MONASH University
Business and Economics

Neglected Equity Issues in Cost Effectiveness Analysis – Part 1:

Severity of Pre-Treatment Condition, Realisation of Potential for Health, Concentration and Dispersion of Health Benefits, and Age-Related Social Preferences

Dr John McKie

Research Fellow, Centre for Health Economics,
Monash University

Professor Jeff Richardson

Foundation Director, Centre for Health Economics,
Monash University

July, 2005

TABLE OF CONTENTS

BOD	burden of disease
CA	conjoint analysis
CEA	cost-effectiveness analysis
CUA	cost-utility analysis
DALY	disability-adjusted life year
EUT	expected utility theory
HRQoL	health-related quality of life
NICU	neo-natal intensive care unit
PTO	person trade-off
PYLL	potential years of life lost
QoL	quality of life
QALY	quality-adjusted life year
RR	rule of rescue
RS	rating scale
STTO	social time trade-off
SG	standard gamble
SWTP	social willingness to pay
TTO	time trade-off
VAS	visual analogue scale
WHO	World Health Organization
WTP	willingness to pay

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

BACKGROUND

Economic evaluations and policy recommendations in the health area often give the impression that the single relevant criterion for ranking programmes and services is the maximisation of health gains within resource constraints. However, recent studies suggest that the general public, in Australia and elsewhere, is willing to sacrifice aggregate health gains in order to achieve a range of other social objectives. The significance of this is reflected in the recent WHO evaluation of national health systems (World Health Organization 2000), which gave an importance weighting of 0.25 to the promotion of health per se and 0.75 to the achievement of social justice as defined by the distribution of health care costs, system responsiveness, and equity of access to health services. In an earlier study of the global burden of disease the WHO adopted a number of equally controversial assumptions (Murray and Lopez 1996): for example, that the social value of a year of life varies with the age of a person, that the burden arising from reduced quality of life should not be revised downwards as individual adaptation to illness and disability occurs, and that a standard (not actual) life expectancy should be adopted in all countries for the calculation of potential years of life lost (PYLL). The WHO team is presently expanding its enquiry into health-related social objectives in a number of countries (Murray, Salomon et al. 2000).

The assumptions underlying both studies (Murray and Lopez 1996; World Health Organization 2000) have been vigorously debated (Richardson, Wildman et al. 2003). Nevertheless, the WHO has succeeded in focusing the attention of researchers and policy makers on a range of issues that have previously been ignored, and has "legitimised" the investigation of issues relating to social justice in the health sector. Thus, for example, publication of *The World Health Report 2000* led immediately to two government workshops in Australia (2000, 2001) and the examination of the need for including responsiveness data in annual data collections. Similarly, *The Global Burden of Disease* study resulted in both a Commonwealth and a Victorian state government re-estimation of the BOD (burden of disease) on the basis of alternative methodologies (Mathers, Vos et al. 1999; Victorian Government Department of Human Services 1999). This focus on community values can be expected to assume increasing importance as the effects of new technology and ageing inevitably increase the pressure on the health budget.

Rationing already exists. This is explicit in the decisions made to list drugs and medical services in the national health scheme, and it is implicit in the limitations placed upon the number of hospital beds and the supply of medical practitioners. It is commonly recognised that rationing is best done using the economic framework of costs and benefits. At present, however, this framework is incomplete. There is compelling evidence that public expectations and values concerning health and health care have not been properly analysed at the level of either the health system or the individual health programme. Social values are neither understood fully nor systematically incorporated in social policy or economic methods. A framework that can meet the need for on-going re-evaluation of the benefits and costs of medical services, and which can incorporate public values in a substantial rather than rhetorical way, is needed if policy recommendations in the health sector are to be seen as fair. The following overview of the literature reveals that for a range of issues there is a discordance between population values and those currently assumed in health economic evaluations and public policy, and indicates where further research needs to be undertaken.

AIMS

The purpose of the review is to provide an overview and evaluation of the empirical evidence concerning social preferences in the health area, concentrating on issues that have, to date, been overlooked in health policy and, in particular, in the theory and practice of economic evaluation. The studies included demonstrate (with varying degrees of rigor) that for a range of issues, such as severity of illness, potential for health improvement, and the treatment of costs, there is a divergence between community values and those assumed in health economic evaluations.

The studies vary from high quality representative samples through to opportunistic sampling of unrepresentative groups. Likewise, some studies attempt to quantify the social preferences elicited, whereas others merely indicate the direction of the preferences. We have divided the issues discussed in these studies into three categories. Category A comprises two issues for which the empirical evidence concerning community preferences is persuasive and growing: severity of pre-treatment condition and realisation of potential for health. Category B comprises two issues for which the empirical evidence is suggestive but inconclusive: concentration and dispersion of individual health benefits and age. Category C comprises several issues for which the evidence concerning community preferences is at present only weak. This does not mean that the issues in this category are irrelevant. Rather, it is an indication that evidence is presently limited, and that more work needs to be done in this area.

METHODS

The review does not claim to be systematic or exhaustive. The selection of issues was based first and foremost on potential policy relevance. For most issues a table is provided in the text giving the country in which the study was conducted, the number of subjects involved, the type of subjects, and the elicitation methods used. Issues are divided into categories A, B and C based on the number of studies, and the quality of the data in those studies with respect to reliability, validity, and framing effects. Other commentators may assess the evidence differently and re-categorise some issues, and we would expect issues to migrate from one category to another as further evidence becomes available.

RESULTS AND RECOMMENDATIONS FOR FUTURE RESEARCH

Two issues are placed in category A:

Severity of Pre-treatment Condition. No importance is attached to severity of pre-treatment condition per se in cost-effectiveness analyses that use unweighted QALYs (quality-adjusted life years) as the unit of effectiveness. However, the evidence from a number of countries indicates that people think the more severely ill should have priority over the less severely ill (other things being equal). While the evidence concerning the social importance of severity of pre-treatment condition is compelling, more work is needed in this area. Only one study has attempted to measure the preferences of the Australian public for the more severely ill. And as Nord, Richardson et al. comment, subjects' responses in this study, which is now over ten years old, were highly dispersed, "indicating the likelihood of a high sampling error for the median values" (Nord, Richardson et al. 1993, p. 467). Furthermore, responses "were elicited from small samples … and the response rates were low" (Nord, Richardson et al. 1993, p. 472). Finally, the survey was conducted among a convenience sample of students and nurses not the general population. Consequently, more reliable data is needed before the quantitative importance of this preference can be determined for health policy.

Realisation of Potential for Health. There is also extensive evidence from Australia and overseas indicating that people reject discrimination on the basis of disability. In other words, the potential to benefit from treatment carries less weight for many people than is implied by conventional CEA. When asked to judge for others - when asked to adopt the perspective of a health administrator, for example - people are often prepared to sacrifice health gains to ensure that the chronically ill and permanently disabled are afforded the opportunity to realise their *potential* for health, even if this is not great. Again, future research needs to concentrate on measuring how much the Australian community is willing to sacrifice in terms of health production to ensure that the permanently disabled and chronically ill are afforded the opportunity to realise their potential for health. Achieving this is complicated by the fact that not just one "equity-efficiency" trade-off is needed, but a number of them, which differ according to whether life-saving or health-improving interventions are in question, on whether the target group *is* disabled or *will be left* disabled after treatment, and so on. No studies have been undertaken that attempt to precisely measure the degree of support for such non-discrimination in the Australian population, although Nord, Richardson and colleagues conducted a study in which it was observed (Nord, Richardson et al. 1995a).

Two issues are in category B:

Concentration and Dispersion of Individual Health Benefits. Within the context of conventional CEA, using unweighted QALYs as the unit of effectiveness, it does not matter whether the aggregate gain from a health service is distributed across many people or concentrated among a few. However, a number of studies suggest that health gains are valued less when they are concentrated among few people, and valued more when they are distributed more widely. Significantly, people are willing to "trade-off" efficiency to ensure that more people share in the benefits of health programmes and services. However, this is governed by a threshold effect: support for the dispersion of aggregate health gains declines when individual health benefits become too small (the "plateau effect"). In Australia, Nord, Street and colleagues measured the preference for concentration and dispersion of health gains both in the context of health-improving and life-saving treatment (Nord, Street et al. 1996). However, there was a low response rate to the survey, and the authors comment on the possibility of bias due to self-selection of respondents. Also, the authors did not investigate the plateau effect – the idea that people prefer to concentrate health gains among few people if dispersing them more widely will lead to many people gaining health benefits deemed too small. In general, good empirical evidence is starting to accumulate on this issue, though more research is needed.

Age. Many studies indicate that people are willing to give certain age groups priority in the competition for limited health care. Some of these studies measure "utilitarian ageism": the preference for the treatment of younger patients because they are expected to live longer after treatment. Other studies measure "productivity ageism": the preference for some age groups (young adults and the middle-aged, in particular) because of their level of productivity at work or home. Yet other studies measure "egalitarian ageism": the preference for younger patients because of an aversion to inequality in age of death. The latter is similar to the "fair innings" argument, according to which there is a social expectation for the achievement of a fair minimum length of life (Williams 1997). While utilitarian ageism is compatible with health maximisation, productivity ageism and egalitarian ageism imply a rejection of the strategy of ranking programmes and services strictly on the basis of health gains. Nord, Street and colleagues carefully distinguished between utilitarian ageism and egalitarian ageism and detected both in the Australian population (Nord, Street et al. 1996).

However, most studies fail to distinguish between utilitarian ageism, productivity ageism and egalitarian ageism, and more empirical work is therefore needed on age preferences.

Six issues are placed in category C:

High Cost Illnesses. Studies in Australia and overseas suggest that many people reject the usual treatment of costs in economic evaluations in the health sector. Allocative (but not productive) efficiency redistributes services between patients, and the public is reluctant to discriminate against those with high-cost illnesses. For example, a study by Nord, Richardson and colleagues found that respondents in Australia are averse to withholding treatment from those "who are unlucky enough to have high-cost illnesses" (Nord, Richardson et al. 1995b). In face-to-face interviews respondents who rejected the conventional role of costs were challenged, to ensure they understood the adverse consequences for patient health if higher cost treatments were selected, with no increase in health benefits and in the context of a fixed budget. This did not change the results. The participants in this study were over-represented by people with tertiary education, and the study was limited to one health state; namely, severe disability. Future research needs to look at the relevance of direct costs when withholding treatment will lead to death, or leave patients in a milder health state than severe disability – in brief, across a range of health states. In general, and perhaps surprisingly, there is a lack of reliable data on community preferences concerning the relevance of direct costs in economic evaluations in the health sector. Results have been assumed on the basis of economic theory.

Indirect Costs and Benefits. It has generally been accepted by economists that indirect costs and benefits should be included in the comparison of the overall costs and benefits of a programme or service. Presently there is debate about the appropriate way to do this: by adopting the orthodox human capital approach or the frictional cost method. However, the inclusion of indirect costs and benefits will systematically favour some groups, particularly working people with high incomes, and disadvantage other groups, such as the young, the elderly, and the mentally handicapped. A few studies in Australia and overseas suggest that people reject these implications of the inclusion of indirect costs and benefits in economic evaluations. In Australia, Nord, Richardson and colleagues found little support for the idea that people in the workforce should have priority over non-working people (Nord, Richardson et al. 1995b). Of course, this does not mean that all indirect costs and benefits should be excluded from economic evaluations. It is a task for future research to determine which indirect costs and benefits are "socially irrelevant" and which are not (Olsen and Richardson 1999).

Preservation of Hope. Some studies suggest that people are reluctant to abandon patients with illnesses and injuries that have only "cost-ineffective" therapies. That is, there is evidence that people are willing to sacrifice health gains in order to preserve the hope of treatment for everyone. This is borne out by some studies of social attitudes towards organ allocation. Ubel and Loewenstein found that people will allocate a supply of life-saving organs to second- and third-time recipients, even though first-time recipients have a better prognosis (Ubel and Loewenstein 1996a). They suggest that people want to preserve the hope of treatment even for "cost-ineffective" patients. At present, however, there is no reliable data on which to estimate the trade-off between health production and the preservation of hope, either in the case of health-improving or life-saving interventions.

The Rule of Rescue. Informal observation suggests that people feel a strong imperative to rescue identifiable individuals facing avoidable death, without giving too much thought to the opportunity cost of doing so. This "Rule of Rescue" is multi-factorial: it covers social preferences for interventions offered in life-threatening as opposed to non-life-threatening situations, for identifiable over statistical lives, and for interventions offered in "dramatic" or "unexpected" circumstances (such as a sailor lost at sea). Further research is needed to determine the significance of these unexplored factors. For example, it is unclear whether the Australian community attaches different values to interventions offered in "shocking" circumstances

compared to those offered in more mundane circumstances, or whether they exhibit a reflective, considered preference for identifiable over statistical lives.

Adaptation. A persuasive body of empirical research shows that different groups – patients, the general public, health professionals - assess the quality of life associated with various health states differently. Generally, disabled and chronically ill patients rate their own quality of life higher than do healthy people imagining themselves in these states, in part because the former adapt to their disability. This raises the question of whose preferences should be used in determining the "(dis)utility" of health states for the purposes of economic evaluation. If the non-adapted preferences of the general public are used in economic evaluations the gains from a health programme or service may appear greater than if the adapted preferences of patients are used, and in some cases the difference may be significant. The ethical arguments for using adapted and non-adapted preferences are analysed by Menzel, Dolan et al. (2002). At present, however, there is little evidence on the social significance of adaptation: that is, whether the general public thinks adapted or non-adapted preferences (or both) should be used in determining the "(dis)utility" of health states for the purposes of economic evaluation.

The ExAnte/Ex Post Distinction. Ex post evaluations of a health state are based on experience of the state: they are evaluations of the state *after* (or while) it is experienced. *Ex ante* evaluations are anticipatory: they are assessment of what a health state *will* be like. If it is thought that economic evaluations should be based on data from interventions that have actually taken place, then *ex post* judgements of the benefits of those interventions should be used. On the other hand, *ex ante* judgements are able to capture sources of utility - such as the elimination of uncertainty concerning the outcome of a treatment, or the avoidance of the possibility of regret over a treatment decision that turns out badly - that arguably should be included in economic evaluations. With the exception of some interesting work on regret theory and uncertainty aversion, there is little evidence on community attitudes towards the *ex ante/ex post* distinction, although some studies suggest that the adoption of an *ex ante* or *ex post* perspective might have potentially significant implications for economic evaluations of health programmes and services (Cook, Richardson et al. 1994; Smith 1996).

CONCLUSION

The studies included in this review suggest that a number of factors currently omitted from health programme evaluation have varying degrees of community support. Moreover, the neglect of these factors threatens the policy-relevance of economic evaluations, as well as their ethical acceptability. For example, confirmation of the importance of severity of pre-treatment condition, potential for health improvement, or the relevance of age, all have the potential to significantly alter the relative importance of the benefits of different programmes and services, and the likelihood of their inclusion in a national health scheme. Likewise, the treatment of indirect costs and benefits may alter the range of variables included in an economic evaluation, or the way in which they are included. The empirical evidence reviewed here indicates that people want a health system that is not only efficient, but equitable in the broadest sense - that promotes a sense of community, respects individual needs and choices, is responsive to community expectations, and so on. To achieve this, more qualitative and quantitative investigation of social objectives in the health sector is clearly needed.

INTRODUCTION

The Quality-Adjusted Life Year (QALY) has become increasingly popular in economic evaluations of health programmes and services. The QALY allows two of the most important consequences of health care - the prevention of premature death and the enhancement of quality of life - to be combined into a single unit of measurement.[1] In its simplest form the QALY represents a year of life that has been weighted by a number representing quality of life. Conventionally, full health has a weighting of 1 and death has a weighting of 0. For example, if a year of life on hospital dialysis is considered to be worth only 60 per cent as much as a year of normal health, other things being equal, then 20 years on dialysis would be equivalent to 20×0.60 = 12 QALYs.

Cost-effectiveness analysis (CEA) using QALYs as the unit of effectiveness provides a conceptually simple method for prioritising health programmes and services: all else equal, the lower the cost per QALY the greater the value for money offered by a programme or service, and hence the higher a priority it should be. As Weinstein and Stason comment (1977, p. 717): "alternative programs or services are then ranked, from the lowest value of this cost-per-effectiveness ratio to the highest, and selected from the top until available resources are exhausted." If "health" is defined by the number of QALYs attained then this simple cost-per-QALY prioritising rule maximises health for a given resource cost,[2] and the argument that health policy should maximise health has intuitive appeal.

However, this method of ranking programmes and services has been criticised on the ground that it discriminates against certain groups (Harris 1987; Harris 1988a). While it will ensure that QALYs are maximised, it will not ensure that those who are the most severely ill receive priority over those who are less severely ill (other things being equal), that the permanently disabled are allowed to realise their potential for health improvement (be this large or small), that the young have priority over the elderly, and so on. In brief, no value is assigned to any distributive or procedural concern beyond what is entailed by health maximisation.

There is a sense in which this form of CEA is egalitarian: programmes and services are ranked according to their potential to produce QALYs (subject to budget constraints[3]) without regard to the way the QALYs are distributed among particular individuals or groups. That is, everyone's QALYs are valued equally: an improvement in *a*'s health is considered to be on a par with a similar improvement in *b*'s health regardless of intelligence, gender, social class, occupation, race, and so on. Moreover, because the QALY is a non-disease specific health-state index, and therefore carries no reference to the cause of the health state, it has the advantage of screening

---

1. As Dolan and Olsen note, another important function of health care, if not an outcome, is care itself. Palliative care provides an obvious example. This is not directly aimed at cure or prevention, but at offering assistance to those whose health cannot be substantially improved. In contrast with cure and prevention, "the interaction between a carer and a patient is not justified by its outcomes, but by such process-related concepts as dignity, respect, autonomy, empathy, and sympathy" (Dolan and Olsen 2002).

2. When some of the costs are "off budget" - for example, patient time costs - then this rule is more complex, as resource costs do not correspond with the fixed budget that forces prioritisation (Richardson 1991).

3. This review is not concerned with the allocation of resources between health and non-health sectors (the latter including education, housing, transport, and so on). It focuses on allocative issues within the health sector.

out public prejudices against certain diseases (such as AIDS or schizophrenia), certain causes of disease or injury (such as boxing), and certain groups (for example, where a disease is prevalent only in females, or primarily in a minority ethnic group) (Edgar 1995, p. 242). In this sense the cost-per-QALY prioritising rule - the rule that, all else being equal, the lower the cost per QALY offered by a health programme or service, the higher a priority it should be - is egalitarian.

QALY-based CEA then is both "fair" and "unfair". It is fair in the sense that it does not discriminate against patient groups on grounds such as race and gender. It is unfair in the sense that it does not distinguish between patients groups when this is morally relevant, such as when groups differ in terms of severity of illness or potential to benefit from treatment. Both of these implications in fact derive from the same source: CEA assigns no value to any distributive or procedural concern beyond what is entailed by health maximisation. This has lead some economists and philosophers to suggest modifications to the strategy of simple QALY maximisation in the interests of fairness. The most important of these suggested modifications, and others that have received little or no attention, are examined in this review. They are divided into three categories depending on the degree of empirical support they can claim: catogory A includes those that have good empirical support, category B includes those for which the empirical support is suggestive but inconclusive, and category C includes those that can only claim poor empirical support at present. The current document examines the issues in categories A and B. Category C is included in *Neglected Equity Issues in Cost-effectiveness Analysis – Part 2.*

CATEGORY A

SEVERITY OF PRE-TREATMENT CONDITION

In conventional CEA the initial health state of a patient is only of importance because health improvement depends upon HRQoL (Health-Related Quality of Life) before and after treatment. However, when informed of the fact that individual patients find two health improvements to be of identical benefit, people generally express a strong preference for allocating resources to those with the worst initial health state. This result has been independently derived in Norway, Australia, the USA, Canada and Spain (see Table 1). It is true that moderately ill patients can only benefit moderately from treatment, while severely ill patients can benefit substantially. However, when moderately ill and severely ill patients are expected to derive the same benefit, and all else is equal, prioritising on the basis of the cost-per-QALY rule will not distinguish between them, and will actually "discriminate" against more severely ill patients when doing so is expected to produce the most QALYs.[4] There is no value associated with the severity of the initial condition itself. The cost-per-QALY prioritising rule disregards entirely the following sentiment: "Our bias, I contend, should be to give priority to persons whose suffering and inability to function in ordinary life is most pronounced, even if the available treatment for them is comparatively less efficacious than for other conditions" (Callahan 1994, p. 463).

