguments against the maximisation of health as a dominant decision-making objective. The social value of a health intervention is not generally related to the characteristics of eligible patient groups, i.e. limited support for setting priorities on the basis of age, gender, social class, ethnicity, social roles. The literature also clearly indicates that the level of benefit (health improvement) is an important social consideration, that efficiency is an important social consideration, and that other factors may impact on the preferred distribution of health care. Potential prime candidates for 'other factors', may be the 'severity of the current (pre-treatment) health state or condition' and the reduction of health inequalities (around socio economic status).

Although specifically identified by NICE, the relative importance of life-extending end-of-life interventions has not been the subject of empirical equity studies (not identified in current search)⁷. A further gap in the evidence base is on 'innovation' (innovative attributes of health interventions), and although this is specifically mentioned as part of the VBP consultation document, this review has not identified empirical studies that have sought to address the relative importance of innovation (vis-a-vis efficiency, or other considerations). This is also the case for the area highlighted in the VBP consultation document on 'wider societal benefits', as there is specifically no reference in the empirical equity literature to the relative importance of factors such as 'impact on carers', vis-a-vis other considerations. However, some of the evidence on inequalities, or patient characteristics may inform in this broad area of 'wider societal benefits'.

A major weakness in the evidence base is that the vast majority of studies have looked at specific social values in isolation e.g. severity versus efficiency, or age versus efficiency, and there is a need to undertake empirical work looking at the relationship between key social values. This is a common recommendation in three previous literature reviews in this area (Sassi *et al* 2001, Schwappach 2002, Dolan *et al* 2005). Although studies that ask respondents to weigh up a number of different issues at the same time are rare, the growth in choice experiments (stated preference DCE) has provided a number of studies seeking to highlight the relative importance of competing social values.

4. DISCUSSION

The case is made above that the current evidence base on empirical equity/ethics is not able to inform on the proposed framework for VBP, specifically in the area of need in VBP around "weightings to the benefits provided by new medicines", and the development of a price threshold structure (both central elements of the VBP framework).⁸ Where are the values for VBP to be derived from? (*Question for HESG*) Is the VBP proposal a feasible one, given the difficulties faced in being specific over the objectives of the NHS, and in determining plausible weightings for relevant priority setting criteria? (*Question for HESG*) In their recent discussion paper (on VBP) Dixon *et al* (2011) close by stating that "... *the immediate challenge [for VBP] will involve identifying methods to assess the social values in a robust manner; twenty years of research has not produced a preferred method, yet we have less than three [years] to derive a suitable set of weights*" (p17). With this challenge in mind, I ask the question of health economics – can it help? Specifically, given the emphasis placed / retained on cost-effectiveness analysis, I ask can economic evaluation help with the challenges faced by VBP?

Whilst I believe that health economics can help with difficult allocation decisions, I do not believe that it is the place of economic evaluation (a cost per QALY, cost per life-year, framework), as currently practiced, to answer the broader question on the social value of health care.

My contention is that economic evaluation is about efficiency, it is about 'production efficiency'. My view is that unless the decision maker objectives are framed (clearly stated) in a way that allows factors other than efficiency to be addressed as part of an economic evaluation then economic evaluation, as it is currently practised, retains the maximand of health gain, and provides one input, on relative efficiency in production of health gain, to the broader allocative (efficiency) deliberations. Such allocative deliberations will, at the present time, remain the remit of the informed committee (decision-making forum), with inputs from other areas, on other factors, relevant to the decision context. That the deliberative committee (or similar) has difficulties exercising, and communicating such a role is a different topic for discussion.

⁷ One of the HESG papers – Baker *et al* – addresses this topic.

⁸ Even if social values were available, the next consideration would be how much should the preferences of the general public (collected in experimental surveys) influence the policy framework?

On this interpretation – with the working objective of maximisation of health gain putting health economics (use of economic evaluation) in the domain of *production efficiency*, and removing it from deliberations on *consumption/allocative efficiency*⁹ (this being the landscape characterised by social values) – it can be argued that economic evaluation is outside of the framework that seeks to capture the social value of health, at least when economic evaluation is seen to be used alone (and/or in a dominant way). Indeed some economists currently contend that social values are not arguments in the social welfare function (Binmore, 2010).