Table 1. Studies of Social Preferences for Severity of Illness

Study	Country	Number of Subjects	Type of Subjects	Elicitation Method
(Nord 1991)	Norway	1,141	Doctors/Bio[a]/GenPop[b]	PTO[d]
(Nord 1993b)	Norway	150	Politicians	DCM[e]
(Nord, Richardson et al. 1993)	Norway/Australia	486	GenPop/Students/Nurses	PTO
(Nord 1993a)	Norway	10	NIPH[c]	PTO
(Abelson, Lomas et al. 1995)	Canada	280	GenPop/Health Officials	DelPol[f]
(Ubel, Loewenstein et al. 1996)	USA	42	Economics Students	PTO
(Prades 1997)	Spain	30	Economics Students	PTO
(Ubel, Spranca et al. 1998)	USA	289	Prospective Jurors	DCM
(Ubel 1999b)	USA	479	Prospective Jurors	DCM

a   Bio = Bioengineers
b   GenPop = General Population
c   NIPH = Staff at the National Institute of Public Health in Oslo
d   PTO = Person Trade-Off
e   DCM = Discrete Choice Method
f   DelPol = Deliberative Polling

Empirical studies show that people reject the idea that social value is determined solely by increases in HRQoL, as measured, for example, by the Standard Gamble (SG), Time Trade-Off (TTO) or personal Willingness To Pay (WTP). They reject the principle put forward by Stokey

---

4. We should not construe such discrimination as intentional; it is often an unrecognised or overlooked implication of the exclusive focus on health-maximisation in conventional CEA. Savulescu calls such unintentional discrimination "*de facto* discrimination" (Savulescu 1998, pp. 215-216). There is no intention to disadvantage those who are more severely ill, or any presumption that they are less worthy of concern and respect. In this respect it differs from racism, sexism, ageism, and the like, as ordinarily understood, underlying which there is an intention to deny some groups and individuals less than others *based on* race, sex, age, and so on.

and Zeckhauser that, "individual welfare is all that matters in policy choices" (Stokey and Zeckhauser 1978, p. 262). When asked to judge for others – when asked to adopt a social perspective – respondents systematically re-weight individual patient preferences according to the severity of the initial health state. As a result, the priority accorded to patients with a severe condition does not reflect the utility – the HRQoL - those patients would personally derive from treatment. Rather, it reflects a social judgement about the distribution of health benefits. As Cohen notes: "society may want to direct resources preferentially to those who are farthest from good health, even if larger aggregate benefits could be obtained under a different distribution" (Cohen 1995, p. 287).

The idea that the worse off have a moral claim for special consideration has long received support among philosophers. Most famously, Rawls argued that social and economic inequalities are justifiable only in so far as they are to the greatest benefit of the least advantaged members of society (Rawls 1971, pp. 302-3). Even philosophers within the utilitarian tradition – from which CEA derives its credentials - have attempted to account for the intuition that the worst off deserve special consideration. For example, Wolf argues from a utilitarian perspective that actions which reduce (or minimise) misery are prima facie obligatory, whereas actions that increase well-being, although good, are not obligatory (Wolf 1997). This gives lexical priority to reducing misery, and thus to helping those who are worst off. Maximising well-being becomes a secondary consideration.

Rawls has been criticised for not allowing natural inequalities, such as those arising from health status, to be a factor in determining who is worst off and therefore more deserving of compensation. For example, Kymlicka comments: "According to Rawls, people born into a disadvantaged class or race not only should not be denied social benefits, but also have a claim to compensation because of that disadvantage. Why treat people born with natural handicaps any differently? Why should they not also have a claim to compensation for their disadvantage?" (Kymlicka 1990, pp. 72-73). Green makes a similar point, arguing that health care is in fact a *social* good: "Access to health care is not only a social primary good, but possibly one of the most important such goods ... [because] disease and ill health interfere with our happiness and undermine our self-confidence and self-respect" (Green 1976, p. 117). As we will see, empirical studies of social preferences in the health area show that many people agree with Kymlicka and Green that inequalities arising from health status – even when they are "natural" rather than "social" - are relevant to prioritising health care, to the extent that people are willing to give priority to the more severely ill even if this means reducing aggregate health gains.

The consideration of severity is not a purely theoretical exercise. In the USA severity is a dominating factor in the allocation of heart and liver transplants (when need exceeds supply). Those with the best prognosis after receipt of an organ are those with the least severe illness, and maximum health gain would be achieved by giving this group priority. By contrast, the actual policy gives a very high weighting to those with the most severe problem. This results in the "perverse" situation where the relatively healthy must wait until their health has deteriorated sufficiently for them to satisfy the severity criterion (Wikler 1989; Ubel, Arnold et al. 1993). This policy can only be described as "perverse", however, if maximising health gains is the overriding social objective. In the present case, health production is explicitly of secondary importance to severity.[5] Support for severity can also be found in official government guidelines in several countries, and in reports of government-appointed commissions (see, Dutch Committee on Choices in Health Care 1992; Campbell and Gillett 1993; Swedish Health Care and Medical

---

5. Kidney transplantation is not included in this policy since dialysis is available as an alternative – that is, the condition without transplantation is not sufficiently severe for inclusion in the policy.

Priorities Commission 1993). In brief, whether or not to weight QALYs for severity of pre-treatment condition can have significant practical implications.

In an early study Nord found that returning one person to full health from the following state – "unable to work, unable to pursue family and leisure activities, strong pain, depressed" – was considered as valuable as returning 50 people to full health from the following state – "unable to work, moderate pain". However, the utility values for these states (assigned by the participants using a rating scale) implied that curing one person in the more severe state should be equivalent to curing two people in the less severe state (Nord 1991; Nord 1999, p. 33). Similarly, returning one person to full health from the first state was considered as valuable as returning 100 people to full health from a state of "moderate pain", even though the utility associated with "moderate pain" implied that curing one person in the more severe state should be equivalent to curing five people in "moderate pain". In brief, the societal value of treating the more severely ill was much higher than would be expected from taking account only of the respective utility scores of patients.

In another study of 150 Norwegian politicians involved in health-care decision making, Nord found that 38 per cent of subjects would give priority to the treatment of an illness that gives "severe" problems in preference to an illness that gives "moderate" problems, even though treatment would help those with the severe illness only "a little" whereas it would help those with the moderate illness "considerably" (Nord 1993b). Another 45 per cent would divide any increase in funding evenly between the two, leaving only 11 per cent who would follow the health maximisation strategy of conventional, QALY-based CEA.

Table 2. Comparison of Norwegian and Australian Results: Distribution in Percentages, Median Values, and Confidence Intervals (CI)

Number cured in state *i* equivalent to 10 lives saved	State			
	W[a]		Z[b]	
	Norway	Australia	Norway	Australia
0-9	4.1	0.0	5.9	6.2
10	8.8	9.3	23.5	21.5
11-19	4.1	2.3	5.9	7.7
20-39	8.3	6.9	11.8	10.8
40-50	4.1	20.9	11.8	20.0
51-100	20.8	16.3	23.5	16.9
101-999	8.3	25.7	5.9	6.1
1000 up	41.7	18.6	11.8	10.8
Total	100.0	100.0	100.0	100.0
Number	24	43	17	65
Median	110	85	50	40
90% CI	100	50	25	30
	200	200	100	50

Source: (Nord, Richardson et al. 1993)

    a   State W = Uses crutches for walking. Light pain intermittently. Unable to work.
    b   State Z = Sits in a wheelchair. Pain most of the time. Unable to work.

In a joint Norwegian-Australian study, Nord, Richardson et al. surveyed members of the general public (in Norway) and students and nurses (in Australia) using the PTO. Subjects were asked to adopt the perspective of members of Parliament, and to choose between two equally expensive

special health units. Unit A would save ten people per year from dying and restore them to full health. Unit B would restore to full health a larger number of patients suffering from a chronic illness (W or Z). Participants were asked to indicate how many patients treated in unit B per year they consider equivalent to saving the ten patients in unit A. Table 2 shows that the median response in Norway was 110 for state W and 50 for state Z, and in Australia was 85 for state W and 40 for state Z. These numbers are higher than would be expected from the individual utility scores for these states. As Nord, Richardson et al. note, the utility scores seriously underestimate the *social value* placed upon the health states when the alternative is death (Nord, Richardson et al. 1993, p. 46). The authors also comment, however, that subjects' responses to the PTO exercise were highly dispersed, "indicating the likelihood of a high sampling error for the median values" (Nord, Richardson et al. 1993, p. 467).

In another study, Nord asked a convenience sample of ten individuals from the National Institute of Public Health in Oslo to compare improvements in health on a seven-level disability scale with approximately equal distances between the levels (Nord 1993a). For example, subjects were asked to indicate how many patients moving from level 5 to level 1 on the scale they considered equivalent to a smaller number of patients moving from level 6 to level 4 (where lower numbers represent better functioning). Again, the results showed a marked preference for treating the more severely ill (see Table 3). Although taking a patient from level 5 to level 1 should be twice as valuable as taking a patient from level 6 to level 4, taking into account only gains in HRQoL, Nord's subjects judged them to be approximately equal: taking 16 patients from level 5 to level 1 and taking 17 patients from level 6 to level 4 were both found to be equivalent to taking one person from dying to a state of healthy (the latter was used as a reference state). According to Nord, these findings "support the hypothesis that the conventional QALY model, which attributes all social appreciation to treatment effect, is misconceived" (Nord 1993a, p. 233). Moreover, because the levels on the disability scale were judged approximately equal by the subjects themselves, it is hard to explain the discrepancy between the utility-based predictions and the direct measurements by arguing that the health states were not placed on an interval scale.

Table 3. Numbers of Different Outcomes That May Be Considered Equivalent in Social Value

		Severity with intervention						
		7	6	5	4	3	2	1
Severity without intervention	1							
	2							128
	3						83	64
	4					54	42	32
	5				35	27	21	16
	6			22	17	13	10	8
	7		15	12	9	7	5	4
	Dead	?	3	2	2	1	1	1

Source: (Nord 1993a)

Seven Step Severity Scale:
1. No problems with walking.
2. Can move about without difficulty anywhere, but has difficulties with walking more than a kilometer.
3. Can move about without difficulty at home, but has difficulties in stairs and outdoors.
4. Moves about without difficulty at home.
5. Can sit. Needs assistance to move about – both at home and outdoors.
6. To some degree bedridden. Can sit in a chair part of the day if helped up by others.
7. Completely bedridden.

General support for the importance of severity was also detected in a Canadian study by Abelson, Lomas, and colleagues. Participants in this study were drawn from five different groups: randomly selected citizens, attendees at town-hall meetings, appointees to district health councils, elected officials and experts in health care and social services. All five groups thought that information about *needs* was most important for health-care and social-service decision making (Abelson, Lomas et al. 1995). With the exception of elected officials, all groups thought that information about *benefits* was next in importance, followed by costs, and finally preferences. Elected officials ranked costs second. It is possible to identify need with potential to benefit - those who are in greatest need are those who can benefit most from treatment – and in this way make need as an allocation criterion compatible with the cost-per-QALY prioritising rule. Alternatively, it is possible to think of those who are in greatest need as those who are worst off – those who are most severely ill (Edgar, Salek et al. 1998, pp. 68-71). The subjects in this study appear to have adopted the latter interpretation, and ranked information on benefits as *less important* than information on need (severity). In this way they confirmed Lockwood's view that "the QALY arithmetic is inherently insensitive to differences in degree of need, except in so far as they happen to correlate with the degree of benefit per unit cost that treatment can confer" (Lockwood 1988, p. 45-46).

Ubel, Spranca and colleagues, using the same seven-step scale as Nord in the previous study (1993a), found that the observed preference for more severely ill patients extends to preventative interventions (Ubel, Spranca et al. 1998). Using prospective jurors as subjects, Ubel found only a slight preference for preventative over curative interventions when they brought similar benefits at similar costs. However, there was a significant preference for helping the more severely ill in both contexts. Moreover, this preference was observed, in both the curative and preventative context, even when the more severely ill would benefit less.

In another study, Pinto-Prades asked subjects in Spain to assume the role of health planners in an exercise designed to compare the implications of the Visual Analogue Scale (VAS), the SG and (three forms of) the PTO for the purposes of priority setting (Prades 1997). The study used the following four EuroQoL health states:

12121
No problems with walking about.
Some problems washing or dressing self.
No problems with performing usual activities (e.g. work,
   study, housework, family or leisure activities).
Moderate pain or discomfort.
Not anxious or depressed.

21312
Some problems with walking about.
No problems with self care.
Unable to perform usual activities.
No pain or discomfort.
Moderately anxious or depressed.

23232
Some problems with walking about.
Unable to wash or dress self.
Some problems with performing usual activities.
Extreme pain or discomfort.
Moderately anxious or depressed.

32331
Confined to bed.
Some problems washing or dressing self.
Unable to perform usual activities.
Extreme pain or discomfort.
Not anxious or depressed.

The values assigned to these health states by the VAS, the SG and the PTO are shown in Table 4. Again, the PTO, which measures social value and incorporates information on severity per se, consistently assigned higher values to these states. Thus, for example, for state 21312 the VAS score was 52, the SG score was 81 and the PTO score was 95. This implies that saving one life would be equivalent to curing two people of condition 21312 according to the VAS ($0.48 \times 2 = 0.96$), five people according to the SG ($0.19 \times 5 = 0.95$), and twenty people according to the PTO ($0.5 \times 20 = 1.0$). Clearly, using the PTO will place a much higher value on life saving than the VAS or the SG.

Table 4. Values of the Health States Using Different Preference Elicitation Methods

	Mean (SE)				
	VAS	SG	PTO-1	PTO-2	PTO-3
12121	75	95	98.5	99.20	99.14
	(2.37)	(1.28)	(0.82)	(0.29)	(0.58)
21312	52	81	95	97.05	93.73
	(2.19)	(2.92)	(1.31)	(0.69)	(2.43)
23232	29	71	84	90.78	81.26
	(2.01)	(3.83)	(3.83)	(2.27)	(4.5)
32331	16	44	59	79.65	59.25
	(1.52)	(5.11)	(6.44)	(5.32)	(6.44)

Source: (Prades 1997)

In a study in the US, Ubel, Loewenstein and colleagues conducted a study with economics students using the VAS, the SG and TTO. Using these techniques they sought to measure the utility associated with three health states – ganglion cyst of the hand, ligament damage to the knee, and severe headache. The results are given in Table 5. They then tested the same students one to three weeks later using the PTO to elicit their rationing choices for groups of patients with the same three conditions (Ubel, Loewenstein et al. 1996). They also added a fourth, fatal condition – appendicitis. Participants were asked the following question concerning appendicitis and meningioma (and similar questions concerning the other conditions):

> **A.** Which do you think would bring the most benefit?
> _____ ten people cured of appendicitis
> _____ * * people cured of meningioma
> _____ indifferent

The questions were tailored to individual participants – that is, the double asterisks were replaced by the number at which each participant was predicted to be indifferent, based on their answers to the utility elicitation questions. The study showed that people do not agree with the rationing implications of their answers to utility elicitation questions. For example, when asked to compare appendicitis with meningioma, 37 subjects chose to treat the appendicitis patients, 2 chose to treat those with meningioma, and 3 were indifferent (see Table 6.) "This preference to treat more severely ill people was consistent across all six rationing choices and all three methods of elicitation" (Ubel, Loewenstein et al. 1996, p. 113). Confirming this, subjects were asked:

> **B.** How many people would have to be cured of meningioma to equal the benefit brought by curing ten people of appendicitis?

Table 5. Mean Utilities for the Health States as Measured by the Three Utility-elicitation Methods

Health Condition	Analogue Scale	Utilities	
		Standard Gamble	Time Tradeoff
Cyst	0.92	0.91	0.99
Knee	0.63	0.83	0.94
Meningioma	0.37	0.75	0.90

Source: (Ubel, Loewenstein et al. 1996)

Table 6. Rationing Choices for Health States Between Which Subjects Should Be Indifferent

Treatment Choice	Analogue Scale	Standard Gamble	Time Tradeoff	Total	Significance of Difference
Appendicitis vs meningioma					$p < 0.0001$
Appendicitis	15	10	12	37	
Indifferent	0	1	2	3	
Meningioma	2	0	0	2	
Appendicitis vs knee					$p < 0.0001$
Appendicitis	17	11	13	41	
Indifferent	0	0	0	0	
Knee	0	0	1	1	
Appendicitis vs cyst					$p < 0.0001$
Appendicitis	17	11	14	42	
Indifferent	0	0	0	0	
Cyst	0	0	0	0	
Meningioma vs cyst					$p < 0.0001$
Meningioma	17	9	14	40	
Indifferent	0	2	0	2	
Cyst	0	0	0	0	
Meningioma vs knee					$p < 0.0001$
Worse	15	9	13	37	
Indifferent	1	2	1	4	
Better	1	0	0	1	
Knee vs cyst					$p < 0.0001$
Knee	15	10	13	38	
Indifferent	1	1	1	3	
Cyst	1	0	0	1	

Source: (Ubel, Loewenstein et al. 1996)

The results are given in Table 7, along with five other pair-wise comparisons. The value 8.3 in the second column means that the point at which subjects were indifferent between the two outcomes (cured of appendicitis versus cured of meningioma), calculated by their answers to question B, was 8.3 times greater than the point at which they were predicted to be indifferent based on their answers to the utility elicitation question (using the VAS). "The results show how far the respondents' answers diverged from those predicted by their answers to the utility survey" (Ubel, Loewenstein et al. 1996, p. 113).

Table 7. Discrepancies Between Indifference Points of Rationing Choices and Those Predicted by Utility-survey Responses

Rationing Scenario	Median Discrepancy*		
	Analogue Scale	Standard Gamble	Time Tradeoff
Appendicitis vs meningioma	8.3	100.0	81.5
Appendicitis vs knee	35.7	10,000.0	720.9
Appendicitis vs cyst	100.0	6.0 EX8	†
Meningioma vs cyst	10.0	1270	100.0
Knee vs cyst	10.0	40.0	5.0
Meningioma vs knee	44.3	3.3	10.0

Source: (Ubel, Loewenstein et al. 1996)

* A value of 1.0 means that the subject's indifference point in the rationing choice was the same as that predicted by the utility survey response (no discrepancy) and 100 means that the subject's indifference point in the rationing survey was one hundred times greater than that predicted by the utility survey response.

† Could not be calculated because the median lay between a numerical answer and one denoting infinity.

In a subsequent study using prospective jurors as subjects, Ubel repeated the earlier study by Nord (1993b). Like Nord, he found that many people are prepared to sacrifice overall health gains in order to benefit those with the worst initial health state. Of 479 subjects, 9 per cent gave priority to patients with "moderate" health problems, 26 per cent gave priority to those with "severe" health problems, and 64 per cent chose to divide resources equally between the two

Table 8. Subjects' Justifications for their Allocation Preferences

Justification	Number of subjects providing justification according to allocation preference[a]			
	Priority to severely ill patients	Priority to moderately ill patients	Divide resources evenly	Total
Fairness	2	1	101	104
Maximize health benefits	0	83	3	86
Severely ill deserve priority	52	0	2	54
Future research benefits	11	1	10	22
Prevent decline in health	2	11	8	21
Future economic benefits	4	11	2	17
Severe illness more urgent	8	0	1	9
Do not discriminate according to treatment benefits	0	0	7	7
Level the playing field	5	0	1	6

Source: (Ubel 1999b)

a Some subjects provided more than one justification.

groups. However, Ubel noted that when subjects are not given the option of dividing resources evenly between patients with "severe" health problems and those with "moderate" health problems, only a small majority favoured giving priority to the former group. This result raises the possibility that subjects in Nord's original study, and perhaps in other studies, were not expressing a preference for severity per se, but "may have simply been unwilling to make a difficult treatment choice" (Ubel 1999b, p. 897). Also, Ubel found that responses were sensitive to the wording of the options. For example, when subjects were reminded about how much improvement each group was expected to undergo ("a little" for the severely ill, and "considerably" for the moderately ill), fewer participants gave priority to the severely ill (6 per cent compared with 26 per cent without the reminder), and more gave priority to the moderately ill (21 per cent compared with 9 per cent without the reminder). This does not negate the significance of severity, as Ubel acknowledges (Ubel 1999b, p. 902), but it does indicate that caution must be exercised in calculating just how much weight it should carry when setting health care priorities. Table 8 shows subjects' reasons for their allocation choices.

Taking into consideration the information on public preferences for severity revealed in the preceding studies, Nord divides health states into three classes – "severe", "considerable" and "moderate" - and assigns them values consistent with the above empirical findings (Nord 1999, pp. 37-38). See Table 9. These values give rise to some "rules of thumb" concerning severity: saving someone from death is something like 3-6 times better - has greater social value - than curing someone of a severe health problem, something like 10-15 times better than curing someone of a moderate problem, and 50-200 times better than curing someone of a moderate problem. According to Nord: "Quantitative models that purport to be useful for estimating the societal value of health care activities in these countries [Australia, England, Norway, Spain, and the US], as well as in other countries with similar values, must reflect this structure of concern" (Nord 1999, p. 38). Notice that, by contrast with the health state (QALY) values used in CEA – derived, for example, by means of the SG, TTO or RS – these social values are higher, particularly at the upper end of the scale. From the societal perspective, conventional CEA underestimates the value of curing severe health problems, including life-saving treatments. The value structure encapsulated in Table 9 therefore "compresses health states to the upper end of the scale" (Nord 1999, p. 38).

Table 9. Rules of Thumb Concerning Severity (after Nord 1999)

Health state	Description	Value range
Death		0.0
Severe	For instance, a person who has to sit in a wheelchair, has pain most of the time, and is unable to work	0.65-0.85
Considerable	For instance, a person who must use crutches to walk, has light pain intermittently, and is unable to work	0.90-0.94
Moderate	For instance, a person who has difficulty moving about outdoors and slight discomfort, but is able to do some work and has only minor difficulties at home	0.98-0.995
Full health		1.0

It should be noted that measuring the severity of different health states is not without its problems. First, there is disagreement about the best way to measure HRQoL. Different techniques, such as the SG, the TTO, and VAS, produce different results, and have their own advantages and disadvantages (Froberg and Kane 1989; Nord 1992; Richardson 1994; Dolan, Gudex et al. 1996; Prades 1997). As Nord observes, "this is naturally a serious problem for

potential users of economic evaluations" (Nord 1999, p. 91). Second, there is disagreement about whose preferences should be used to value health states in the first place. There is evidence that different groups – patients, health professionals, "well-informed" members of the general public – rate health states differently (Boyd, Sutherland et al. 1990; Rothman 1991; Hurst, Jobanputra et al. 1994; Polsky, Willke et al. 2001; Menzel, Dolan et al. 2002).

Finally, preference elicitation in general is subject to framing effects (such as loss aversion and diminishing sensitivity), judgement biases (such as anchoring), and context effects (such as menu expansion and contraction), (Kahneman, Slovic et al. 1982; Kahneman, Knetsch et al. 1991; Tversky and Kahneman 1991; Rabin 1998). These problems confronting preference elicitation call into question the underlying assumption that people have stable, coherent preferences that are waiting to be elicited, and raise the suspicion that in some cases preferences might be *determined* by the elicitation method. However, despite these caveats - which will affect any attempt to incorporate social preferences into health care decision-making systematically - there is a growing body of persuasive empirical evidence indicating that severity of illness is a significant factor for many people when allocating limited health care resources, to the extent that sacrificing overall health production to help the more severely ill is a justifiable option for many people. There is reason to think that the notion of "effectiveness" used in cost-effectiveness analyses should be expanded to encompass not just potential health gains, but a weighting of such health gains to reflect the severity of the initial condition of patients.

REALISATION OF POTENTIAL FOR HEALTH

Conventional CEA, using unweighted QALYs as the unit of effectiveness, places no value upon patients realising their potential for health improvement per se. Saving the life of a paraplegic, B, will count for less than saving the life of a non-paraplegic, A, other things being equal, due to the lesser value of life years with paraplegia (see Figure 1). For the same reason, saving the life of someone who will be left a paraplegic, D, will count for less than saving the life of someone who will regain full health, C (see Figure 2). The same is true when health improvement rather than life saving is at issue. A treatment that will restore a disabled patient, E, to full health will always be a higher priority, other things being equal, than a treatment that will produce a smaller but perhaps still significant benefit for a similarly disabled patient, F (see Figure 3). And a treatment that will secure a given benefit for a disabled patient who is better off, G, will always be a higher priority, other things being equal, than a treatment that will secure an equivalent relative benefit for a patient who is worse off, H (see Figure 4). However, there is mounting evidence that people do not want to discriminate against those with a reduced capacity to benefit from treatment to the extent implied by conventional CEA.

Figure 1. Potential for Health: Pre-existing Disability Case
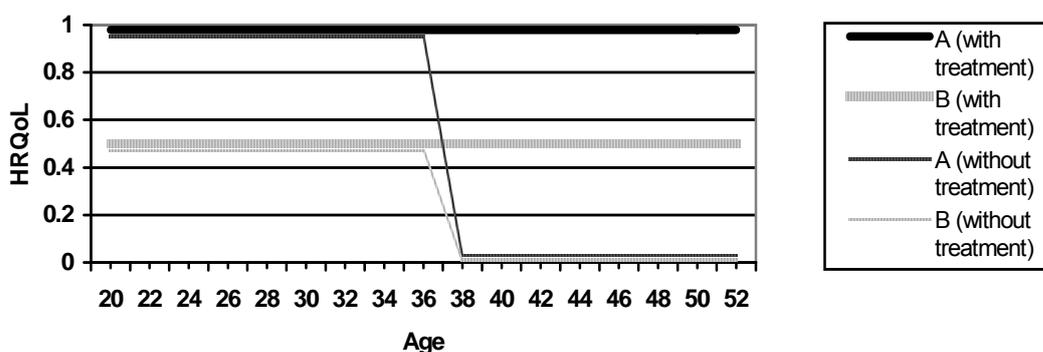
Figure 2. Potential for Health: Disability Onset Case
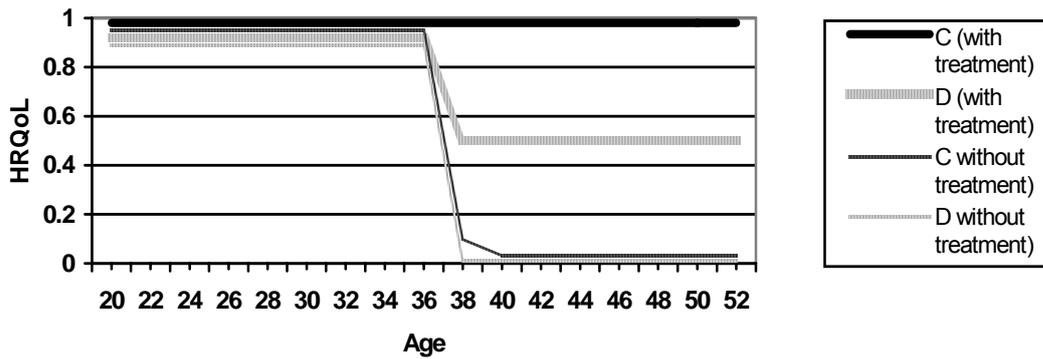
Figure 3. Potential for Health: Unequal Improvement Case
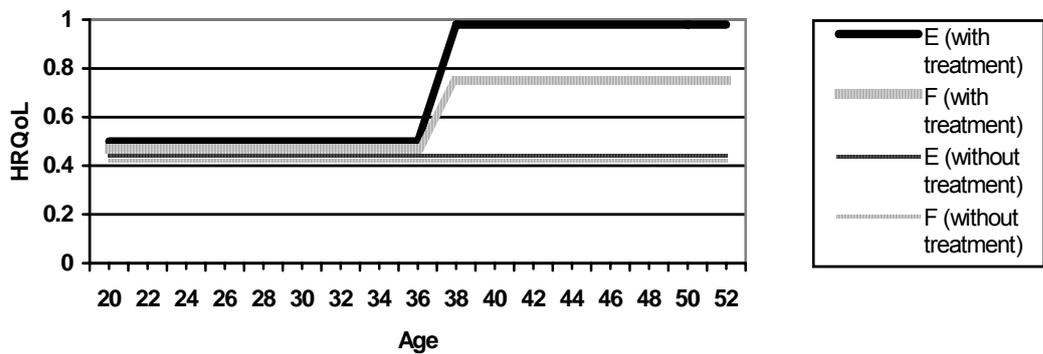
Figure 4. Potential for Health: Equivalent Improvement Case
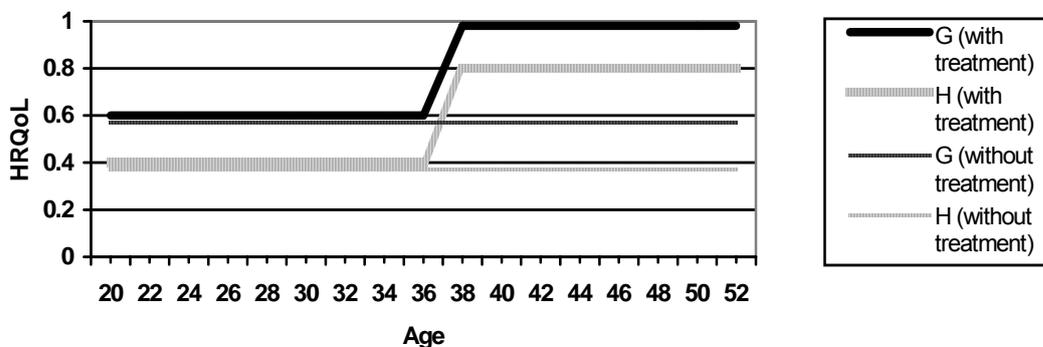
Harris argues that in the life saving context discriminating against the permanently disabled and chronically ill, on the basis that fewer QALYs will be produced, involves a form of "double jeopardy":

> QALYs dictate that because an individual is unfortunate, because she has once become a victim of disaster, we are required to visit upon her a second and perhaps graver misfortune. The first disaster leaves her with a poor quality of life and QALYs then require that in virtue of this she be ruled out as a candidate for lifesaving treatment, or at best, that she be given little or no chance of benefiting from what little amelioration her condition admits of (Harris 1987, p. 120).

According to Harris, the cost-per-QALY prioritising rule discriminates unfairly against those who, when faced with a life-threatening illness, are unfortunate enough already to have a permanent disability, because the lower an individual's anticipated HRQoL after treatment the lower their

QALY score (other things being equal). In this sense the permanently disabled suffer a "double jeopardy". Not only are they unfortunate enough to be disabled, but also as a result of this they are disadvantaged in the competition for limited resources that may save their lives. Scanlon makes the same point: "It seems right … that a person who suffers from physical disabilities making it difficult and expensive for him to enjoy any of the normal pleasures of life should receive special assistance…. [T]he principle of equal satisfaction may set this level too high, but the general tendency to recognise such claims does not seem wrong" (Scanlon 1975, p. 659). The cost-per-QALY prioritising rule does not recognize such claims, however, and, when strictly applied, will direct resources away from the permanently disabled and chronically ill if their ability to benefit from treatment is less than an able-bodied person.[6] The United States Secretary for Health and Human Services used a form of "double jeopardy" argument in rejecting the Oregon plan for prioritising the State's health services in 1992 (Daniels 1993).

It would be wrong to assume that the disabled are always "inefficient utility producers". As Stein cautions, we must not assume that low *total* welfare implies low *marginal* welfare from resources (Stein 2002, p. 4). A disabled person may derive more utility from a wheelchair than an able-bodied person would derive from a new set of golf clubs, even if they cost the same. Nevertheless, when disability and illness do interfere with a person's ability to benefit from the health care they receive they will be disadvantaged by conventional CEA, and it is the fairness of this that Harris, Scanlon, and others, question. The following studies suggest that many citizens in Norway, Australia, the USA and Spain share this reluctance to discriminate against the permanently disabled and chronically ill.

In a Norwegian study, Nord used the PTO to investigate people's views on the relevance of health state after treatment in prioritising between different patients (Nord 1993d). One group was asked to choose between two 60-year old housewives, each of whom had contracted a life-threatening disease. An operation would make Mrs Anderson "completely healthy", whereas it would give Mrs Peterson a life "with moderate pain and dependency on crutches for walking". Being informed that it may not be possible to operate on both patients, subjects were asked whether they thought Mrs Anderson should be operated on first, whether Mrs Peterson should be operated on first, or whether they should be taken in the order in which they were admitted to the hospital. The results are summarised in Table 10. They show that the majority of participants, 78.7 per cent, chose the option that afforded both patients an equal chance of being treated (order of admission), whereas only 14.8 per cent favoured the most cost-effective option (treating

Table 10. Prioritising at Hospital Level

	Number	%
Mrs Anderson should be operated on first	9	14.8
Mrs Peterson should be operated on first	2	3.2
They should be operated on in the order in which they were admitted to the hospital	48	78.7
Uncertain	2	3.2
Total	61	99.9

Source: (Nord 1993d)

6. Note also Dworkin in the special case of children and wealth distribution: "When the question arises how wealth should be distributed among children, for example, those who are seriously physically or mentally handicapped do seem to have, in all fairness, a claim to more than others" (Dworkin 1981, p. 189).

Mrs Anderson). "The subjects rationalized their responses by pointing to equality between individuals in value of life and entitlement to treatment irrespective of differences in potential for health" (Nord 1999, p. 39).

Twenty-three participants were also asked to adopt a budget-level rather than a hospital-level perspective, and to choose between two equally costly units, A and B. Unit A would save 10 people per year from dying and make them "completely well", whereas unit B would save 10 people from dying, but give them a life "with moderate pain and dependency on crutches for walking". Subjects were asked how many patients treated in unit B they consider equivalent to 10 patients being treated in unit A. As Table 11 shows, just over half of those who chose not to discriminate between Mrs Anderson and Mrs Peterson, 11 out of 19, also considered 10 patients treated in unit B equivalent to 10 patients treated in unit A. Subjects were also asked about life-enhancing as opposed to life-saving treatments. In this case, unit A would move patients from the state "must sit in a wheelchair and endure strong pain" to "completely well", whereas unit B would move patients from the state "must sit in a wheelchair and endure strong pain" to a life "with moderate pain and dependency on crutches for walking". As Table 11 also shows, 9 out of 18 of those who chose not to discriminate between Mrs Anderson and Mrs Peterson also considered 10 patients treated in unit B equivalent to 10 patients treated in unit A. On the basis of telephone interviews with the subjects taking part in this study, Nord concludes: "what seems to count for the vast majority of the subjects, is the view that individuals are equally valuable and equally entitled to treatment" (Nord 1993d, p. 39). As Nord notes, from this, it cannot be concluded that health state after treatment is irrelevant in prioritising between different patients. When the potential of patients to benefit is very limited – when the final health state is much worse than "moderate pain and dependency on crutches for walking" – subjects may place considerably more weight on health outcome. Nonetheless, "the possible negative implications for the use of cost per QALY as a criterion for prioritising among different health care programmes are obvious" (Nord 1993d, p. 41).

Table 11. Distribution of Responses with Respect to Equivalence Numbers at Budget Level by Respondent Category in Question at Hospital Level

	Life-saving		Chronically ill	
Equivalence number	Mrs Anderson	Equal priority	Mrs Anderson	Equal priority
10	1	11		9
11		7		6
15		1		1
20	1		2	1
30	1		1	1
50	1			
100			1	
All	4	19	4	18

Source: (Nord 1993d)

Abellan-Perpiñán and Pinto-Prades repeated Nord's study with Spanish subjects (Abellan-Perpiñán and Prades 1999). However, they changed the sex of the patients, and their ages from 60 to 20 years. They also split the subjects into two groups. The first group (X) was asked to choose between a life saving operation that would make Charles "completely healthy", and a life saving operation that would give James a life "with moderate pain and dependency on crutches for walking". They were then asked to choose between a life saving operation that would make

Charles "completely healthy", and a life saving operation that would leave James in a more severe health state: "frequently in bed, can move around the house but after 2-3 hours has to return to bed and stay there for about 5-6 hours resting." The second group (Y) received the same two questions but in the reverse order. The results are given in Table 12. They show that the order in which questions are asked can have a strong influence upon responses. They also support Nord's conclusion that potential for improvement does not have the importance for the general public that an exclusive focus on health maximisation would suggest. The results "cast doubt on the relevance of the health maximization principle in governing resource allocation in health care, and they signal the importance of taking equity principles into account in social decision making" (Abellan-Perpiñán and Prades 1999, p. 704).

Table 12. Number and percentage of respondents supporting different criteria for prioritizing between Charles and James when Charles' final health state is full health and James' final health state is worse than full health

Criteria	Group X				Group Y			
	First question, James with mild final health state		Second question, James with severe final health state		Second question, James with mild final health state		First question, James with severe final health state	
	n	%	*n*	%	*n*	%	*n*	%
Charles first	9	10.9	20	25.1	29	31.8	44	48.4
Admission order	69	83.1	58	69.9	57	62.6	42	46.1
James first	3	3.6	3	3.6	0	0	0	0
No answer	2	2.4	2	2.4	5	5.5	5	5.5
Total	83	100	83	100	91	100	91	100

Source: (Abellan-Perpiñán and Prades 1999)

In the study previously referred to (Nord 1993b), Nord also asked 150 Norwegian politicians involved in health-care decision making to allocate an increase in funds between two diseases that cause the same degree of suffering. Equally costly treatment would help patients with the first disease, A, "a little" whereas it would help patients with the second disease, B, "considerably". Subjects were asked whether most of the increase should be allocated to treatments for illness B, on the ground that the effects here are greater, or whether the increase should be divided evenly between the two groups, on the ground that they suffer equally and are equally entitled to treatment. Nord found that only 24 per cent chose the utilitarian option (B), whereas 72 per cent chose the egalitarian option (A).

In an Australian study designed to test the level of support for utilitarian and egalitarian views in the general public, Nord, Richardson and colleagues asked subjects to choose between the following two options (Nord, Richardson et al. 1995a):

1. Among patients who are suffering equally, some priority should be given to those who will be helped most from treatment.

2. Among patients who are suffering equally, those who can become a little better should have the same priority as those who can become much better.

They found no marked preference for either view. As Table 13 indicates, 52.8 per cent of participants chose option 1 and 47.2 per cent chose option 2. It should be noted, however, that option 1 used the expression "some priority", which tends to imply weak preference, and probably

attracted more support for this option than the word "priority" unqualified would have. Moreover, option 2 asked subjects to compare those who can become only "a little better" with those who can become "much better". If this difference had been smaller, option 2 might have attracted more support. Taking these two points into consideration, the authors conclude: "the observed distribution does not suggest very strong support in the Australian public for the assumptions of health benefit maximization and distributive neutrality" (Nord, Richardson et al. 1995a, p. 1434).

Table 13. Percent Response to Potential Issue and Difficulty of Choice

Issue	Option	Percent choosing each option	Difficulty of Choice (%)		
			Very Difficult	Slightly Difficult	Not Difficult
POTENTIAL	Favour most helped	52.8	29.4	30.9	39.7
	Equal priority	47.2	22.2	30.0	47.7

Source: (Nord, Richardson et al. 1995a)

In the Spanish study previously mentioned (Abellan-Perpiñán and Prades 1999), Abellan-Perpiñán and Pinto-Prades also asked 149 undergraduate students to allocate a fifty-million peseta budget between two patients. The subjects were told that it would cost one million pesetas to produce a year of life for John and two million pesetas to produce a year of life for Andrew. When Group 1 (*N* = 78) was asked how they would split the money, on the assumption that Andrew and John "will live a normal life" after treatment, 74 per cent chose to ensure that both patients would gain 16.6 years, although this meant spending twice as much on Andrew (33.3 million pesetas) as on John (16.6 million pesetas) (see Table 14). When the same subjects were asked to imagine that Andrew would live the rest of his life undergoing dialysis, 84 per cent did not change their view (see Table 15). This supports the findings by Nord, Richardson et al. (1995a), and suggests that health state after treatment is relatively unimportant. However, in a separate exercise, subjects in Group 2 (*N* = 71) were told from the beginning that Andrew would live the rest of his life undergoing dialysis. In this case only 29 per cent chose to ensure both Andrew and John would gain 16.6 years (by allocating 33.3 million pesetas to Andrew and 16.6 million pesetas to John). By contrast with the earlier finding, this suggests that

Table 14. Number and percentage of respondents reporting different criteria for distributing money between Andrew and John when final health state is the same: Group 1

Money (million pesetas)		Years		*n* (%)
Andrew	John	Andrew	John	
33.3	16.6	16.6	16.6	57 (74)
30	20	15	20	8 (10)
25	25	12.5	25	9 (12)
20	30	10	30	2 (3)
10	40	5	40	1 (1)
0	50	0	50	0 (0)
31.34	18.59	15.67	18.59	Mean
(4.15)	(4.18)	(2.07)	(4.18)	(S.D.)

Source: (Abellan-Perpiñán and Prades 1999)

health state after treatment does matter in prioritising between patients. However, the support for health maximisation is still weak, with only 27 per cent of subjects being prepared to spend more

on John than on Andrew, although this would produce more life years of better quality (see Table 16). This suggests that health state after treatment does matter, but not to the extent that QALY maximisation indicates. As the authors conclude: "social preferences may be somewhere in between two opposite extremes, which are that discrimination based on the degree of health improvement is never acceptable and that discrimination based on the degree of health improvement is always acceptable" (Abellan-Perpiñán and Prades 1999, p. 701).