In considering what it is that economic evaluation seeks to provide (does provide) as commonly practiced, it can be simply characterised as maximising units of health gain (e.g. QALY) subject to constraints. This can be presented as an 'economic' consideration, with options appraised as either 'economic' (e.g. cost-effective) or 'uneconomic'. Whereas consideration of 'other relevant factors', other than health gain, is (as currently practised) outside of the framework for optimising inputs (i.e. if objectives cannot be addressed as part of the optimisation problem), and as such these can be regarded as 'non-economic' considerations.¹⁰ Typically, when an activity is judged 'uneconomic' it is judged as wasteful, or using another term does not represent value for money. For example, more health gain can be obtained from use of resources i.e. better value for money can be obtained elsewhere. However, this is typically about efficiency in the use of inputs (e.g. investment of funds), and efficiency is only one input to many important decisions. Indeed, society (or an individual, or group) may wish to undertake an activity for 'non-economic' reasons – social, aesthetic, moral or political – but this does not alter the analysis that the activity (action) is 'uneconomic', when judged against typical criteria.

With such a view of economic evaluation, as set entirely in the context of efficiency in production, it provides the decision maker with information on the production possibilities, yet the decision maker (with multiple potential and conflicting objectives) needs to decide which of the production possibilities to undertake (which health interventions will be funded). Such decisions will typically be based on economic considerations in some scenarios, but on both economic and non-economic considerations in other scenarios. In this latter scenario, other 'non economic' value judgements are needed. In a theoretical world, of welfare economics, the social welfare function may be an analytical approach to help in the decision on consumption efficiency (allocative efficiency). However, at a practical level that is not an option. At a practical level the economic evaluation of health technologies (e.g. cost per QALY), as currently practiced, contributes to the deliberations needed on the social (consumption) value of health care through the provision of information on the relative efficiency of competing activities.

Rehearsing this argument above, on a hopefully friendly HESG readership/audience, I would argue that whilst economics can help with difficult resource allocation decisions (e.g. via multi-attribute decision framework), economic evaluation – as currently practiced in health care – is not the vehicle for assessing the broader value associated with a health intervention.

I suspect (very much hope) there may be some discussion on this view, as I am no longer surprised (I was once) when many academic health economists look at me with some 'concern' and fervently disagree when this statement is made¹¹. For those who would argue that an economic evaluation, presenting a cost per QALY estimate, informs on factors other than efficiency, I would ask how? Producing life years from health care does not offer any insight on the consumption value of those life years. Nor does the use of a weight to adjust a life-year, to a quality-adjusted life-year, offer any insight on the consumption value of the QALY. The QALY is merely a means of addressing a concern, not held by all, that not all life-years are equal in terms of productive output. The QALY weight is a technical adjustment, using a benchmark

⁹ Other than through provision of information on the most efficient uses of health care resources (relative to comparator, or a general benchmark [threshold]).

¹⁰ This terminology was used by EF Schumacher in his discussion of the role of economics; 1973, Ch.3.

¹¹ That is – cost per QALY/cost per life-year estimates are only informing on efficiency, i.e. the cost of producing an extra unit of output.

(upper value) of 1.00 unit of production to reflect a life year equivalent to full health (best possible unit of production, using such a metric). This life year may not be in full health, it may involve impairments to health, but it is considered, in such circumstances that the health impairments/ problems do not merit the sacrifice of future life expectancy [in a TTO task] or the acceptance of a risk of death [SG task] that will on average lead to a loss of life expectancy [expected value]. This may be regarded as a stylised scenario, but it is one which prevails in the central conduct of health technology appraisal in the UK (NICE 2008). For example, in many economic evaluations using the EQ-5D there are a proportion of participants with a health state description showing no problems on the 5 dimensions of health status, and subsequently a health state value of 1.00 (top score), however it is likely (and shown in many studies) that such participants would have measures of health status that indicated health problems/impairments. Is it such characteristics of the QALY that introduce some consideration of consumption benefits/efficiency? ... I would argue that is not the case.