There are several possible explanations for the different responses given by Group 1 and Group 2 when Andrew would be in a worse final health state after treatment. Abellan-Perpiñán and Pinto-Prades speculate that, because the subjects in Group 1 were asked not only to make the

Table 15. Number and percentage of respondents reporting different criteria for distributing money between Andrew and John when final health state of Andrew is worse than John's: Group 1

Money (million pesetas)		Years		n (%)
Andrew	John	Andrew	John	
33.3	16.6	16.6	16.6	52 (68)
30	20	15	20	8 (10)
25	25	12.5	25	8 (10)
20	30	10	30	5 (6)
10	40	5	40	1 (1)
0	50	0	50	3 (4)
30.02	19.91	15.01	19.91	Mean
(7.31)	(7.35)	(3.66)	(7.35)	(S.D.)

Source: (Abellan-Perpiñán and Prades 1999)

inter-personal comparison between John in full health and Andrew on dialysis, but also the intra-personal comparison between Andrew in full health and Andrew on dialysis, the concern of participants for equitable treatment "came to play a more important role" than in Group 2

Table 16. Number and percentage of respondents reporting different criteria for distributing money between Andrew and John when final health state of Andrew is worse than John's: Group 2

Money (million pesetas)		Years		n (%)
Andrew	John	Andrew	John	
33.3	16.6	16.6	16.6	22 (29)
30	20	15	20	11 (14)
25	25	12.5	25	17 (22)
20	30	10	30	12 (16)
15	35	7.5	35	1 (1)
10	40	5	40	6 (8)
5	45	2.5	45	1 (1)
0	50	0	50	1 (1)
25.46	24.51	12.71	24.51	Mean
(8.06)	(8.09)	(4.02)	(8.09)	(S.D.)

Source: (Abellan-Perpiñán and Prades 1999)

(Abellan-Perpiñán and Prades 1999, p. 705). Abellan-Perpiñán and Pinto-Prades see this as a

confounding factor, and place greater veracity on inferring the significance of health state after treatment by comparing the responses of different groups, since this approach runs less risk of making subjects feel personally responsible for making discriminatory choices. Nord offers an alternative explanation for the different responses given by Group 1 and Group 2. He does not disagree with Abellan-Perpiñán and Pinto-Prades that concern for equitable treatment may have played a more important role for Group 1 than for Group 2. But he does not see this as necessarily a bad thing. "Subjects in group 1 might, to a greater extent than subjects in group 2, have been stimulated to think about the ethical justification of discriminating against a person with a lesser potential for health, and on the basis of such serious reflection found it unjustified" (Nord 1999, p. 41).

In a US study, Ubel, Richardson et al. explicitly distinguished between a "pre-existing paraplegia scenario," in which subjects were "asked how many paraplegics' lives would have to be saved to be just as important as saving 100 'normal' people's lives," and a "paraplegia onset scenario," in which they were "asked how many patients who would experience the onset of paraplegia need to be saved to equal the benefit of saving 100 'normal' lives" (Ubel, Richardson et al. 2002). They found, as in the previous studies, that subjects' indifference points were inconsistent with the principle of QALY maximisation. Whether a person begins as a paraplegic or ends up a paraplegic has less significance for the general population than CEA implies. The authors also varied the order in which subjects were given these two scenarios and found that PTO elicitations are susceptible to significant order effects. For example, among those who received the pre-existing scenario first, the median PTO indifference point for this scenario was 100 (meaning that 100 paraplegics' lives would have to be saved to equal the benefit of saving 100 'normal' lives). By contrast, among those who received the paraplegia onset scenario first, the median PTO indifference point for the pre-existing scenario was 200. The results are given in Table 17. The authors attempted to eliminate the possibility that the observed order effect was due entirely to numerical anchoring by requiring non-numerical responses in some cases. The effect was still observed, though its magnitude was significantly reduced.

Table 17. PTO Indifference Points and Order Effects

Questionnaire design		PTO indifference points: 25th, 50th and 75th percentiles						
Questionnaire version	Scenario order		Pre-existing paraplegia scenario			Paraplegia onset scenario		
		N	25	50	75	25	50	75
1	(1) Pre-existing paraplegia (2) Paraplegia onset	100	100	100	200	100	126	1000
2	(1) Paraplegia onset (2) Pre-existing paraplegia	91	100	200	5000	101	1000	1 000 000

Source: (Ubel, Richardson et al. 2002)

A qualitative study in England by Dolan and Cookson (2000) confirms the suspicion, raised in several of these studies, that people place greater emphasis on health state after treatment when the potential of patients to benefit is very limited – that is, either when the final health state of one group falls below an absolute limit, or when the relative difference between the final health states of the two groups is sufficiently large. Dolan and Cookson asked members of the public (randomly chosen from two GP practice lists) to choose between two groups of patients, A and B, who would either benefit from treatment by differing amounts, or would benefit to the same degree but from different starting points (Dolan and Cookson 2000). The six pair-wise choices are given in Table 18. The first three involve life-saving treatments, and the last three life-

enhancing treatments. Participants were asked whether they would give the same priority to both groups or whether they would give priority to the group that would benefit most. Based on their answers to this question, the health gain to the group that would gain the least was adjusted to ascertain each subject's indifference point. "That is, respondents were asked by the moderator to state the point at which they would 'draw the line' between giving the same priority to both groups, and giving priority to the group who would benefit most from treatment" (Dolan and Cookson 2000, p. 21). In those cases where the health gain was identical (questions 3 and 6), subjects were asked whether they would give the same priority to both groups or whether they would give priority to one group rather than the other based on their starting points. The results are given in Table 19. Subjects' initial choices - that is, prior to being asked to "draw the line" – indicate strong support for giving equal priority to the two groups. "Many respondents appeared to start with a horizontal equity concern for treating everybody equally" (Dolan and Cookson 2000, p. 27).

Table 18. The Questions

Question		Without treatment	With treatment
1	Group A	0 years	0 years
	Group B	0 years	20 years
2	Group A	5 years	10 years
	Group B	5 years	20 years
3	Group A	30 years	40 years
	Group B	10 years	20 years
4	Group A	0%	50%
	Group B	0%	100%
5	Group A	25%	50%
	Group B	25%	100%
6	Group A	75%	100%
	Group B	25%	50%

Source: (Dolan and Cookson 2000)

In their qualitative analysis of the group discussions, Dolan and Cookson divided the reasons given by participants for their decisions into five "principles" (generalised rules for priority-setting) and five "factors" (relevant aspects of the situation not captured by principles, such as the age or health status of a patient). The number of times each principle was mentioned, and the number of groups in which it was mentioned, are given in Table 20. Ideas such as that "a life is a life,"

Table 19. Individual Results

Question	Initial choice			Always give the same priority (%)
	Priority to A (%)	Priority to B (%)	Same priority (%)	
1		42	58	17
2		43	57	15
3	2	50	48	n/a
4		37	63	17
5		42	58	20
6	8	32	60	n/a

Source: (Dolan and Cookson 2000)

that "everyone is equal," and that "equal treatment is fair" were mentioned often and widely, and explain why many subjects (especially in qualitative work) are unwilling to automatically give higher priority to the group that would gain more from treatment. However, when asked to "draw the line", Dolan and Cookson found that many subjects were willing to give higher priority to the group with the greatest potential when the potential of the other group was very limited. "[A]lthough many respondents initially wanted to treat both groups equally, most were willing to trade this principle off for the vertical equity consideration of giving greater priority to those who gain most from treatment when the size of the gain to one group was no longer considered large enough" (Dolan and Cookson 2000, pp. 27-28). As Table 20 indicates, some participants mentioned an absolute threshold – they were willing to give higher priority to the group that would gain most when the potential of the other group dropped below a certain limit, whereas others mentioned a relative threshold – they were willing to give priority to the group that would gain most when the relative difference in potential between the two groups became sufficiently large.

Table 20. Principles Mentioned

Principle	Number of mentions in:	
	Total	Groups
1. Priority for bigger gains to:		
a. length of life	12	5
b. quality of life	2	2
c. health in general	8	4
2. Priority to:		
a. life-threatening conditions	2	2
b. disability in general	4	1
3. Same priority:		
a. a life is a life and everyone is equal	27	8
b. equal treatment is moral, fair, humane	14	4
c. give more individuals a chance	8	5
d. sit on the fence	5	3
4. Priority according to threshold of end-point:		
a. absolute threshold	15	4
b. relative threshold	3	2
5. Priority according to threshold of difference		
a. end-point based absolute threshold	3	1
b. end-point based relative threshold	3	1

Source: (Dolan and Cookson 2000)

The above studies are somewhat at odds with an early study by Patrick, Bush et al. who found a preference for saving the lives of non-disabled people (Patrick, Bush et al. 1973). For example, they found that saving the lives of 100 people "as healthy as possible" was considered equivalent to saving approximately 300 people "aged 40-61," able to "travel with difficulty," "move independently in a wheelchair," "perform major activities with limitations," with "impairment of one foot or leg, such as a fracture, burn, cut, deformity, or paralysis." Nord points to the change in attitudes in most developed countries towards the rights of the disabled as a possible explanation for the divergence between this early study and the more recent studies (Nord 1999). Support for this explanation can perhaps be found in the US government's rejection of the Oregon plan for prioritising that State's health services, on the ground that it would discriminate against the disabled.

Ubel points out a potential difficulty with building into health programme evaluation the social aversion to discrimination against the disabled that is evident in most of the studies above (Ubel 2000, pp. 161-168). If saving the life of someone with paraplegia, for example, is considered as valuable as saving the life of someone who can be returned to full health, then curing paraplegia will have no value. Within the conceptual framework of QALY-based CEA, the utility involved in saving a healthy person from death is 1-0 = 1. If saving the life of someone with paraplegia has the same value, then paraplegia must have a utility of 1 also. "Since the best we can do is to hope to improve the condition so that paraplegics have full health, and since full health and paraplegia have the same utility value …, curing paraplegia would not improve utility" (Ubel 2000, p. 162). Ubel calls this the "QALY trap". The problem arises because CEA is focused exclusively on utility gains – in the present context, QALY gains. It takes no account of the *social value* of different health programmes and services. "The term 'societal value' refers to the strength of public preferences for giving priority to various competing health care programs" (Ubel 2000, p. 164). Once we distinguish between the utility arising from a health intervention, measured for example by the SG, TTO or VAS, and the social value of that intervention, measured for example by the PTO, we can consistently maintain that saving the life of someone with paraplegia is no less valuable than saving the life of someone who can be returned to full

Table 21. Studies of Social Preferences for Potential to Benefit from Treatment

Study	Country	Number of subjects	Type of subjects	Elicitation method	Pre-treatment HRQoL*	Intervention
(Patrick, Bush et al. 1973)	USA	46	Students/HL[a]	PTO	Unequal	LS[b]
(Nord 1993d)	Norway	61	NIPH[c]	PTO[d]	Equal	LS/LE[e]
(Nord 1993b)	Norway	150	Politicians	DCM[f]	Equal	LE
(Nord, Richardson et al. 1995a)	Australia	551	GenPop[g]	DCM	Equal	LE
(Abellan-Perpiñán and Prades 1999) Experiment 1	Spain	174	Undergraduates	PTO	Equal	LS
(Abellan-Perpiñán and Prades 1999) Experiment 2	Spain	149	Undergraduates	SWTP[h]	Unequal	LS
(Dolan and Cookson 2000)	England	60	GenPop	SGD[i]	Equal/Unequal	LS/LE
(Ubel, Richardson et al. 2002)	USA	319	Jurors	PTO	Equal/Unequal	LS

a   HL = Health Leaders (e.g. members of the New York State Health Planning Commission and Advisory Council)
b   LS = Life-Saving
c   NIPH = Staff at the National Institute of Public Health in Oslo
d   PTO = Person Trade-Off
e   LE = Life-Enhancing
f   DCM = Discrete Choice Method
g   GenPop = General Population
h   SWTP = Social Willingness to Pay
i   SGD = Small Group Discussion

*   *A* may have a lower potential to benefit from treatment compared with *B* because *A* has a different illness from *B*. Even if *A* and *B* have the same illness, *A* may have an asymptomatic (e.g. genetic) condition that affects his or her capacity to benefit. In either case, the pre-treatment HRQoL of *A* and *B* may be equal. By contrast, *A* may have a lower potential to benefit from treatment compared with *B* due to a pre-existing condition that manifests itself in a lower HRQoL, such as paraplegia or chronic lung disease. In this case the pre-treatment HRQoL of *A* and *B* are unequal.

health – that both interventions have the same *social value* - and yet maintain that curing paraplegia brings significant benefit to paraplegics in the form of increased (health-related) *utility*.

"When utility measures are supplemented with separate societal value measures, we will no longer be caught in the trap" (Ubel 2000, p. 164).

In summary, the literature in this section suggests that potential to benefit from treatment carries less weight for many people than is implied by conventional CEA. Only one study of the Australian community's attitude towards potential to benefit from treatment has been undertaken. Moreover, this study, conducted by Nord, Richardson et al. (1995a), did not look at the significance of potential in the life-saving context, nor whether the significance of potential varies when initial health status is different. Also, Nord, Richardson and colleagues did not attempt to quantify the egalitarian preferences it detected in this study. It is clear that more research needs to be done before the significance of potential is known.

CATEGORY B

CONCENTRATION AND DISPERSION OF HEALTH BENEFITS

Appealing solely to the cost-per-QALY prioritising rule, there is no difference in value between giving many QALYs to a few people and giving a few QALYs to many people, provided that the total number of QALYs is the same. The sole aim is to maximise the health of the community with the resources available, where this is measured by the sum of unweighted QALYs. A number of studies suggest, however, that the social value of an additional QALY may differ depending on how many QALYs a person receives: QALYs are valued less when they are concentrated among few people, and valued more when they are distributed more widely.

For example, in a study in Norway, Olsen surveyed 90 undergraduate economics students and 44 doctors using the PTO, the aim being to derive implicit social weights for health improvements with different durations (Olsen 1994). Subjects were asked to state how many people (p) each of whom would gain relief from a "chronic state of dysfunction and periodic pain" for 20 years, they considered as good as 100 people gaining relief from the same health state for t years. For the doctors t was set at 5 years; the students were randomly assigned one of two versions of the questionnaire, one with t set at 5 years, the other with t set at 10 years. Subjects were given the following options for p when t was set at 5 years: 25, 30, 40, 50, 70, 90. When t was set at 10 years they had the following options: 50, 60, 70, 80, 90, 100. Table 22 shows the implicit weights Olsen derived from the stated indifference values. The majority of respondents, 77 per cent, chose to make a trade-off between the duration of the benefit and the number of people receiving it. When t was set at 5 years – when subjects were comparing a 5-year benefit with a 20-year benefit – 8.3 per cent of respondents thought that the number of people receiving the benefit should be the same. This increased to 17.4 per cent when t was set at 10 years – that is, when the relative difference in duration of benefit was smaller. The results indicate that duration of benefit matters. However, they do not support the assumption implicit in conventional CEA that the social value of a QALY gained is independent of the number of QALYs a person receives. Rather, they suggest that QALYs are valued less when they are concentrated among few people, and valued more when they are distributed more widely. The results also suggest that the preference for spreading benefits is greater when the relative difference in the benefits accruing to the two groups is smaller.

It should be noted that the discounting of additional QALYs is different from the time discounting of future QALYs. Time discounting implies that a QALY gained in the future is valued less by society than a QALY gained now. This is so whether those QALYs are gained by the same

individual or different individuals, and it is irrelevant how many QALYs that individual, or those individuals, will receive. By contrast, the study by Olsen suggests that more and more QALYs accruing to the same individual or group are valued less and less by society. Duration discounting, in contrast to time discounting, is thus a function, inter alia, of the number of QALYs an individual or group will receive. On this difference see (Nord 1999, pp. 55-57).

Table 22. Implicit "Social Weight Rates" (%)

Sample	Students		Doctors	Sum
t	10	5	5	5, 10
Median	14.9	13.0	13.0	13.0
Mean	14.6	16.1	16.6	15.7
Standard Error	1.2	2.3	2.8	1.2
Valid *N*	46	44	40	130

Source: (Olsen 1994)

In another study, Nord, Street and colleagues sought to quantify the preference for duration of effect among the Australian population by asking subjects to choose between two highly specialised hospital units (Nord, Street et al. 1996). Unit A would allow the treatment of more patients than unit B, but the patients in unit B would derive a greater individual benefit, either in terms of life-expectancy after treatment or quality of life. Assuming that they themselves might need treatment one day, subjects were asked how many people receiving the greater individual benefit they considered equivalent to the greater number of people receiving the smaller individual benefit. They found, for example, that saving the lives of 10 patients for 1 year was considered equivalent to saving the lives of 3.5 patients for 5 years (see Table 23). Contrary to

Table 23. Life Extending Interventions (Including Self Interest)

Years extended		Equivalent numbers of people helped			
Group A	Group B	A		B (median)	IQR (B)
1 year	5 years	10	=	3.5	1.5-6.6
5 years	10 years	10	=	6.5	5.0-9.1
10 years	20 years	10	=	7.0	5.5-9.0
20 years	30 years	10	=	8.5	7.5-9.5

Source: (Nord, Street et al. 1996)

Table 24. Health Improving Interventions (Including Self Interest)

Years enhanced		Equivalent numbers of people helped			
Group A	Group B	A		B (median)	IQR (B)
1 year	5 years	10	=	2.5	1.5-6.0
5 years	10 years	10	=	6.8	4.9-9.1
10 years	20 years	10	=	6.5	5.5-9.0
20 years	30 years	10	=	7.5	6.4-9.1

Source: (Nord, Street et al. 1996)

the expectations of pure health maximisation, treating fewer people with a longer life expectancy was regarded as equally valuable as treating more people with a shorter life expectancy, even though the total health gain in the former case was greater. In general, Nord, Street and colleagues found that "there is a marked discounting of additional life years gained in unit B, both

in the context of life saving and health improvement" (Nord, Street et al. 1996, p. 107). The results obtained for health improving rather than life saving treatments are given in Table 24.

In a Swedish study, Johannesson and Gerdtham observed the same discounting of additional QALYs (Johannesson and Gerdtham 1996). They asked participants to choose between two societies, A and B, that differed with respect to the remaining life-expectancy (in full health) of people in two equi-numerous groups, 1 and 2. Subjects were told they faced a 50 per cent chance of belonging to either group. That is, they were effectively placed behind a "veil of ignorance". Participants were then divided into two sub-samples. For half of the participants, Group 1 in society A had 20 remaining life-years and Group 2 had 10 remaining life-years – that is, the relative difference in remaining life-years was 100%. For the other half, Group 1 in society A had 16 remaining life-years and Group 2 had 10 remaining life-years – that is, the relative difference in remaining life-years was 60%. In both sub-samples, the number of remaining life-years in Group 1 in society B was decreased by 3 (20–3 = 17 in one case, 16–3 = 13 in the other). The number of remaining life-years in Group 2 in society B was increased randomly by 0, 1, 2, or 3 life years. These four increases for Group 2 represent four different possible trade-offs between Group 1 and Group 2: namely, 0.00, 0.33, 0.67 and 1.00. The percentage of individuals choosing society A for these four trade-offs is shown in Table 25. The first row indicates, for example, that when the relative difference in remaining life-years was 60 per cent, and when the trade-off between Group 1 and Group 2 in society B was 0.33, 80 per cent of subjects chose society A. It is noticeable that the percentage of respondents choosing society A decreases as the trade-off increases. In the limit, when society A offers 20 remaining life-years for Group 1 and 10 remaining life-years for Group 2 (total = 30 life-years), and society B offers 17 remaining life-years for Group 1 and 13 remaining life-years for Group 2 (total = 30 life-years) – that is, when the total number of life-years is the same and no trade-off is required - no one chose society A. In this case, the more equitable distribution of society B could be "purchased" at no cost (in efficiency) at all. "According to the results," Johannesson and Gerdtham state, "the respondents are willing to give up 1 QALY in the group with more QALYs to gain 0.45 QALYs in the group with fewer QALYs" (Johannesson and Gerdtham 1996, pp. 365-366). This study detected no significant difference when the relative difference in remaining life-years was 100% and when it was 60%. "It was thus as if the respondents focused on inequality as such rather than the size of the inequality" (Johannesson and Gerdtham 1996, p. 366).

Table 25. The percentage of individuals choosing society A for the four different marginal trade-offs of QALYs between group 1 and 2 and the two relative differences in QALYs between group 1 and 2

	Marginal trade-off			
Relative difference	0.00	0.33	0.67	1.00
60%	70	80	20	10
100%	100	60	30	0
Full sample	85	70	25	5

Source: (Johannesson and Gerdtham 1996)

In a similar study, Andersson and Lyttkens asked 225 economics students in Sweden to choose from behind a "veil of ignorance" between two societies, A and B (Andersson and Lyttkens 1999). Each society comprised two groups of people whose citizens had different total life expectancies (rather than remaining life expectancies, as in the study by Johannesson and Gerdtham). Subjects were divided into two sub-samples. For one, the life expectancy of Group 1 in society A was 88 years and for Group 2 was 66 years. For the other, the life expectancy of Group 1 in

society A was 82 years and for Group 2 was 68 years. The relative difference between Group 1 and Group 2 in society A was therefore greater. In both sub-samples, the life expectancy of Group 1 was reduced by 6 years in society B and the life expectancy of Group 2 was increased by 2, 4, or 6 years. These represent the same trade-offs tested by Johannesson and Gerdtham: namely, 0.33, 0.67 and 1.00. The results are given in Table 26, and show two things. First, that the proportion of subjects choosing the more egalitarian society, B, increased as the sacrifice of 6 years for Group 1 "bought" more years for Group 2. Second, that the proportion of subjects choosing the more egalitarian society was greatest when the relative difference between the two groups was greatest. By contrast, Johannesson and Gerdtham found that the relative-difference variable had no significant impact on subjects' choices. Overall, approximately 71 per cent of subjects choose society B, "which indicates fairly egalitarian preferences" (Andersson and Lyttkens, p. 372). Whereas Johannesson and Gerdtham found a median marginal trade-off of 0.45 years for the worst off in exchange for one year lost by the better off, Andersson and Lyttkens found a median marginal trade-off of 0.35 when the relative difference was small and 0.11 when the relative difference was large.

It is worth emphasising that subjects in this study, as in the study by Johannesson and Gerdtham, and many others, collectively express neither purely egalitarian preferences from behind a veil of ignorance – preferences to "maximin" - nor purely utilitarian preferences – preferences to maximise. Rather subjects are willing to trade-off efficiency against equity depending on a series of variables (such as relative difference in life expectancy). Another difference between this study and the one by Johannesson and Gerdtham is that in this study the effects of risk and uncertainty were tested, by informing one group of subjects that they had a 50 per cent chance of belonging to Group 1 or Group 2 (as in the Johannesson and Gerdtham study), and the other that nothing is known about the probability of belonging to one group or the other. However, no significant difference was noted in relation to this variable: "the impact of uncertainty is almost completely indiscernible in the data" (Andersson and Lyttkens, p. 373).

Table 26. Proportion of Subjects Choosing Society B

Trade-off	Relative difference	
	Small	Large
0.33	0.50	0.60
0.67	0.68	0.81
1.00	0.80	0.89

Source: (Andersson and Lyttkens 1999)

When total QALYs are equal, respondents prefer to give a few QALYs to many people rather than give many QALYs to a few people. As the study by Johannesson and Gerdtham indicates, the discounting of additional health benefits can also be observed when the two groups of potential recipients are the same size. In fact, in an early study, Kahneman and Varey observed this in the case of individuals (Kahneman and Varey 1991). They asked subjects to play the role of doctor, and allocate a daily supply of forty-eight pain-relieving pills between two patients. Patient A's metabolism was said to be such that it takes three pills to give him one hour of relief, whereas patient B's metabolism was such that it takes one pill to give him one hour of relief. An obvious option would be to give 24 pills to patient B, who is most "efficient" at converting pills into pain relief, and give the remainder to A, who would gain 8 hours of relief. This would both maximise pain relief (32 hours total), and involve splitting the pills equally between A and B. Kahneman and Varey report, however: "most of our respondents opted to equalize hours of pain, and rejected an equal split of pills or a compromise solution; 77% of respondents endorsed an equal-pain allocation in this question" (Kahneman and Varey 1991, p. 156). This supports the idea that

the social value of an additional QALY may differ depending on how many QALYs a person receives (Yaari and Bar-Hillel 1984).

The assumption that subjects invariably favour more equitable distributions of QALYs is too simple however. Other evidence suggests that people are not willing to sacrifice overall QALY gains to achieve a more equitable distribution if the individual benefits accruing to the larger number of people fall below a (possibly context-dependent) threshold. For example, in a Canadian study, Choudhry, Slaughter and colleagues asked senior health officials to choose between two health programmes, A and B, that would provide different individual benefits, measured in terms of increases in life-expectancy, to different numbers of patients, and in one case to patients of different ages also (Choudhry, Slaughter et al. 1997). The programmes were described as similar in all other respects. In particular, the aggregate life years each programme would produce was the same. As Table 27 indicates, when 1000 people would gain 20 years and 4000 people would gain 5 years a large number of respondents, 35.0 per cent, were either unable to decide between the programmes or had no preference. Of those who were able to decide, a majority (53.8 per cent) chose to spread the gains rather than concentrate them among the few. However, when 500 people would gain 20 years and 10,000 people would gain 1 year - when the smaller individual benefit was decreased and the difference between the individual benefits was increased - a majority of respondents, 55.8 per cent, chose to concentrate rather than spread the benefits. This shows that the same subjects may have preferences for spreading or concentrating health benefits depending on the size of the benefits in question. It is unclear from this study whether it was the difference in absolute size of the smaller benefit, that is

Table 27. Decision-makers' Responses to Programs where Beneficiaries Differ in Distributions of Benefits – All Respondents

	Participants (%, *n*)			Comments[a]
	Favoring Program A	Favoring Program B	Unable to decide/ no preference	
All respondents	55.8 (47)	18.8 (15)	22.5 (18)	Program A: 500 people gain 20 years of life-expectancy
Respondents able to decide	75.8 (47)	25.2 (15)	N/A	Program B: 10,000 people gain 1 year of life-expectancy
All respondents	30.0 (24)	25.0 (28)	35.0 (28)	Program A: 1000 people gain 20 years of life-expectancy
Respondents able to decide	46.2 (24)	53.8 (28)	N/A	Program B: 4000 people gain 5 years of life-expectancy
All respondents	53.8 (43)	21.3 (17)	25.0 (20)	Program A: 500 30-year-olds gain 20 years of life-expectancy
Respondents able to decide	71.7 (43)	28.3 (17)	N/A	Program B: 2000 50-year-olds gain 5 years of life-expectancy

Source: (Choudhry, Slaughter et al. 1997)

a   The programs are equal in all respects except those outlined in the table; increases in life-expectancy and ages of patients affected by a given disease are expressed as averages.

5 years compared to 1 year, or the change in the relative difference between the smaller benefit and the larger benefit, from 19 years to 15 years, that was responsible for subjects' variable preferences. The final two rows of Table 27 again show a preference for concentrating benefits, but in this case the results are confounded by the different ages of the potential beneficiaries. The authors conclude: "health resource allocation must explicitly consider not only net and aggregate health gains and the incremental costs to achieve them, but also the distribution of

benefits (and harms) and the implications of that distribution with respect to societal values such as equitable access to care" (Choudhry, Slaughter et al. 1997).

The idea that mild health states sit on a "plateau" at the top of the value scale, and are considered equivalent to full health for trade-off purposes, was tested by Prades and Lopez-Nicolás using the PTO (Prades and Lopez-Nicolás 1998). In their two-part study, participants were initially asked to evaluate, from a personal point of view, six EuroQol health states using a VAS:

11121
I have no problem in walking about
I have no problem with selfcare
I have no problem with performing my usual activities
I have moderate pain or discomfort
I am not anxious or depressed

12111
I have no problem in walking about
I have some problems washing or dressing mysel
I have no problem with performing my usual activities
I have no pain or discomfort
I am not anxious or depressed

21232
I have some problems in walking about
I have no problem with self care
I have some problems with performing my usual activities
I have extreme pain or discomfort
I am moderately anxious or depressed

22322
I have some problems in walking about
I have some problems with self care
I am unable to perform my usual activities
I have moderate pain or discomfort
I am moderately anxious or depressed

22233
I have some problems in walking about
I have some problems washing or dressing myself
I have some problems with performing my usual activities
I have extreme pain or discomfort
I am extremely anxious or depressed

33322
I am confined to bed
I am unable to wash or dress myself
I am unable to perform my usual activities
I have moderate pain or discomfort
I am moderately anxious or depressed

In the second part of the study subjects were asked to adopt a social perspective and choose between two equally costly health programmes. Programme 1 would improve a neo-natal intensive care unit (NICU) and save the lives of 10 newborns each year. Programme 2 would fund a new medicine, without which patients would remain in health state A for the rest of their lives. Health state A was replaced by one of the six EuroQol health states. Subjects were asked

how many patients would have to be treated under Programme 2 each year for them to be indifferent between the two programmes. (Subjects who indicated they would always give priority to the NICU were taken as implicitly giving the value 1 to health state A (equivalent to full health). Subjects who chose to give priority to Programme 2 were taken as implicitly giving the value 0 to health state A (equivalent to being dead).) The results are given in Table 28. Column (1) shows the EuroQol health states arranged in order of increasing severity, as judged by subjects using the VAS. Column (4) shows the derived values of these health states, as judged by subjects using the PTO. It is noticeable that the derived values decrease less rapidly than the interview

Table 28. Results Obtained with The Visual Analogue Scale (VAS) and The Person Tradeoff (PTO) (Medians; n = 83)

	Visual Analogue Scale		Person Tradeoff	
	(1)	(2)	(3)	(4)
EuroQol Health State	Interview Value (Median)	Full-health Equivalent[a]	Full-health Equivalent[b]	Derived Value[c]
111211	0.80	5	No tradeoff[d]	1.00
12111	0.75	4	No tradeoff	1.00
211232	0.40	1.7	22.5	0.95
22322	0.30	1.4	10	0.90
22233	0.18	1.2	6	0.83
33322	0.10	1.1	2.3	0.56

Source: (Prades and Lopez-Nicolás 1998)

a   Number returned to full health equivalent to one saved life using VAS values obtained from column (1). E.g. $(1 - 1/5) = 0.80$; $(1 - 1/4) = 0.75$.

b   Number returned to full health equivalent to one saved life directly elicited from interviews.

c   PTO values obtained from column (3). E.g. $1 - 1/22.5 = 0.95$; $1 - 1/10 = 0.90$.

d   More than 50% of the interviewees said that they would always choose to save the lives of ten newborns even if the number of people cured from mild health problems was very high (e.g. 100,000 people).

values. The "plateau effect" is one consequence of this. Relatively mild health states have little "trade-off value". In fact, the first two health states received a rating of "no tradeoff" (column (3)), indicating that more than 50 per cent of subjects would give priority to the NICU, which involves saving lives, even if large numbers of people - e.g. 100,000 - could be relieved of these health states. (Using regression analysis the authors also detected a plateau effect in the middle of the range (p. 291)). In other words, although people generally favour more equitable distributions of QALYs, they do not prefer large numbers of people getting relatively minor benefits – for example, being relieved of "moderate pain or discomfort" - to fewer people gaining significant benefits - for example, life-saving treatment.

In Norway, Olsen also sought to determine when respondents prefer to allocate a small benefit to many people, and when instead they prefer to allocate a large benefit to a few people (Olsen 2000). He calls those who prefer to distribute a given sum of QALYs so that each of many receives a small benefit, rather than each of few receiving a large benefit, *diffusors*, because of their preference for spreading gains. By contrast, those who prefer that the few receive substantial improvements, rather than the many receiving insignificant improvements, he calls *concentrators*. Those who are concerned solely with ensuring the largest possible health gains and are indifferent about whether those gains are diffused or concentrated, he calls *maximisers*. Olsen initially asked subjects to choose between a programme, X, that would relieve 100 people of a disease (involving "reduced functioning and periodic pain") for *t* years each, and another

programme, Y, that would relieve 5 people of the same disease for 20 years each. Subjects could also indicate whether they considered increased funding to X or Y "equally good" or "don't know". Three different versions of the questionnaire were administered with t taking the values 1, 2 and 5 years. Subjects who indicated a definite preference for either programme X or Y were then given a second task. From among the following six alternatives - 2, 4, 10, 30, 60, 100 - they were asked to choose the number of people they consider would have to be relieved of symptoms for 20 years in order for programme Y to be as good as programme X. Table 29 shows the percentage of *diffusors*, *concentrators*, *maximisers*, and "don't know" in each population sub-sample ($t = 1$, $t = 2$, $t = 5$). Observing the different results for the sub-samples, Olsen concludes that "the threshold quantity does not refer to an absolute value independent of the comparator. It depends on how small the smaller gain is relative to the larger gain" (Olsen 2000, p. 544). Table 29 indicates that *diffusors* outnumbered *concentrators* in all three sub-samples, and that the proportion of *concentrators* compared with *diffusors* was greatest when t was smallest: 21 *concentrators* compared with 30 *diffusors* when $t = 1$. "This lends support to the hypothesis that a threshold value may exist, and that it lies somewhere around $t = 1$ when $T = 20$" (Olsen 2000, p. 547). It should be noted, however, that only 68 per cent of those who expressed a preference for concentrating gains completed the second task, and quite a few of these, 69 per cent, gave inconsistent answers. This means that the results have to be treated with caution. Nonetheless, Olsen's summary conclusion seems justified: "The study referred to in this paper challenges the conventional wisdom in the literature that there is a universal preference towards a more equal distribution of health gains than the health maximising one" (Olsen 2000, p. 549).

Table 29. Percent Distribution of Subjects in the Different Versions of the Questionnaire

	$t = 1$	$t = 2$	$t = 5$	Total	Indifference[b]	Inconsistent
Diffusors $p_i > Pt/T$[a]	30	38	39	36	94% of 258	1% of 242
Concentrators $p_i < Pt/T$	21	14	18	18	68% of 126	69% of 86
Maximisers $p_i = Pt/T$	38	38	37	38		
Don't know	10	9	6	9		
	100%	100%	100%	100%		
N	231	258	227	716		

Source: (Olsen 2000)

    a   Where p_i = subject's indifference point, $P = 100$ persons and $T = 20$ years
    b   The percentage of subjects who completed the second task

Rodríguez-Míguez and Prades also used the PTO to analyse how the size of individual health gains influences social preferences for concentration and dispersion of total health gains (Rodríguez-Míguez and Prades 2002). Participants in this study were presented with a number of health programmes, which differed in the number of patients that would benefit from their implementation (p), and the number of years in full health each patient would receive (t). More particularly, the following five variations were used: (1, 100), (2, 50), (5, 20), (20, 5), (50, 2), where the first element is t and the second element is p. In terms of total QALYs gained, all of these programmes are equal. However, their distributive implications are obviously different. For each programme, subjects were asked to specify how many patients (p^*) would have to receive a ten-year benefit for the two programmes to be of equivalent social value. (All patients were assumed to be twenty years old.) The mean values of p^*, designated $\bar{p}^*$, are given in Table 30. From this data the authors calculated the preference for concentration and dispersion of benefits among those participating in the study. Table 31 shows that, contrary to the assumption of a

Table 30. Assessment of Health Programmes

Health gain, t (years)	Number of patients, p	$\overline{p}^{*a}$ (t-test)[b]
10	10	10
20	5	9.23 (-1.90)
5	20	8.93 (-2.17)
2	50	7.49 (-2.99)
50	2	7.40 (-5.39)
1	100	6.74 (-3.49)

Source: (Rodríguez-Míguez and Prades 2002)

   a   Number of patients who would have to receive a 10 life-year increase in order that this programme be indifferent to the (t, p) programme

   b H_0: $\overline{p}^*$ = 10; H_1: $\overline{p}^*$ / = 10; n = 45

Table 31. Distributive Preferences Based on Individual Gain (Years)[a]

(Gain, patients)	(2, 50)	(5, 20)	(10, 10)	(20, 5)	(50, 20)
(1, 100)	C*	C**	C**	C**	C*
(2, 50)		C**	C**	C**	S
(5, 20)			C**	C	S**
(10, 10)				S*	S**
(20, 5)					S**

Source: (Rodríguez-Míguez and Prades 2002)

   a   C: preferences for concentrating; S: preferences for spreading; ** significative at 5% level; * significative at 10% level

universal preference for spreading health gains, "there are cases where subjects, on average, prefer to concentrate …, and cases where they prefer to spread out health gains…" (Rodríguez-Míguez and Prades 2002, p. 48). For example, when comparing (1, 100) with (20, 5), the number of patients who will benefit from the former programme is obviously greater (100 > 5), the total QALY gains are identical (100 years), and yet the number of patients who would have to receive 10 years to match the first programme, 6.74, is less than the number of patients who would have to receive 10 years to match the second programme, 9.23. This reveals a preference for *concentrating* benefits. By contrast, when comparing (20, 5) with (50, 2), the number of patients who will benefit from the former programme is greater (5 > 2), the total QALY gains are identical (100 years), and the number of patients who would have to receive 10 years to match the first programme, 9.23, is greater than the number of patients who would have to receive 10 years to match the second programme, 7.40. This reveals a preference for *spreading* benefits. Analysing these preferences, Rodríguez-Míguez and Prades conclude: "when gains are under 9.1 years, participants on average prefer to concentrate those gains, but if the gains are greater than this *threshold* gain they prefer to distribute them" (Rodríguez-Míguez and Prades 2002, p. 49). It should be noted that the preferences of 26 per cent of subjects in this study were excluded from the analysis - those who did not make trade-offs, and that the convenience sample of undergraduate students may not have been representative. Nevertheless, like the previous studies in this section, this one suggests that the assumption implicit in conventional CEA - that each additional QALY has the same social value independent of the number of QALYs received - lacks empirical support.

Table 32. Studies of Concentration and Dispersion

Study	Country	Number of Subjects	Type of Subjects	Elicitation Method
(Olsen 1994)	Norway	134	Students/Physicians	PTO
(Nord, Street et al. 1996) G1[a]	Australia	44	GenPop[c]	PTO
(Nord, Street et al. 1996) G2[b]	Australia	42	GenPop	PTO
(Johannesson and Gerdtham 1996)	Sweden	80	Students	DCM[d]
(Choudhry, Slaughter et al. 1997)	Canada	80	MOH[e]	DCM
(Prades and Lopez-Nicolás 1998)	Spain	83	Students/Physicians	PTO
(Andersson and Lyttkens 1999)	Sweden	225	Students	DCM
(Olsen 2000)	Norway	716	GenPop	PTO
(Rodríguez-Míguez and Prades 2002)	Spain	61	Students	PTO

a   G1 = Group 1
b   G2 = Group 2
c   GenPop = General Population
d   DCM = Discrete Choice Method
e   MOH = Senior Officials in the Ministry of Health of Ontario

AGE-RELATED SOCIAL PREFERENCES

Empirical studies suggest that many people are willing to give certain age groups priority in the competition for limited health care. Age-related preferences are not homogeneous however. Different studies appear to pick up different things. Using the tripartite division of Tsuchiya (Tsuchiya 1999), some studies appear to measure "utilitarian ageism" (Nord, Street et al. 1996, p. 103) or, more specifically, "health-maximisation ageism" (Tsuchiya, Dolan et al. 2003, p. 688): the preference for the treatment of younger patients because they are expected to live longer after treatment. This form of ageism is compatible with prioritising programmes and services on a cost-per-QALY basis (see Table 33). In contrast, other studies appear to measure "productivity ageism": the preference for some age groups because of their level of productivity at work or home. Yet other studies appear to measure "egalitarian ageism" (Nord, Street et al. 1996, p. 103) or "fair-innings ageism" (Tsuchiya, Dolan et al. 2003, p. 688): the preference for younger patients because of an aversion to inequality in age of death. These latter two forms of ageism imply a rejection of the strategy of prioritising health programmes and services strictly on a cost-per-QALY basis.

In some studies a mixture of age-related preferences is probably responsible for observed choices. For measurement and computational purposes, however, it is important that these different forms of ageism be kept separate. As Richardson points out, for example, the application of age weights to QALYs raises the possibility of double counting. If age weights reflect peoples' (correct) expectations that the quality of life will decline with age, or that people in their middle years contribute to both their family and to the economy, *and these factors are already taken into account in a QALY study*, then age weights will not be capturing social value not included elsewhere (Richardson 2002c, p. 667). To the extent that the different forms of ageism are not carefully distinguished the danger of such double counting increases.

Table 33. Categorisation of the Main Types of Ageism (after Tsuchiya, Street et al. 1996)

	Efficiency-based age weighting		Equity-based age weighting
	Compatible with standard QALY paradigm[a]	Incompatible with standard QALY paradigm	
Utilitarian ageism (Health-maximisation ageism)	✓		
Productivity ageism		✓	
Egalitarian ageism (Fair-innings ageism)			✓

a   According to the standard paradigm, QALYs are of equal social value no matter who gets them.

In an early study, Charny, Lewis and colleagues asked a representative sample of the general population in Wales to choose between patients of different ages needing life-saving treatment (Charny, Lewis et al. 1989). The results are summarised in Table 34. They show a preference for treating younger patients, except that 33.8 per cent of subjects chose an 8-year old over a 2-year old, compared with 20.7 per cent who chose the younger patient. However, the greatest majority (45.5 per cent) could not choose between these two patients, despite being encouraged to do so. In reporting their qualitative study of the tape-recorded interviews, Charny, Lewis et al. mention, among the reasons why subjects preferred the five-year old to the seventy-year old, "future economic productive capacity", and among the reasons why they preferred the thirty-five-year old to the sixty-year-old, "greater responsibilities". This suggests productivity ageism, although the study does not allow this inference to be drawn with any confidence. It is likely that utilitarian ageism and egalitarian ageism also influenced choices.