This has been a bit of an explorative, potentially dangerous, and informal discussion. However (as space runs out) I would appreciate it if the HESG audience would give some thought to the latter points here – i.e. is economic evaluation able to address (the best way to address) choices on allocative/consumption efficiency?

The main substance of the paper has been on trying to show that the empirical evidence is not available to support the proposed move to VBP. There are other broader questions on VBP, such as how NICE can appraise all new branded medicines, and how the benefits foregone (opportunity cost) can be set out in an identifiable way, so that there is appropriate consideration of other relevant factors on that side of the equation. Furthermore, there are wider considerations on the way that innovation could/should be address by the Government. But, on the weightings for health benefits – in the first instance – where will the values come from? How will such values be collected? is it to be a pragmatic 'decision-maker knows best' approach!

REFERENCES (short list; other listing of review references available on request, and at HESG session)

- Binmore K. Social norms or social preferences? *Mind & Society* 9:139-157; 2010.
- Bleichrodt H, Doctor J, Stolk E. A nonparametric elicitation of the equity-efficiency trade-off in cost-utility analysis. *Journal of Health Economics* 2005;24(4):655-78.
- Bowling A. Health care rationing: the public's debate. *British Medical Journal* 1996; 16;312(7032):670-4.
- Bryan S, Williams I, Mclver S. Seeing the NICE side of cost-effectiveness analysis: A qualitative investigation of the use of CEA in NICE technology appraisals. *Health Economics* 2006;16(2):179-93.
- Claxton K, Sculpher M, Carroll S. Value-based pricing for pharmaceuticals: Its role. Specification and prospects in a newly devolved NHS. CHE Research Paper 60, University of York, 2011
- Culyer AJ. Equity - Some theory and its policy implications. *J Med Ethics* 2001; 27(4):275-83.
- Daniels N, Sabin J. The ethics of accountability in managed care reform. *Health Affairs* 1998;17:50-64.
- Department of Health. A new value-based approach to the pricing of branded medicines. A consultation. 2010 (Dept of Health, Medicines, Pharmacy and Industry Group)
- Dixon S, Tsuchiya A, Brazier J, et al. Proceed with caution: an economic perspective on the UK's value based pricing proposals. HEDS (SchARR) Discussion Paper 11/004, University of Sheffield 2011.
- Dolan P, Tsuchiya A. Health priorities and public preferences: the relative importance of past health experience and future health prospects. *Journal of Health Economics* 2005;24(4):703-14.
- Dolan P, Shaw R, Tsuchiya A, Williams A. QALY maximisation and people's preferences: A methodological review of the literature. *Health Economics* 2005;14(2):197-208.
- Hauck K, Smith PC, Goddard M. *The economics of priority setting for health care. A Literature review.* Washington, USA: Health Nutrition and Population (NHP) Discussion Paper, The World Bank; 2003.
- NICE. Appraising life-extending, end of life treatments. 2009. <http://www.nice.org.uk/>
- NICE. *Social value judgments. Principles for the development of NICE guidance.* 2nd Edition. London: National Institute for Health and Clinical Excellence (www.nice.org.uk) 2008.
- NICE. Guide to the methods of technology appraisal. 2008. <http://www.nice.org.uk/>
- Nord E, Pinto JL, Richardson J, Menzel P, Ubel P. Incorporating societal concerns for fairness in numerical valuations of health programmes. *Health Economics* 1999;8(1):25-39.
- Rawlins MD, Culyer AJ. National Institute for Clinical Excellence and its value judgements. *British Medical Journal* 2004;329:224-6.
- Rawls J. *A Theory of Justice.* New York: Oxford University Press; 1971.
- Richardson & McKie 2005 Soc Sci Med
- Sassi F, Le GJ, Archard L. Equity versus efficiency: A dilemma for the NHS. *British Medical Journal* 2001;323(7316):762-3.
- Sassi F. Setting priorities for the evaluation of health interventions: when theory does not meet practice. *Health Policy* 2003;63(2):141-54.
- Schumacher EF *Small is Beautiful*, 1973.
- Schwappach DL. Resource allocation, social values and the QALY: a review of the debate and empirical evidence. *Health Expectations* 2002;5(3):210-22.
- Schwappach DL, Strasmann TJ. "Quick and dirty numbers"? The reliability of a stated-preference technique for the measurement of preferences for resource allocation. *Journal of Health Economics* 2006;25(3):432-48.
- Shah K. Severity of illness and priority-setting in health care: A review of the literature. *H Policy.* 2009
- Wagstaff A. QALYs and the equity-efficiency trade-off. *Journal of Health Economics* 1991; 10(1):21-41.
- Williams A. Ethics and efficiency in the provision of health care. In: Bell M, Mendus S, editors. *Philosophy and Medical Welfare.* Cambridge: Cambridge University Press; 1988.
- Williams A. QALYs and ethics: a health economist's perspective. *Social Science and Medicine* 1996;43(12):1795-804.
- Williams I, Bryan S. Understanding the limited impact of economic evaluation in health care resource allocation: A conceptual framework. *Health Policy* 2007;80(1):135-143.
- Wilmot S, Ratcliffe J, Allen C. How well do members of the public deal with a distributive justice problem in health care? *Journal of Health Services Research and Policy* 2004;9(1):7-13.