Table 34. Percentage of Population Registering Choices

Person Y	Person Z	Easy Y	Diff Y	No Choice	Diff Z	Easy Z
5-yr-old	70-yr-old	69.9	23.9	5.1	0.5	0.6
35-yr-old	60-yr-old	37.9	42.9	13.2	3.6	2.3
2-yr-old	8-yr-old	3.0	17.7	45.5	28.7	5.1

Source: (Charny, Lewis et al. 1989)

In an Australian postal survey, Nord, Richardson et al. presented members of the general public with the following three options (Nord, Richardson et al. 1995a):

1. Among people with life threatening illnesses, younger patients should have some priority over older patients.

2. People should have the same priority with respect to life saving treatment, unless they are very old.

3. People should have the same priority with respect to life saving treatment, no matter what their age is.

Of respondents, 17.6 per cent chose option 1, 40.5 per cent chose option 2, and 41.9 per cent chose option 3. This indicates a tendency towards equal consideration of the young and the old. Supporting this, "those choosing to discriminate expressed greater difficulty making their choice than did those opting for equal priority" (Nord, Richardson et al. 1995a, p. 1432). When presented with a similar choice about treatment intended to improve quality of life rather than

save life, the results were similar, with slightly more subjects preferring to give priority to the young (see Table 35). As in the Charny, Lewis et al. study, when asked to choose between a "young child" and a "newborn infant" 44.2 per cent chose the young child, 54.7 per cent chose not to discriminate, and only 1.2 per cent chose the newborn infant. No attempt was made in this study to discern whether utilitarian ageism, productivity ageism, or egalitarian ageism lay behind subjects' choices.

Table 35. Percent Response to Age issues and Difficulty of Choice

Issue	Option	Percent choosing each option	Difficulty of Choice (%)		
			Very Difficult	Slightly Difficult	Not Difficult
AGE/LIFE	Favour young	17.6	28.0	35.5	36.6
	Against very old	40.5	26.2	48.6	25.2
	Equal priority	41.9	16.7	24.0	59.3
AGE/QoL	Favour young	21.5	36.0	35.1	28.8
	Favour old	2.9	13.3	46.7	40.0
	Equal priority	75.6	15.3	23.8	60.9
NEWBORN	Favour young child	44.2	40.1	30.4	29.5
	Favour newborn	1.2	66.7	16.7	16.7
	Equal priority	54.7	43.1	16.0	40.9

Source: (Nord, Richardson et al. 1995a)

In a British study, Bowling, Jacobson and colleagues asked the general public and doctors to rank 16 health services in order of importance, by placing four of them in the category "essential", four in the category "most important", four in the category "important", and four in the category "less important" (Bowling, Jacobson et al. 1993). Three of these services allowed for the expression of age-related preferences, and are shown in Table 36. "Treatments for children with life threatening illnesses (e.g. leukaemia)" was ranked highest by both sub-samples of the general population, suggesting a preference for younger patients. By contrast, "long stay care (e.g. hospital and nursing homes for the elderly)" was ranked tenth by the community groups and eighth by the random sample of the public. Note, however, that the former explicitly concerns life-saving treatment, whereas the latter does not. It is therefore difficult to discern the extent to which subjects' choices in this study were influenced by ageism sentiments and the extent to which they were influenced by other considerations. Corroborating the Charny, Lewis et al. study (1989) and the Nord, Richardson et al. study (1995a), both the general public and doctors gave a relatively low ranking to "intensive care for premature babies who weigh less than one and a half pounds and are unlikely to survive", compared with "teatments for children with life threatening illnesses (e.g. leukaemia)."

In a subsequent study, Bowling asked 2005 members of the British public to rank in order of importance 12 of the 16 health services included in the 1993 study (Bowling 1996). Four of the services allowed for the expression of age-related preferences, and are shown in Table 37. As in the 1993 study, there is a noticeable preference for the young over the elderly. "Treatment for children with life threatening illnesses" was ranked first again, with "long stay hospital care for elderly people" and "treatment for people aged 75 and over with life threatening illness" being ranked tenth and twelfth respectively. Also like the 1993 study, the preference for young people over the elderly did not extend to a preference for infants over young children, with "intensive care for premature babies who weigh less than 680 g with only a slight chance of survival" being ranked ninth out of twelve.

Table 36. Mean Priority Ranks of the Public and the Doctors (in Order of Priority 1-16, 1=highest, 16=lowest)

	Public community groups	Public random sample	GPs	Consultants	Public health doctors[a]
Treatments for children with life threatening illnesses (e.g. leukaemia)	1	1	5	2	5 (9)
Intensive care for premature babies who weigh less than one and a half pounds and are unlikely to survive	9	11	13	13	10† (15)
Long stay care (e.g. hospital and nursing homes for the elderly)	10	8	3	6*	6 (10)
Number of respondents	322-335	298-305	63-66	112-116	4-6

Source: (Bowling, Jacobson et al. 1993)

*   *P* < 0.0001 (public in comparison with GPs and consultants).

†   This intervention received an equal tenth ranking with treatments for infertility (e.g. test tube babies). The small number of public health doctors, and lack of divergence in their opinions, lead to joint rankings for some interventions. The tenth ranking was the lowest.

a   Because of joint rankings, the public health doctors' scale ran from 1-10 rather than 1-16. Alternative rankings are given in brackets to facilitate comparison.

Bowling also asked respondents whether they agreed or disagreed with the following two statements (among others): (1) "if resources must be rationed, higher priority should be given to treating the young than the elderly" and (2) "high cost technology (for example, transplantation and kidney machines) should be available to all regardless of age." The somewhat contradictory results are given in Table 38. They show, for example, that 50 per cent of respondents either agreed or strongly agreed with the former statement (with 21 per cent neither disagreeing nor agreeing), which is consistent with the ranking of the 12 health services (keeping in mind the low ranking of intensive care for infants with a poor prognosis). However, an even greater majority of respondents, 77 per cent, either agreed or strongly agreed with the latter statement, which appears to be at odds with the lowest ranking given to "treatment for people aged 75 and over with life threatening illness." It is possible, however, that responses to statement (2) reflected attitudes towards the "high cost" of the technology. In general, as in the 1993 study, it is not possible to tell whether utilitarian ageism, productivity ageism, or egalitarian ageism was responsible for subjects' choices in this study.

Myllykangas, Ryynänen et al. conducted a large survey in Finland in which they asked nurses, doctors, politicians, and the general public for their views on a range of health care issues, including the relevance of age in setting health care priorities (Myllykangas, Ryynänen et al. 1996). Two statements on age, and subjects' responses to them, are given in Table 39. A majority of respondents in all groups agreed with the statement that children and young people should be treated before elderly people. However, an even greater majority in all groups disagreed with the statement that expensive examinations or treatments for the elderly should not be subsidised by the community. This expressed some of the ambiguity detected in the 1996 Bowling study: respondents do not want to abandon the elderly, even if their care is expensive, but they give a higher priority to children and young people.

Table 37. Priority Rating of Health Services. Figures are Percentages (Numbers)

	Priority	Treatments for children with life threatening illness	Intensive care for premature babies who weigh less than 680 g with only a slight chance of survival	Long stay hospital care for elderly people	Treatment for people aged 75 and over with life threatening illness	Number of respondents
			Health services			
Priority rank	1	34 (674)	3 (56)	2 (44)	2 (30)	1975
	2	21 (409)	6 (111)	4 (70)	3 (69)	1974
	3	9 (185)	9 (167)	7 (144)	5 (96)	1972
	4	7 (143)	7 (133)	7 (146)	4 (83)	1969
	5	10 (196)	5 (87)	5 (98)	3 (65)	1944
	6	6 (119)	7 (144)	7 (135)	6 (113)	1941
	7	3 (64)	8 (156)	12 (233)	9 (179)	1939
	8	3 (59)	8 (154)	15 (287)	9 (175)	1939
	9	2 (36)	4 (70)	3 (52)	4 (71)	1949
	10	3 (63)	16 (309)	4 (77)	8 (162)	1945
	11	1 (14)	20 (392)	10 (205)	29 (346)	1944
	12	1 (3)	8 (169)	24 (469)	18 (346)	1940
	Mean	3-2	7-7	7-9	8-7	
	Mean p rank	1	9	10	12	

Source: (Bowling 1996)

Table 38. Attitudes about Priorities. Figures are Percentages (Numbers)

Possible answers	High cost technology (for example, transplantation and kidney machines) should be available to all regardless of age	If resources must be rationed, higher priority should be given to treating the young than the elderly
	Statements	
Strongly disagree	2 (32)	5 (94)
Disagree	11 (216)	24 (476)
Neither disagree nor agree	7 (133)	21 (422)
Agree	55 (1092)	40 (776)
Strongly agree	25 (506)	10 (203)
Number of respondents	1978	1971

Source: (Bowling 1996)

Utilitarian Ageism

The studies of dispersion and concentration considered in the previous section show that duration of effect matters to people. This is closely correlated with age, for young people typically have more years left than older people. In the study previously mentioned, Nord, Street and colleagues elicited this preference for duration of effect in its "pure" form by specifying that the ages of the two groups of patients are the same (Nord, Street et al. 1996). When the ages of the two groups are the same it is not possible for subjects to express their "ageist" preferences – their preferences for different age groups. Nevertheless, when subjects choose to save the lives of

younger patients rather than older patients *because* they are expected to live longer after treatment - because of duration of effect - they are giving expression to what Nord, Street et al.

Table 39. Responses to Statements (%)

Statement	Nurses (*n* = 682)	Doctors (*n* = 837)	Politicians (*n* = 1133)	General Public
Expensive examinations or treatments for the elderly should not be subsidised by the community				
Agree	27	15	21	18
No opinion	2	5	3	6
Disagree	70	80	76	76
Children and young people should be treated before elderly people				
Agree	64	55	59	58
No opinion	3	11	6	8
Disagree	32	35	35	34

Source: (Myllykangas, Ryynänen et al. 1996)

call "utilitarian ageism" (1996). On the other hand, when groups are of different ages, and life saving treatment is on offer, some people will die younger than others without treatment. That is, when subjects choose to save the lives of younger patients rather than older patients, their observed preferences are likely to be a mixture of "utilitarian ageism" and "egalitarian ageism" - the preference for younger patients because of an aversion to inequality in age of death. Tsuchiya calls this "hybrid" utilitarian ageism (Tsuchiya 1999, p. 272).

In a Spanish study, Rodríguez and Pinto sought to test whether age-related preferences depend on the size of the health gains (Rodríguez and Pinto 2000). Most studies implicitly assume that age weights for large durations – for example, ten years – can be derived by integrating the age weights for small durations – for example, one or two years. As Rodríguez and Pinto observe, "this is by no means a self-evident assumption" (Rodríguez and Pinto 2000, p. 618). They showed subjects a list of 11 health-care programmes that would produce different health gains (measured in years in full health), t, for different patient age groups, a. Subjects were asked to specify the benefit (in years in full health), t^*, that would have to be gained by 20 year-olds (the reference group) for them to be indifferent between the two programmes ((a,t) and (20, t^*)). The results are given in Table 40, where $\bar{t}^*$ denotes the average of t^*. The first line indicates that a programme that gains 2 years for one-year olds is, on average, equivalent to a programme that gains 0.9 of a year for 20 year-olds. Figure 5 shows the weights of the health gains based on age, derived from the ratios $\bar{t}^*/t$ in Table 40. It can be seen that for 1-year olds the greater the health benefit the higher the value, whereas for 60 year-olds the greater the health benefit the lower the value. Health benefits for 20 year-olds and 40 year-olds are rated similarly when they are of short duration, but receive a lower value for 40 year-olds when they are of longer duration. Rodríguez and Pinto therefore reject the hypothesis "that the relative social value of health gains is constant when comparing people of different ages" (Rodríguez and Pinto 2000, p. 618). However, based on the results in Table 39, they accept the hypothesis that "the relative social value of a health gain t for age a can be estimated as the sum of successive values of small
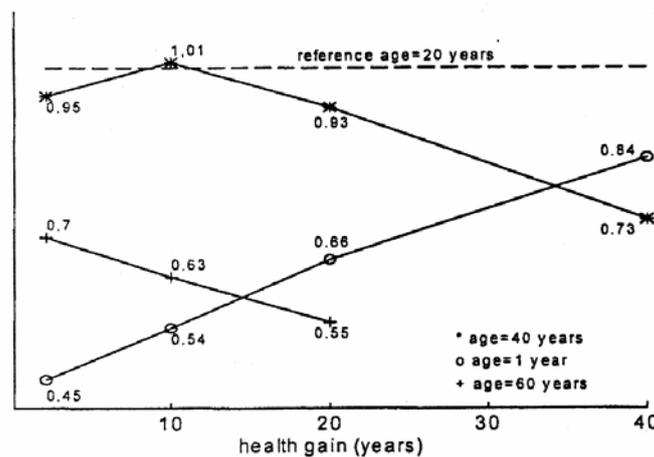
Table 40. Average Social Value of Health Gains (Years) Based on Age

Health gain, t	Patient age, a	Equivalent health gain for 20 year olds, t^* (t-student[a])
2	1	0.9 (-8.41)
	40	1.9 (-0.41)
	60	1.4 (-2.93)
10	1	5.4 (-7.45)
	40	10.1 (0.20)
	60	6.3 (-8.23)
20	1	13.2 (-5.55)
	40	18.6 (-1.03)
	60	10.9 (-10.58)
40	1	33.4 (-3.19)
	40	29.1 (-6.55)

Source: (Rodríguez and Pinto 2000)

a H_0: t^* = t; H1: $t^* \neq$ t.

Figure 5. Weights of Health Gains Based on Age

Source: (Rodríguez and Pinto 2000)

health gains from a to $a + t$" (Rodríguez and Pinto 2000, p. 615). In other words, age-related preferences for particular age groups do not depend on the size of the health gain. The authors confirmed the results by asking subjects to rank cards corresponding to the previously assessed programmes. A high correlation was obtained at both the individual and aggregate levels. They also found high test-retest reliability by asking subjects to repeat some of the tasks after two weeks. No attempt was made in this study to distinguish between utilitarian ageism, egalitarian ageism and productivity ageism: "both efficiency and equity could be influencing the age-related preferences … elicited" (Rodríguez and Pinto 2000, p. 618).

In a US telephone survey, Cropper, Aydede et al. asked members of the general population to choose between two equally-expensive medical programmes, A and B, that would save the lives of patients of different ages. They found, inter alia, that "eight 60-year-olds are judged equivalent to saving one 20-year-old, and seven 60-year-olds are judged equivalent to saving one 40-year-old. By contrast, eleven 60-year-olds are judged equivalent to saving one 30-year old…" (Cropper, Aydede et al. 1994, p. 245). To those surveyed, the age at which lives are saved

matters, but "most respondents do not weight people strictly by their life expectancy" (Cropper, Aydede et al. 1994, p. 259). The result is a hump-shaped function, with the peak at around age 28. Because the potential beneficiaries of the medical programmes would die at different ages if not treated, the results probably reflect a mixture of utilitarian ageism and egalitarian ageism, or what Tsuchiya calls "hybrid" utilitarian ageism. It is possible that productivity ageism also influenced the choices of some participants.

In a follow-up study, Johannesson and Johansson extended the study by Cropper, Aydede et al. They used similar questions to those used in the earlier study, requiring subjects to choose between two programmes, A and B, that would prevent deaths in different age groups. They then went a step further. Using Swedish data on life expectancy at different ages and age-related quality-of-life weights, they estimated the implied number of life-years and QALYs gained among 50-year olds and 70-year olds that are judged to be equivalent to one life-year and one QALY gained among 30-year olds (Johannesson and Johansson 1997b). The results, using logistic regression, are shown in Table 41. They indicate, for example, that 4.6 QALYs gained for 50-year olds is equivalent to one QALY gained for 30-year olds. As in the study by Cropper, Aydede et al., subject's preferences for programme A or B were probably a mixture of utilitarian ageism and egalitarian ageism – that is, "hybrid" utilitarian ageism. And it is not possible to exclude the influence of productivity ageism. Johannesson and Johansson conclude that "the valuation of life-years and QALYs gained seem to depend on the age at which the life-years and QALYs gained occur, in sharp contrast to what is typically assumed in cost-effectiveness analysis" (Johannesson and Johansson 1997b, p. 597).

Table 41. The number of lives, life-years and QALYs saved among 50-year-olds and 70-year-olds that are judged equivalent to saving one life, life-year, and QALY among 30-year-olds

Age	Lives saved	Life-years gained	QALYs gained
50-year-olds	7.7	4.7	4.6
70-year-olds	40.8	11.3	10.5

Source: (Johannesson and Johansson 1997b)

Richardson urges caution in the interpretation of the results of such studies. For example, it is an implication of the study by Johannesson and Johansson that the average value of remaining life years to someone at age 30 is three times greater than the average value of remaining life years to someone at age 50, despite a difference of only 20 years.[7] According to Richardson, "it is probable that survey respondents would not persist with this pattern of responses if these implications were drawn to their attention" (Richardson 2002c, p. 668). It is not that the results should be rejected, but that they should be treated with caution, and subject to re-testing.

---

7. Johannesson and Johansson report that 5 lives saved at age 50 are approximately equivalent to saving one life at age 30. Assuming a life expectancy of 80, Richardson calculates as follows (Richardson 2002c, p. 675):

$$5 \times 50 \text{ year old} = 1 \times 30 \text{ year old}$$

$$5 \times \text{life expectancy of 30} = 1 \times \text{life expectancy of 50}$$

$$150 \text{ years after age 50} = 50 \text{ years after age 30}$$

$$3 \text{ years after age 50} = 1 \text{ year after age 30}$$

Productivity Ageism

The authors of the *World Development Report 1993* used a relative of the QALY, the Disability-Adjusted Life Year (DALY), to measure the burden of disease in various countries (World Health Organization 1993). DALYs are life years weighted not only for quality of life, as are QALYs, but also for age (and include a time discounting factor as well). Figure 6 illustrates the pattern of age weights used in the *Report*. The value for each year of life rises from zero at birth to a peak at age 25 and then declines gradually with increasing age. The result is that, assuming normal quality of life, slightly more DALYs will be lost by the death of a ten year old (36.71 for males, 36.86 for females) than by the death of a newborn (32.34 for males, 32.45 for females), for although the newborn will live longer (assuming standard life expectancy) the extra years of life from 0 to 10 receive a low weighting. This age-weight profile was applied because "[m]ost societies attach more importance to a year of life lived by a young or middle-aged adult than to a year of life lived by a child or an elderly person" (World Health Organization. 1993, p. 213).

Figure 6. Age-weight Function

····· Age weighted ⎯⎯ Uniform age weights
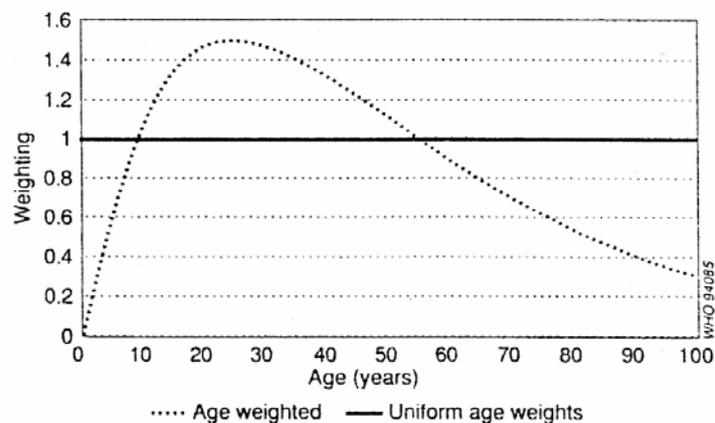
Source: World Bank data

In other words, the age-weight profile used in the *Report* was explicitly meant to reflect productivity ageism: the different levels of dependency experienced at different ages (Murray and Acharya 2002). In particular, "the young, and often the elderly, depend on the rest of society for physical, emotional and financial support" (World Health Organization. 1993, p. 434). Conversely, the providers of this support tend to be those in their twenties and thirties (as child carers) and those in their forties to sixties (as carers of the elderly). Note that the productivity of young and middle-aged adults incorporates more than their economic productivity: "The concept of dependency and social role is broader than formal sector wage productivity and is not linked to total income levels" (Murray 1994, p. 435). The specific age weights used in the *Report*, however, were derived from a convenience group of public health experts, and the factors influencing their decisions is unclear.