Table 2: Summary Characteristics of the Empirical Studies Included in the Review

Authors	Pub. Year	Approach*	Sample Size	Country of Study	Design*	Sample	QA Methods
Abellan-Perpignan & Prades	1999	EXP	78	Spain	SC	CONV (students)	Low
Anderson & Lytkens	1999	EXP	225	Sweden	SC	CONV (students)	Low
Annand & Wailoo	2000	ATT	118	UK	SC	RAND (public)	Moderate
Baker <i>et al</i>	2010	EXP	587 / 409	UK	INT	RAND / CONV (public)	
Baltussen <i>et al</i>	2006	EXP	30	Ghana	SC	CONV (decision-makers)	Low
Baron & Ubel	2001	EXP	70	USA	SC (online)	CONV (public)	Low
Battista <i>et al</i>	1995	ATT	55	Canada	SC	CONV (mixed)	Low
Beach <i>et al</i>	2003	EXP	781	USA	SC	CONV (public)	Low
Bjork & Rosen	1993	EXP / ATT	60	Sweden	INT	CONV (politicians)	Moderate
Bleichrodt <i>et al</i>	2005	EXP	65 / 179	Netherlands	INT/SC	CONV (students/public)	Moderate
Block <i>et al</i>	2001	ATT	122	Mexico/South America	INT/GRP	CONV (public)	Moderate
Bowling	1996	ATT	2005	UK	INT	RAND (public)	High
Bowling <i>et al</i>	2002	ATT	337/242	UK	INT	RAND (public)	High
Brennan <i>et al</i>	2008	ATT	191	Australia	SC	RAND (Dentists)	Moderate
Browning & Thomas	2001	EXP / ATT	238	Australia	SC	RAND/CONV (mixed)	Moderate
Bryan <i>et al</i>	2002	EXP	909	UK	INT	RAND (public)	High
Charny <i>et al</i>	1989	EXP	719	UK	INT	RAND (public)	High
Choudry <i>et al</i>	1997	EXP / ATT	80	Canada	SC/PQ	CONV (Healthcare)	Low
Cookson & Dolan	1999	EXP / ATT	60	UK	SC /GRP INT	RAND (public)	High/Moderate
Cropper <i>et al</i>	1994	EXP	1000 / 564 / 1000	USA	TEL INT (SURV)	RAND (public)	High/Moderate
Dicker & Armstrong	1995	ATT	16	UK	INT	CONV (public/patients)	Moderate
Dolan & Cookson	2000	EXP / ATT	60	UK	GRP INT (focus group)	RAND (public)	High/Moderate
Dolan & Green	1998	EXP	28	UK	INT	CONV (public)	Moderate
Dolan & Robinson	2001	EXP	71	UK	SC (GRP/INT)	CONV (students)	Low
Dolan & Shaw	2003	ATT	23	UK	SC (GRP)	PURPOSIVE (public)	Low
Dolan & Shaw	2004	ATT	23	UK	INT (GRP)	PURPOSIVE (public)	Moderate
Dolan & Tsuchiya	2005	EXP	128	UK	SC (GRP)	RAND (public)	Moderate
Dolan <i>et al</i>	1999	ATT / EXP	60	UK	SC (GRP)	RAND (public)	Moderate
Edwards <i>et al</i>	2003	EXP	1101	UK	SC	RAND (public/clinical/dec-makers)	Moderate
Emmelin <i>et al</i>	1999	ATT / EXP	451	Sweden	SC/PQ	CONV (politicians)	Low
Farrar <i>et al</i>	2000	EXP	130	UK	SC	CONV (healthcare)	Low
Fowler <i>et al</i>	1994	ATT	206	USA	INT (TEL)	RAND (public)	High