In an early study, Williams surveyed the general population in the UK on the importance of good health during different stages of life (Williams 1988). Ten stages of life from infancy to old age were selected and subjects were asked to choose the three at which they considered it most important to keep people well. They were then asked to order these three from most important to least important. The results are given in Table 42. The life stages most often selected were "when bringing up children", "as infants", and "when getting very old". The special importance attached to the child-raising years is evident in many studies of age-related preferences. By contrast, the choice of infants over young children appears anomalous, and can perhaps be

explained by, inter alia, the implicit exclusion of life-threatening illnesses – that is, subjects were not asked in what circumstances they considered it important to keep people *alive*, but in what circumstances they considered it important "to keep people in general well" (Williams 1988, p. 114). The high value placed on good health "when getting very old" may perhaps be explained by subjects attempting to compensate the elderly for their greater susceptibility to illness and injury. From this it cannot be concluded that a HRQoL score of 1.0 is more important to an elderly person than to a child starting school.

Table 42. York Health Evaluation Survey

	Most important		Next most important		Third choice	
	N	*%*	*N*	*%*	*N*	*%*
As infants	103	27.3	36	9.5	25	6.6
When starting school	24	6.4	16	4.2	25	6.6
When starting work	14	3.7	37	9.8	42	11.1
When starting up home for the first time	10	2.7	21	5.6	17	4.5
When bringing up children	124	32.9	102	27.1	42	11.1
When the peak of their earning power	15	4.0	19	5.0	23	6.1
When looking after elderly relatives	19	5.0	44	11.7	38	10.1
When just having retired from work	16	4.2	36	9.5	63	16.7
When coping with the death of a husband or wife	23	6.1	31	8.2	45	11.9
When getting very old	27	7.2	33	8.8	53	14.1
(Unusable responses)	(2)	(0.5)	(2)	(0.5)	(4)	(1.1)
Total	377	100.0	377	100.0	377	100.0

Source: (Williams 1988)

In a study of students (average age 23) and elderly people (average age 78) in the Netherlands, Busschbach, Hessing et al. sought to elicit age-related preferences by means of a TTO exercise (Busschbach, Hessing et al. 1993). Normally, subjects are asked to adopt the "self-interest" perspective in TTO exercises – that is, they are asked to indicate how much of *their own life* they would be willing to give up for an improvement in health. In this study by Busschbach, Hessing

Figure 7. The Ratio of the Utility of Health at Age 35 and the Utility of Health at Ages 5-70
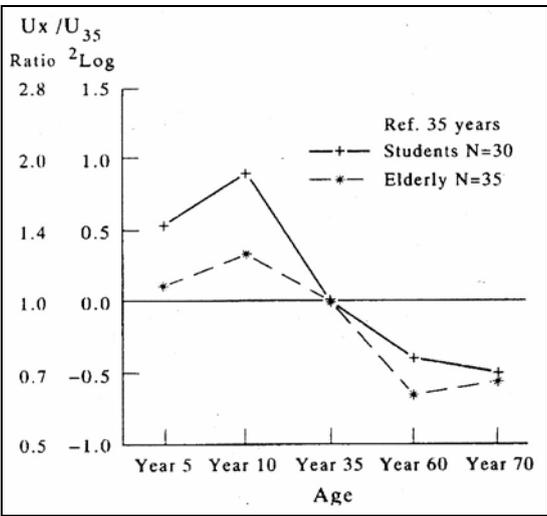
Source: (Busschbach, Hessing et al. 1993)

et al., subjects were asked to compare different periods in ill health (waiting for a kidney transplant) for two patients differing (apart from waiting time) only in their ages. If a subject indicated that it is preferable for a 70 year-old to wait two years for a transplant than for a 35-year old to wait the same time, the period of time the 35-year old would have to wait was progressively halved until the subject changed his/her mind. Subjects were asked to assume that the patients had been ill for the same time, that they would return to normal health after a transplant, and that they would both reach the same age. Participants performed the task twice, once using the age of 35 as the reference point, and once using the age of 60 as the reference point. The results using age 35 are shown in Figure 7. The graph indicates that being in good health at age 10 is valued twice as much as being in good health at age 60 (ratio 2.0). Similar results were obtained using 60 as the reference point (see Figure 8).

Figure 8. The Ratio of the Utility of Health at Age 60 and the Utility of Health at Ages 5-70
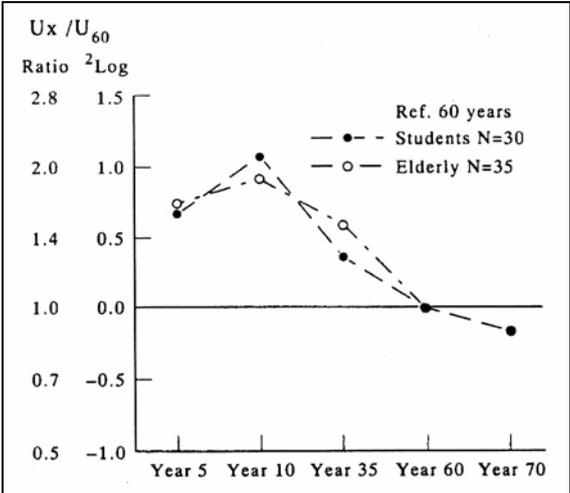
Source: (Busschbach, Hessing et al. 1993)

This pattern of age weights resembles that used in the *World Development Report 1993*. The value of good health at age 10 is valued more highly than good health at age 5, but after age 10 the utility of health gradually decreases. It appears that this study is picking up productivity ageism – that is, the social value of health at different ages - since each patient will receive a transplant eventually (the only differences are the waiting periods and the ages of the patients) and each is expected to reach the same age regardless of the time of transplantation. As Tsuchiya notes, however, any two patients of different ages, even if they are expected to reach

Table 43. Percent Response to Age Issues and Difficulty of Choice

Issue	Option	Percent choosing each option	Difficulty of Choice (%)		
			Very Difficult	Slightly Difficult	Not Difficult
CHILDREN	Favour parents	33.4	30.0	34.7	35.3
	Equal priority	66.6	10.0	14.7	75.2

Source: (Nord, Richardson et al. 1995a)

the same age, also differ in the length of time they have enjoyed up to now and the length of time still to come. If this is influencing subjects' choices, the results may not reflect pure productivity ageism, but may also reflect preferences connected with the *number* of life years prior to or

posterior to the health intervention (organ transplant), and thus "the actual numbers elicited need to be interpreted with care" (Tsuchiya 1999, p. 272).

Qualified support for one aspect of this age-weighting profile – specifically, the public preference for giving priority to parents with dependent children – can be seen in a study by Nord, Richardson et al. (1995a). They found that one-third of subjects would give priority to parents in the allocation of limited health care. On the other hand, two-thirds of subjects would give equal priority to parents and non-parents (see Table 43). Similar results were reported by Dolan, Cookson et al. (1999). They found that 33.4 per cent of respondents would give higher priority to those with children, and 66.6 per cent thought they should be given the same priority. Interestingly, after deliberation, the proportion of respondents who would give higher priority to those with children fell to 20 per cent, and 77 per cent thought they should be given the same priority.

Egalitarian Ageism

The age weighting used in the *World Development Report 1993*, and picked up in the study by Busschbach et al., reflects the relative social value of life at different ages. The life years of all 30-year olds therefore receive the same weighting, as do the life years of all 60-year olds. Another type of weighting is more comprehensive than this, and takes into account the likelihood that a person will enjoy, over the course of their entire life, a "fair share" of good health. This type of weighting is based on egalitarian ageism.

Imagine two patients are in competition for a life saving treatment that only one can receive. If the only difference between them is their age – in particular, if both will gain the same number of QALYs - it would seem unfair to save the life of the older patient. For example, Lockwood comments: "To treat the older person, letting the younger person die, would ... be inherently inequitable in terms of years of life lived: the younger person would get no more years than the relatively few he has already had, whereas the older person, who has already had more than the younger person, will get several more years" (Lockwood 1988, p. 50). Kappel and Sandøe make the same point, commenting on the example of a 20-year old and a 60-year old who will both gain 10 years from a liver transplant: "If we give the liver to the older person, he will get another ten years on top of those 60 that he has already got. And he will end up with 70 life years. Whereas if we give the liver to the young person he will only end up with a total of thirty life years. To give the liver to the older person is like giving money to the rich rather than to the poor" (Kappel and Sandøe 1992, pp. 313-314).

Note that conventional, QALY-based CEA attaches no significance to age per se. The fact that giving priority to one group of patients rather than another will secure more life years for those who have already lived longest, while (in effect) depriving of further life years another group who have lived a shorter time, is not a relevant consideration. Appealing solely to the cost-per-QALY prioritising rule, the lower the cost per QALY the higher a priority a programme or service should be, regardless of the ages of those likely to benefit.

One way of overcoming this apparent shortcoming of CEA is suggested by Williams. He recommends surveying the general public to find out how many life years, adjusted for quality of life, constitutes a "fair" lifetime allocation (within a given society), and then applying weights to life years depending on a person's prospects of achieving this "fair innings". The life years of those with a smaller Quality-Adjusted Life Expectancy (QALE) at birth will be given added weight, while

those with a greater QALE at birth will receive a lower weight.[8] Williams gives an example of how this might work in practice (Williams 1997, p. 121). In the UK, the survival rate for males in the professional and managerial classes is 72.5 years, whereas for semi-skilled and unskilled manual workers it is 67.7 years. This is a difference of about 5 years. When morbidity is taken into account (semi-skilled and unskilled manual workers also experience a poorer HRQoL) the difference between the classes in terms of QALE at birth is about 9 QALYs in favour of the professional and managerial classes. Williams notes that this discrepancy could be lessened by applying weights to life years depending on the social class of the potential recipient. This would go some way towards equalising the prospects of a long life of good quality between the two classes, at the expense of reducing the total number of QALYs shared between the two groups.

Table 44. Life Extending Interventions (for Other People)

Age of recipient				Equivalent numbers of people helped	
Group A	Group B	A		B (median)	IQR (B)
20 year old	10 year old	10	=	9.5	9.0-10.0
60 year old	20 year old	10	=	4.0	2.0-5.5
80 year old	20 year old	10	=	1.0	0.02-2.4

Source: (Nord, Street et al. 1996)

Table 45. Health Improving Interventions (for Other People)

Age of recipient				Equivalent numbers of people helped	
Group A	Group B	A		B (median)	IQR (B)
20 year old	10 year old	10	=	9.0	4.5-9.5
60 year old	20 year old	10	=	4.0	1.0-6.0
80 year old	20 year old	10	=	1.0	0.02-2.5

Source: (Nord, Street et al. 1996)

In the previously mentioned study, in which they sought to quantify the preference for duration of effect, Nord, Street and colleagues also sought to measure egalitarian ageism among the Australian population (Nord, Street et al. 1996). Subjects were asked to choose between two equally costly projects that would enable patients to continue living in normal health for 10 years. Project A would allow the treatment of more patients, but project B would benefit younger patients. Subjects were asked to adopt the role of health administrators – that is, to assume they would not personally be affected by the decision made. They were then asked how many patients treated under project B they would consider equivalent to 10 patients being treated under project A. They found, for example, that saving the lives of 10 twenty-year olds was considered equivalent to saving the lives of 9.5 ten-year olds, and that saving the lives of 10 eighty-year olds was considered equivalent to saving the life of 1 twenty-year old (see Table 44). This obviously conflicts with the cost-per-QALY approach, since all patients, young and old, were stipulated to derive the same benefit – that is, 10 years of good health. In general, "[t]he preference for the young is more pronounced the greater the disparity in the two age groups" (Nord, Street et al.

---

8. As Tsuchiya, Dolan et al. note, egalitarian ageism or "fair innings ageism" may take a more or a less extreme form: "In its extreme form, FIA [fair innings ageism] will set some discontinuous threshold or cut-off corresponding to the fair innings, beyond which people will be denied health care" (Tsuchiya, Dolan et al. 2003, p. 688). By contrast, the form of ageism Williams defends is less extreme: "In its less extreme form, the fair innings represents the point at which people will be given a relative weight of one in the priority-setting calculus. If an individual has poorer (better) lifetime health prospects than the fair innings, they will be given relative weights larger (smaller) than one" (Tsuchiya, Dolan et al. 2003, p. 688).

1996, p. 106). Similar results were obtained for health-improving rather than life-saving treatments (see Table 45).

As Tsuchiya points out, it is possible that subjects' choices in this study partly reflected productivity ageism, since, although it was made clear that the young and old would derive the same benefit in terms of QALYs, there is no indication that subjects were told to consider the young and the old equivalent in terms of their social contribution. "Egalitarianism may appear to be supported, not only for equity concerns, but also for efficiency concerns" (Tsuchiya 1999, p. 274). Furthermore, because the young and the old would derive the same benefit (10 years of good health), it is not possible to draw any conclusions about an equity-efficiency trade-off – it is not possible to conclude that subjects would continue to favour the young if they would benefit less in health terms. Using Tsuchiya's terminology, the study only supports *weak* egalitarian ageism – "a supplement to conventional QALY maximisation that comes into effect when the latter is indifferent between two parties and one of them is younger than the other" - not *strong* egalitarian ageism – which "will continue to favour the younger patient, as long as the expected age of death of the younger patient remains less than the present age of the older patient" (Tsuchiya 1999, p. 274).

In a Canadian study, Choudhry, Slaughter and colleagues asked senior health officials to choose between two health programmes, A and B, that would benefit patients of different ages (Choudhry, Slaughter et al. 1997). The programmes were described as similar in terms of the benefits patients would receive (expressed in terms of life expectancy) and the harms they would be exposed to (expressed in terms of side-effects). As Table 46 indicates, a majority of respondents - 56.3 per cent - were either unable to decide between a programme that would

Table 46. Decision-makers' Responses to Programs where Beneficiaries Differ in Average Ages – All Respondents

	Participants (%,*n*)			Comments[a]
	Favoring Program A	Favoring Program B	Unable to decide/ no preference	
All respondents	42.5 (34)	1.3 (1)	56.3 (45)	Program A: 30-year-old patients
Respondents able to decide	97.1 (34)	2.9 (1)	N/A	Program B: 50-year-old patients
All respondents	57.5 (46)	8.8 (7)	33.8 (27)	Program A: 5-year-old patients
Respondents able to decide	76.8 (46)	13.2 (7)	N/A	Program B: 65-year-old patients

Source: (Choudhry, Slaughter et al. 1997)

a   The programs were described as equal in all respects except those outlined in the table; increases in life-expectancy and ages of patients affected by a given disease are expressed as averages.

benefit 30-year olds (Programme A) and a programme that would benefit 50-year olds (Programme B), or had no preference. Of those who were able to decide, 97.1 per cent chose Programme A. When the choice was between a programme that would benefit 5-year olds (Programme A) and a programme that would benefit 65-year olds (Programme B) fewer subjects - 33.8 per cent - could not decide. However, a clear majority of those who were able to decide - 76.8 per cent - still favoured Programme A. Because the benefits and harms associated with these programmes were described as similar (e.g. a 5-year increase in life expectancy), it is difficult to ascribe the results to "utilitarian ageism", unless subjects were importing into their deliberations unintended assumptions. The results probably reflect a combination of "egalitarian ageism" and "productivity ageism".

In a British study, Ratclife used "social" conjoint analysis (CA) to test public preferences for the allocation of donor liver grafts for transplantation (Ratcliffe 2000). In the first section of the mail questionnaire, individuals were asked (inter alia) to indicate their degree of agreement with the statement "preference should be given to younger rather than older people" in the context of scarce organ allocation. Of respondents, 14.8 per cent strongly agreed with the statement, 51.5 per cent agreed with the statement, 18.9 per cent disagreed with the statement, and 5.1 per cent strongly disagreed with the statement (9.8 per cent had no opinion). In the second section of the questionnaire respondents were asked to allocate 100 donor livers between two groups of 100 patients awaiting a life-saving transplant. The two groups differed in the following five attributes: expected length of survival, whether the patient's liver disease was naturally occurring or alcohol induced, the time spent on the waiting list for a transplant, the age of the patient, and whether the patient was a primary or re-transplant candidate. The age of the patients ranged over three levels: 40 years, 50 years, and 60 years. Likewise, the expected length of survival ranged over 5 years, 10 years, and 15 years. In allocating the livers between the two groups, respondents "generally exhibited a preference for younger people over older people" (Ratcliffe 2000, p. 144). By including length of survival as an additional attribute, Ratcliffe sought to derive the extent of egalitarian ageism: "One of the aims of this study was to attempt to isolate equity reasons for ageism by including the expected length of survival as an additional attribute, and thereby reducing the potential for efficiency reasons for ageism to enter into individual's preferences" (Ratcliffe 2000, p. 140). However, this neglects productivity ageism, which might also have influenced subjects' choices.

Tsuchiya, Dolan et al. conducted a study in the UK aimed at uncovering why people prefer some age groups over others (Tsuchiya, Dolan et al. 2003). They asked subjects to rank five different age groups: 5-year olds, 20-year olds, 35-year olds, 55-year olds and 70-years olds. Subjects were told that people in each of these groups would die in a few days if not treated, and would gain 5 years in full health if treated. A fixed duration of 5 years was chosen because this would be useful for generating age weights for use in cost-per-QALY studies. For this purpose, weights are needed that reflect the relative values of a single year of life at different ages. By contrast, a remaining-lifetime benefit would produce the relative values of various durations starting at different ages. Also, using a remaining-lifetime benefit allows subjects to import into their deliberations confounding assumptions about normal life expectancy at different ages. Years of life in full health, rather than health improvement, was chosen as the unit of benefit for similar reasons, for subjects may think that a given health problem represents a larger loss to a younger person than to an older one (contrary to the assumption of "distributive neutrality" in the QALY approach), thus again confounding the findings. The results of the ranking exercise are given in Table 47. As Tsuchiya, Dolan et al. observe: "The most common response was to rank the ages in ascending order, from 5-year olds first to 70-year olds last - 45% of respondents had this particular ranking and the Borda scores decrease monotonically from age 5 to age 70" (Tsuchiya, Dolan et al. 2003, p. 692).[9] Subjects were also asked to give reasons for their choices, which were coded by the interviewer into one of the following six categories:

   (1) Lived less life/had a 'fair innings'
   (2) Benefit to society
   (3) Family and/or other responsibilities
   (4) Five years may be more valuable to some people than others
   (5) Capacity to benefit

---

9. "Borda scores are calculated by treating the ordinal information as if they were expressed on an interval scale. For each respondent, the first through fifth ranked ages are given scores of 5 through 1. These scores are then added up across the respondents for each age, so that the age with the largest total score is ranked first and the age with the lowest score is ranked last" (Tsuchiya, Dolan et al. 2003, p. 691).

(6) Other (any other response that the interviewer recorded verbatim)

The results are given in Table 48. A clear majority of respondents (76 per cent) cited the "fair innings" argument in support of their chosen ranking.

Table 47. *Ranking Results in Round One – 5-Year Benefit Question*

Age	Rank 1 (%)	Rank 2 (%)	Rank 3 (%)	Rank 4 (%)	Rank 5 (%)	Borda score
5	**53**	10	23	10	3	4.0
20	20	**60**	10	8	2	3.9
35	22	20	**53**	5	0	3.6
55	2	10	10	**75**	3	2.3
70	3	0	3	2	**92**	1.3

Source: (Tsuchiya, Dolan et al. 2003)

Modal rank is in bold.

Table 48. *Reasons for Rankings (Numbers are Percentages of Respondents)*

Reason	Round 1: 5-year benefit
1. Have lived less life (FIA)	76
2. Greater benefit to society (PA)	8
3. Family responsibilities, etc. (PA)	22
4. Five years is more valuable to some (PA)	0
5. Capacity to benefit (HMA)	0
6. Other	3

Source: (Tsuchiya, Dolan et al. 2003)

Reason (1) is consistent with fair innings ageism.
Reasons (2)-(4) are consistent with productivity ageism.
Reason (5) is consistent with health maximisation ageism.

Although a fixed benefit of 5 years in full health was stipulated in this study, it is possible that subjects imported into their deliberations assumptions about normal life expectancy at different ages. To test this, Tsuchiya, Dolan et al. conducted a second round of interviews. This time they first asked subjects to rank the five age groups on the basis that they would live a normal life expectancy if treated. Subjects were then asked to rank them on the basis that they would gain 5 years in full health. "The juxtaposition of the 5-year benefit question with a 'full-life' benefit question should highlight the fixed and limited nature of the benefit in the main ranking question" (Tsuchiya, Dolan et al. 2003, p. 693). Table 49 shows how respondents ranked the different age groups when the benefit was remaining lifetime in full health. An even greater majority of respondents ranked the ages in ascending order. Table 50 shows how respondents ranked the different age groups when the benefit was five years in full health. A comparison of Table 49 and Table 50 shows that the aggregated ranking of the age groups in the 5-year benefit question in round one is significantly different from the ranking of the age groups in the 5-year benefit question in round two. The number of subjects ranking 5-year olds first dropped from 53 per cent to 39 per cent, and the number of respondents ranking 35-year olds first increased from 22 per cent to 43 per cent. "The Borda scores imply a hump-shaped profile with age 35 as the peak age" (Tsuchiya, Dolan et al. 2003, p. 694).