- Work in progress: Not for citation without authors' permission -

Authors	Pub. Year	Approach*	Sample Size	Country of Study	Design*	Sample	QA Methods
Furnham <i>et al</i>	2000	EXP	107 / 24	UK	SC	CONV (mixed / students)	Low
Gallego <i>et al</i>	2007	ATT / EXP	200	Australia	INT (strt)	CONV (public)	Moderate
Gerard	2005	EXP	798	UK	SC	CONV (public)	Low
Green	2009	EXP	261	UK	INT	RAND (public)	High
Green & Gerard	2009	EXP	259	UK	INT	RAND (public)	High
Green <i>et al</i>	2000	ATT	100 / 101	USA	SC	CONV (public / politicians)	Low
Gyrd-Hansen	2004	EXP	3201	Denmark	INT	RAND (public)	High
Gyrd-Hansen & Slothuus	2002	EXP	1895	Denmark	INT	RAND (public)	High
Gyrd-Hansen & Sogaard	2001	EXP	750	Denmark	INT	RAND (public)	High
Gyrd-Hansen & Kristiansen	2008	EXP	2923	Denmark	INT	RAND (public)	High
Hasman <i>et al</i>	2008	ATT	22	UK	INT	CONV (decision makers)	Moderate
Holmes	1997	EXP	72	USA	SC	CONV	Low
Jehu-Appiah <i>et al</i>	2008	EXP	63	Ghana	SC (GRP)	CONV (decision-makers)	Low
Johannesson & Gerdtham	1996	EXP	80	Sweden	INT (GRP)	CONV (students)	Moderate
Johannesson & Johansson	1996	EXP	1000 / 2000	Sweden	SC	RAND (public)	Moderate High
Johannesson & Johansson	1997	EXP	780	Sweden	INT (TEL)	RAND (public)	High
Johri <i>et al</i>	2005	EXP	160	USA	SC	CONV (public)	Low
Kinnunen <i>et al</i>	1998	ATT / EXP	1178/682/837/1133	Finland	SC	RAND (public / nursing / Drs) CONV (politicians)	Moderate/Low
Kluge & Tomasson	2002	ATT	5	Canda	INT	CONV (healthcare)	Moderate
Kuder & Roeder	1995	ATT / EXP	46	USA	INT (GRP)	CONV (public)	Moderate
Lees <i>et al</i>	2002	ATT	1004 / 357	UK	SC	CONV (public / clinicians)	Low
Lewis & Charny	1989	EXP	721	UK	INT	RAND (public)	High
Lian	2001	ATT	152	Norway	SC	CONV (clinicians)	Low
Linblad <i>et al</i>	2002	ATT	22	Sweden	INT	CONV (patients)	Moderate
Lindholm & Rosen	1998	EXP	449	Sweden	SC	CONV (politicians)	Low
Lindholm <i>et al</i>	1997	EXP	449	Sweden	SC	CONV (politicians)	Low
Mariotto <i>et al</i>	1999	ATT	443 / 189	Italy	INT	RAND (elderly public) / CONV (nurse/aide)	High / Low
Menon & Stafinski	2008	EXP / ATT	16	Canda	INT (citizens jury)	RAND (public)	High
Mooney <i>et al</i>	1995	EXP / ATT	283	Australia	SC	CONV (decision makers)	Low
Mullet <i>et al</i>	1996	ATT	6	Canada	SC	CONV (clinicians)	Low
Myllykangas <i>et al</i>	2003	ATT		Finland	SC	RAND (public / nursing / Drs)	Moderate/Low

- Work in progress: Not for citation without authors' permission -

Authors	Pub. Year	Approach*	Sample Size	Country of Study	Design*	Sample	QA Methods
			1178/682/837/1133			CONV (politicians)	
Neuberger <i>et al</i>	1998	EXP / ATT	1000 / 200 / 100	UK	INT / INT / SC	QUOTA (pub / Drs / specialist Dr)	Moderate
Nord	1993	EXP	61 / 25	Norway	SC	CONV (public)	Low
Nord	1993	EXP	10	Norway	INT	CONV (healthcare)	Moderate
Nord	1995	EXP	14 / 53 / 52	Norway	SC	CONV	Low
Nord	1995	EXP	8 seminars (10-13)	Norway	SC	CONV	Low
Nord	2009	EXP					
Nord <i>et al</i>	1995	ATT	551	Australia	SC	CONV (partial quota sampling)	Low
Nord <i>et al</i>	1995	EXP / ATT	119	Australia	INT	CONV (public)	Moderate
Nord <i>et al</i>	1996	ATT / EXP	551/44/42	Australia	SC / INT / INT	CONV (public)	Low / Moderate
						RAND (4 of 7 groups) - Phys/Pols/public/public	Moderate
Oddsson	2003	EXP / ATT	913	Iceland	SC		
Oliver	2004	EXP	25	UK	SC	CONV (university)	Low
Oliver	2009	EXP					
Olsen	1994	EXP	134	Norway	SC	CONV (students/clinicians)	Low
Ottersen <i>et al</i>	2008	ATT	63	Tanzania	SC (Grp)	CONV (decision makers)	Low
Pernerger <i>et al</i>	2002	EXP	1170	Switzerland	SC	Pop'n of physicians	Moderate
Ratcliffe	2000	EXP	303	UK	SC	CONV (mixed)	Low
Roberts <i>et al</i>	1999	EXP	91	UK	INT	CONV (public)	Moderate
Rodriquez-Miguez <i>et al</i>	2000	EXP	61	Spain	SC / GRP	CONV (students)	Low
Rodriquez-Miguez <i>et al</i>	2002	EXP	45	Spain	SC / GRP	CONV (students)	Low
Rosen & Karlberg	2002	ATT	1194 / 427	Sweden	SC	RAND (public / decision-makers) RAND (pub / nursing / physicians);	Moderate Moderate/Low
Ryynanen <i>et al</i>	1999	ATT	1156/667/803/1096	Finland	SC	CONV (pol)	
Schwappach	2002	EXP	127	Germany	SC	CONV (public)	Low
Schwappach	2003	EXP	154	Germany	SC (online)	CONV (students)	Low
Schwappach	2005	EXP	1253	Germany	SC (online)	CONV (public)	Low
Schwappach & Strasmann	2006	EXP	716	Germany	SC (online)	CONV (public)	Low
Shmueli	2000	ATT	2030	Israel	INT	RAND (public)	High
Shmueli	1999	EXP	2006	Israel	INT	RAND (public)	High
Skitka & Tetlock	1992	EXP / ATT	198 / 37	USA	SC	CONV (students)	Low
Stolk <i>et al</i>	2005	EXP	41	Netherlands	SC	CONV (students/healthcare)	Low
Stronks <i>et al</i>	1997	ATT / EXP	45	Netherlands	INT (GRP)	CONV (mixed)	Moderate
Tappenden <i>et al</i>	2006	EXP	37	UK	SC	CONV (decision-makers)	Low

- Work in progress: Not for citation without authors' permission -

Authors	Pub. Year	Approach*	Sample Size	Country of Study	Design*	Sample	QA Methods
Tsuchiya & Dolan	2007	EXP	271/220	UK	SC	RAND/CONV (public/clinicians)	Moderate
Tsuchiya & Dolan	2009	EXP	123	UK	SC (GRP context)	RAND (purposive)	Moderate
Tsuchiya <i>et al</i>	2003	EXP / ATT	140	UK	INT	RAND (purposive)	High
Tsuchiya <i>et al</i>	2005	ATT	87	UK	SC	CONV (public)	Low
Ubel	1999	EXP	479	USA	SC	CONV (public)	Low
Ubel and Lowenstein	1995	EXP	138	USA	SC	CONV (public)	Low
Ubel and Lowenstein	1996	EXP	169	USA	SC	CONV (public)	Low
Ubel <i>et al</i>	1996	EXP	568 / 74 / 73	USA	SC	CONV (public/ethicists/dec-makers)	Low
Ubel <i>et al</i>	1998	EXP	289	USA	SC	CONV (public)	Low
Ubel <i>et al</i>	1999	EXP	241 / 66 / 74	USA	SC	CONV (public)	Low
Ubel <i>et al</i>	2001	EXP/ATT	408	USA	SC	CONV (public)	Low
Ubel <i>et al</i>	2001	EXP	615 / 68	USA	SC	CONV (public)	Low
Valtonen	2009	EXP	498	Finland	SC	RAND (healthcare)	Moderate
van Busschbach <i>et al</i>	1993	EXP	30 / 47	Netherlands	INT	CONV (students / public)	Moderate
Varekamp <i>et al</i>	1998	ATT	31	Netherlands	INT	CONV (healthcare)	Moderate
Wailoo & Annand	2005	ATT	118	UK	SC/PQ	RAND (public)	Moderate
Wilmot <i>et al</i>	2004	EXP/ATT	22	UK	GRP/INT	CONV (public)	Moderate
Williams	1988	ATT	N/A	UK	SC	CONV (mixed)	Low
Wiseman <i>et al</i>	2003	ATT	373	Australia	SC	CONV (public/patient)	Low
Wiseman	2005	ATT/EXP	373 / 43	Australia	SC	CONV (public/patient/healthcare)	Low
Woolhead <i>et al</i>	2002	ATT	25	UK	INT	Selected patients	Moderate
Zweibel <i>et al</i>	1993	EXP / ATT	505	USA	INT (TEL)	RAND (public)	High

* **KEY:**

Approach: EXP = experimental, ATT = Attitudinal

Design: SC = self complete, PQ = postal questionnaire, INT = interview, GRP = group, TEL = telephone

APPENDIX – SEARCH STRATEGY

1. Search terms:

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

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

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

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

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

2. SEARCH: *Inclusion / exclusion criteria*

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

1. *Excludes:*

- References purely on patient-level clinical ethics were excluded.
- References comprising of editorial or commentary style contributions were excluded.

2. *Health context:*

- Papers were sought which were related to the area of health and health care (e.g. not issues of criminal justice)
- Papers were sought which were able to inform on the areas of equity and resource allocation decisions within health, health care and general health policy.

3. *Content:*

- Following the above general sift against criteria 1 and 2 (above), references had to present empirical findings of relevance to the analysis/consideration of equity/fairness (i.e. empirical insights to the dimensions of equity, however defined, and/or trade-off between relevant dimensions of equity).

Given the methodological nature of the literature search it was necessary to make a judgment on inclusion of studies on a case-by-case basis. Also, given the methodological nature of the review, and the judgements needed regarding inclusion/exclusion, it is accepted that the search will not be exhaustive, and that there may be references missing that should/could have been included (please advise on any known omissions).