The reasons for the rankings in round two are shown in Table 51. When the benefit was remaining lifetime in full health, one-third of respondents cited the "fair innings" argument as the

main reason for their ranking (33 per cent), one-third cited reasons consistent with productivity ageism (33 per cent), and one-third cited reasons consistent with health maximisation ageism (33 per cent). Most significantly, the reasons differed between round one and round two when the benefit was 5 years in full health. Whereas 76 per cent of respondents - a clear majority -

Table 49. Ranking Results in Round Two – Full-life Benefit Question

Age	Rank 1 (%)	Rank 2 (%)	Rank 3 (%)	Rank 4 (%)	Rank 5 (%)	Borda score
5	**76**	13	10	1	0	4.6
20	13	**76**	10	1	0	4.0
35	10	7	**80**	1	1	3.2
55	0	3	0	**96**	1	2.0
70	1	1	0	0	**97**	1.1

Source: (Tsuchiya, Dolan et al. 2003)

Modal rank is in bold.

Table 50. Ranking Results in Round Two – 5-Year Benefit Question

Age	Rank 1 (%)	Rank 2 (%)	Rank 3 (%)	Rank 4 (%)	Rank 5 (%)	Borda score
5	**39**	3	17	24	17	3.2
20	14	**56**	17	13	0	3.7
35	**43**	17	39	0	1	4.0
55	3	24	23	**50**	0	2.8
70	1	0	4	13	**81**	1.3

Source: (Tsuchiya, Dolan et al. 2003)

Modal rank is in bold.

Table 51. Reasons for Rankings (Numbers are Percentages of Respondents)

Reason	Round 2: Full-life benefit	Round 2: 5-year benefit
1. Have lived less life (FIA)	33	37
2. Greater benefit to society (PA)	23	17
3. Family responsibilities, etc. (PA)	10	23
4. Five years is more valuable to some (PA)	0	21
5. Capacity to benefit (HMA)	33	0
6. Other	1	2

Source: (Tsuchiya, Dolan et al. 2003)

Reason (1) is consistent with fair innings ageism.
Reasons (2)-(4) are consistent with productivity ageism.
Reason (5) is consistent with health maximisation ageism.

cited the "fair innings" argument in support of their chosen ranking in round one, only 37 per cent cited this reason in round two. And whereas productivity ageism was cited by only 30 per cent of respondents in round one, this increased to 61 per cent in round two. In general, the reasons for the rankings in round two are more evenly spread across the categories. The different results for the 5-year benefit question in rounds one and two "provide quite strong support for the idea that a proportion of respondents in round one failed to appreciate – or edited out – the fixed nature of the benefit" (Tsuchiya, Dolan et al. 2003, p. 695).

It is clear from the preceding that there are different ways in which age might be deemed relevant in prioritising between different patient groups. It is equally clear that some studies of public preferences fail to distinguish carefully between these different forms of ageism. As a result, it is sometimes difficult to ascertain just what is being measured or detected. More studies are needed that quantify, specifically, social preferences for utilitarian ageism, productivity ageism, and egalitarian ageism.

Table 52. Studies of Age-related Social Preferences

Study	Country	Number of Subjects	Type of Subjects	Elicitation Method	Type of Ageism
(Williams 1988)	UK	377	GenPop[c]	DCM[d]	Prod[h]
(Charny, Lewis et al. 1989)	Wales	719	GenPop	DCM	?
(Busschbach, Hessing et al. 1993)	Netherlands	65	Students/Elderly	STTO[e]	Prod
(World Health Organization. 1993)	USA	?	Public Health Experts	MDM[f]	Prod
(Cropper, Aydede et al. 1994)	USA	3,000	GenPop	PTO[g]/DCM	Hybrid[i]
(Nord, Richardson et al. 1995a)	Australia	551	GenPop	DCM	?
(Bowling, Jacobson et al. 1993)	Britain	828	GenPop/Doctors	DCM	?
(Bowling 1996)	Britain	2005	GenPop	DCM	?
(Nord, Street et al. 1996) G1[a]	Australia	44	GenPop	PTO	Egal[j]
(Nord, Street et al. 1996) G2[b]	Australia	42	GenPop	PTO	Util
(Johannesson and Johansson 1997b)	Sweden	1,000	GenPop	PTO/DCM	Hybrid
(Choudhry, Slaughter et al. 1997)	Canada	80	MOH[l]	DCM	Egal
(Dolan, Cookson et al. 1999)	England	60	Patients	SGD[m]	Prod
(Rodríguez and Pinto 2000)	Spain	61	Students	DCM	?
(Ratcliffe 2000)	England	303	University Employees	SCA[n]	Egal
(Tsuchiya, Dolan et al. 2003)	England	140	GenPop	DCM	Util/Prod/Egal

a   G1 = Group 1
b   G2 = Group 2
c   GenPop = General Population
d   DCM = Discrete Choice Method
e   STTO = Social Time Trade-Off
f   MDM = Modified Delphi Method
g   PTO = Person Trade-Off
h   Prod = Productivity Ageism
I   Hybrid = Hybrid Utilitarian Ageism
j   Egal = Egalitarian Ageism
k   Util = Utilitarian Ageism
l   MOH = Senior Officials in the Ministry of Health of Ontario
m   SGD = Small Group Discussion
n   SCA = Social Conjoint Analysis

Considering the empirical findings on age-related preferences, Tsuchiya draws four main conclusions (Tsuchiya 1999, p. 271):

1. there is no support for a horizontal age weight profile
2. there is broad agreement that the profile declines after middle age
3. there is some disagreement on whether middle age or childhood is the peak and
4. the general pattern is independent of respondent age.

To these it seems safe to add a fifth:

5. there is broad agreement that a newborn infant attracts less weight than a young child or young adult when life-saving treatment is at issue.

CONCLUSION

It is becoming more and more evident that economic evaluation studies that neglect the distributional effects of alternative programmes, services and policies may not reflect community values. In Part 1 we have looked at several issues where the evidence regarding community expectations is becoming clear, and is potentially at odds with economic orthodoxy, which implies ranking programmes according to their cost-effectiveness ratio. *Neglected Equity Issues in Cost Effectiveness Analysis – Part 2* looks at some further equity issues that present a challenge to the traditional economic approach, but where the evidence concerning community preferences is at present only suggestive.

REFERENCES

Abellan-Perpiñán, J.-M. and J.-L. P. Prades (1999). "Health state after treatment: A reason for discrimination?" *Health Economics* **8**(8): 701-707.

Abelson, J., J. Lomas, J. Eyles, S. Birch and G. Veenstra (1995). "Does the Community Want Devolved Authority? Results of Deliberative Polling in Ontario." *Canadian Medical Association Journal* **153**(4): 403-412.

Andersson, F. and C. H. Lyttkens (1999). "Preferences for equity in health behind a veil of ignorance." *Health Economics* **8**(5): 369-378.

Bowling, A. (1996). "Health care rationing: The public's debate." *British Medical Journal* **312**(7032): 670-674.

Bowling, A., B. Jacobson and L. Southgate (1993). "Explorations in Consultation of the Public and Health Professionals on Priority Setting in an Inner London Health District." *Social Science and Medicine* **37**(7): 851-857.

Boyd, N. F., H. J. Sutherland, Z. K. Heasman, D. L. Trichler and B. J. Cummings (1990). "Whose Utilities for Decision Analysis?" *Medical Decision Making* **10**(1): 58-67.

Busschbach, J. J. V., D. J. Hessing and F. T. d. Charro (1993). "The utility of health at different stages of life: A quantitative approach." *Social Science and Medicine* **37**(2): 153-158.

Callahan, D. (1994). "Setting Mental Health Priorities: Problems and Possibilities." *The Milbank Quarterly* **72**(3): 451-470.

Campbell, A. and G. Gillett (1993). Justice and the right to health care. *Ethical Issues in Defining Core Services*. Wellington, The National Advisory Committee on Core Health and Disability Support Services.

Charny, M. C., P. A. Lewis and S. C. Farrow (1989). "Choosing who shall not be treated in the NHS." *Social Science and Medicine* **28**(12): 1331-1338.

Choudhry, N., P. Slaughter, K. Sykora and C. D. Naylor (1997). "Distributional dilemmas in health policy: Large benefits for a few or small benefits for many?" *Journal of Health Services Research and Policy* **2**(4): 212-216.

Cohen, B. J. (1995). "Utility measurement and the allocation of health care resources." *Medical Decision Making* **15**: 287-288.

Cook, J., J. Richardson and A. Street (1994). "A cost-utility analysis of treatment options for gallstone disease: Methodological Issues and Results." *Health Economics* **3**(3): 157-168.

Cropper, M. L., S. K. Aydede and P. R. Portney (1994). "Preferences for Life Saving Programs: How the Public Discounts Time and Age." *Journal of Risk and Uncertainty* **8**: 243-265.

Daniels, N. (1993). "Rationing Fairly: Programmatic Considerations." *Bioethics* **7**(2/3): 224-233.

Dolan, P. and R. Cookson (2000). "A qualitative study of the extent to which health gain matters when choosing between groups of patients." *Health Policy* **51**(1): 19-30.

Dolan, P., R. Cookson and B. Ferguson (1999). "Effect of discussion and deliberation on the public's views of priority setting in health care: Focus group study." *British Medical Journal* **318**(7188): 916-919.

Dolan, P., C. Gudex, P. Kind and A. Williams (1996). "Valuing health states: A comparison of methods." *Journal of Health Economics* **15**(2): 209-231.

Dolan, P. and J. A. Olsen (2002). *Distributing Health Care: Economics and Ethical Issues*. Oxford, Oxford University Press.

Dutch Committee on Choices in Health Care (1992). *Choices in Health Care*. Rijswijk, Ministry of Welfare, Health and Cultural Affairs.

Dworkin, R. (1981). "What is Equality? Part 1: Equality of Welfare." *Philosophy and Public Affairs* **10**(3): 185-246.

Edgar, A. (1995). "Weighting Health States and Strong Evaluation." *Bioethics* **9**(3-4): 240-251.

Edgar, A., S. Salek, D. Shickle and D. Cohen (1998). *The Ethical QALY: Ethical Issues in Healthcare Resource Allocations*. Haslemere, UK, Euromed Communications.

Froberg, D. G. and R. L. Kane (1989). "Methodology for measuring health-state preferences." *Journal of Clinical Epidemiology* **42**(6): 345-354, 459-471, 585-592, 675-685.

Green, R. M. (1976). Health care and justice in contract theory perspective. *Ethics and Health Policy*. R. Branson. Cambridge, Massachusetts, Ballinger Publishing Co**:** 111-126.

Harris, J. (1987). "QALYfying the value of life." *Journal of Medical Ethics* **13**(3): 117-123.

Harris, J. (1988a). More and Better Justice. *Philosophy and Medical Welfare*. J. M. Bell and S. Mendus. Cambridge, Cambridge University Press**:** 75-96.

Hurst, N. P., P. Jobanputra, M. Hunter, M. Lambert, A. Lochhead and H. Brown (1994). "Validity of Euroqol - a Generic Health Status Instrument - in Patients with Rheumatoid Arthritis." *British Journal of Rheumatology* **33**(7): 655-662.

Johannesson, M. and U.-G. Gerdtham (1996). "A Note on the Estimation of the Equity-Efficiency Trade-Off for Qalys." *Journal of Health Economics* **15**(3): 359-368.

Johannesson, M. and P.-O. Johansson (1997b). "Is the Valuation of a Qaly Gained Independent of Age? Some Empirical Evidence." *Journal of Health Economics* **16**(5): 589-599.

Kahneman, D., J. L. Knetsch and R. H. Thaler (1991). "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias." *Journal of Economic Perspectives* **5**(5): 193-206.

Kahneman, D., P. Slovic and A. Tversky, Eds. (1982). *Judgement Under Uncertainty: Heuristics and Biases*. Cambridge, Cambridge University Press.

Kahneman, D. and C. Varey (1991). Notes on the psychology of utility. *Interpersonal Comparisons of Well-Being*. J. E. Roemer. Cambridge, Cambridge University Press**:** 127-161.

Kappel, K. and P. Sandøe (1992). "QALYs, age and fairness." *Bioethics* **6**(4): 297-316.

Kymlicka, W. (1990). *Contemporary Moral Philosophy*. Oxford, Oxford University Press.

Lockwood, M. (1988). Quality of Life and Resource Allocation. *Philosophy and Medical Welfare*. S. Mendus. Cambridge, Cambridge University Press**:** 33-55.

Mathers, C. D., E. T. Vos and C. E. Stevenson (1999). The burden of disease and injury in Australia. Canberra, Australian Institute of Health and Welfare.

Menzel, P., P. Dolan, J. Richardson and J. A. Olsen (2002). "The role of adaptation to disability and disease in health state valuation: A preliminary normative analysis." *Social Science and Medicine* **55**(12): 2149-2158.

Murray, C. J. L. (1994). "Quantifying the burden of disease: The technical basis for disability-adjusted life years." *Bulletin of the World Health Organization* **72**(3): 429-445.

Murray, C. J. L. and A. Acharya (2002). Age Weights and Discounting in Health Gaps Reconsidered. *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications*. C. J. L. Murray, J. A. Salomon, C. Mathers and A. Lopez. Geneva, World Health Organization**:** 677-684.

Murray, C. J. L. and A. D. Lopez, Eds. (1996). *The global burden of disease: A comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020.* Cambridge, Massachusetts, Harvard University Press.

Murray, C. J. L., J. A. Salomon and C. Mathers (2000). "A critical examination of summary measures of population health." *Bulletin of the World Health Organization* **78**(8): 981-994.

Myllykangas, M., O.-P. Ryynänen, J. Kinnunen and J. Takala (1996). "Comparison of doctors', nurses', politicians' and public attitudes to health care priorities." *Journal of Health Services Research and Policy* **1**(4): 212-216.

Nord, E. (1991). "The validity of a visual analogue scale in determining social utility weights for health states." *International Journal of Health Planning and Management* **6**: 234-242.

Nord, E. (1992). "Methods for quality adjustment of life years." *Social Science and Medicine* **34**(5): 559-569.

Nord, E. (1993a). "The trade-off between severity of illness and treatment effect in cost-value analysis of health care." *Health Policy* **24**: 227-238.

Nord, E. (1993b). "Helsepolitikere ønsker ikke mest mulig helse per krone (Health politicians do not wish to maximize health benefits)." *Journal of the Norwegian Medical Association* **113**: 1171-1173.

Nord, E. (1993d). "The relevance of health state after treatment in prioritising between different patients." *Journal of Medical Ethics* **19**(1): 37-42.

Nord, E. (1999). *Cost-Value Analysis in Health Care.* Cambridge, Cambridge University Press.

Nord, E., J. Richardson and K. Macarounas-Kirchmann (1993). "Social evaluation of health care versus personal evaluation of health states: Evidence on the validity of four health state scaling instruments using Norwegian and Australian Surveys." *International Journal of Technology Assessment in Health Care* **9**(4): 463-478.

Nord, E., J. Richardson, A. Street, H. Kuhse and P. Singer (1995a). "Maximizing Health Benefits Vs Egalitarianism: An Australian Survey of Health Issues." *Social Science and Medicine* **41**(10): 1429-1437.

Nord, E., J. Richardson, A. Street, H. Kuhse. and P. Singer (1995b). "Who cares about cost? Does economic analysis impose or reflect social values?" *Health Policy* **34**(2): 79-94.

Nord, E., A. Street, J. Richardson, H. Kuhse and P. Singer (1996). "The Significance of Age and Duration of Effect in Social Evaluation of Health Care." *Health Care Analysis* **4**(2): 103-111.

Olsen, J. A. (1994). "Persons vs years: Two ways of eliciting implicit weights." *Health Economics* **3**: 39-46.

Olsen, J. A. (2000). "A note on eliciting distributive preferences for health." *Journal of Health Economics* **19**(4): 541-550.

Olsen, J. A. and J. Richardson (1999). "Production gains from health care: What should be included in cost-effectiveness analyses?" *Social Science and Medicine* **49**(1): 17-26.

Patrick, D. L., J. W. Bush and M. M. Chen (1973). "Methods for Measuring Levels of Well-being for a Health Status Index." *Health Services Research* **8**: 228-245.

Polsky, D., R. J. Willke, K. Scott, K. A. Schulman and H. A. Glick (2001). "A comparison of scoring weights for the EuroQol[(c)] derived from patients and the general public." *Health Economics* **10**(1): 27-37.

Prades, J.-L. P. (1997). "Is the person trade-off a valid method for allocating health care resources?" *Health Economics* **6**(1): 71-81.

Prades, J.-L. P. and A. Lopez-Nicolás (1998). "More Evidence of the Plateau Effect: A Social Perspective." *Medical Decision Making* **18**(3): 287-294.

Rabin, M. (1998). "Psychology and Economics." *Journal of Economic Literature* **36**(1): 11-46.

Ratcliffe, J. (2000). "Public preferences for the allocation of donor liver grafts for transplantation." *Health Economics* **9**(2): 137-148.

Rawls, J. (1971). *A Theory of Justice*. Cambridge, Harvard University Press.

Richardson, J. (1991). "Economic assessment of health care: Theory and Practice." *Australian Economic Review*(First Quarter): 2-21.

Richardson, J. (1994). "Cost Utility Analysis: What Should Be Measured?" *Social Science and Medicine* **39**(1): 7-21.

Richardson, J. (2002c). Age Weighting and Time Discounting: Technical Imperatives Versus Social Choice. *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications*. C. J. L. Murray, J. A. Salomon, C. D. Mathers and A. D. Lopez. Geneva, World Health Organization**:** 663-676.

Richardson, J., J. Wildman and I. K. Robertson (2003). "A critique of the World Health Organisation's evaluation of health system performance." *Health Economics* **12**(5): 355-366.

Rodríguez, E. and J. L. Pinto (2000). "The social value of health programmes: Is age a relevant factor?" *Health Economics* **9**(7): 611-621.

Rodríguez-Míguez, E. and J.-L. P. Prades (2002). "Measuring the social importance of concentration or dispersion of individual health benefits." *Health Economics* **11**(1): 43-53.

Rothman, M. L. (1991). "The validity of proxy-generated scores as measures of patient health status." *Medical Care* **29**(2): 115-124.

Savulescu, J. (1998). "Consequentialism, Reasons, Value and Justice." *Bioethics* **12**(3): 212-235.

Scanlon, T. M. (1975). "Preference and Urgency." *The Journal of Philosophy* **LXXII**(19): 655-669.

Smith, R. D. (1996). "Is Regret Theory an alternative basis for estimating the value of healthcare interventions?" *Health Policy* **37**(2-3): 105-115.

Stein, M. S. (2002). "Utilitarianism and the disabled: Distribution of resources." *Bioethics* **16**(1): 1-19.

Stokey, E. and R. Zeckhauser (1978). *A Primer for Policy Analysis*. New York, W. W. Norton and Co.

Swedish Health Care and Medical Priorities Commission (1993). *No Easy Choices: The Difficulties of Health Care. Sveriges offentlige utredninger, 1993*. Stockholm, The Ministry of Health and Social Affairs.

Tsuchiya, A. (1999). "Age-related preferences and age weighting health benefits." *Social Science and Medicine* **48**(2): 267-276.

Tsuchiya, A., P. Dolan and R. Shaw (2003). "Measuring people's preferences regarding ageism in health: Some methodological issues and some fresh evidence." *Social Science and Medicine* **57**(4): 687-696.

Tversky, A. and D. Kahneman (1991). "Loss Aversion in Riskless Choice: A Reference-Dependent Model." *Quarterly Journal of Economics* **106**(4): 1039-1061.

Ubel, P. A. (1999b). "How stable are people's preferences for giving priority to severely ill patients?" *Social Science and Medicine* **49**(7): 895-903.

Ubel, P. A. (2000). *Pricing Life: Why It's Time for Health Care Rationing.* A. Caplan. Cambridge, Massachusetts, Massachusetts Institute of Technology Press.

Ubel, P. A., R. M. Arnold and A. C. Caplan (1993). "Rationing Failure: The Ethical Lessons of the Retransplantation of Scarce Vital Organs." *Journal of the American Medical Association* **270**(20): 2469-2474.

Ubel, P. A. and G. Loewenstein (1996a). "Distributing scarce livers: The moral reasoning of the general public." *Social Science and Medicine* **42**(7): 1049-1055.

Ubel, P. A., G. Loewenstein, D. Scanlon and M. Kamlet (1996). "Individual Utilities Are Inconsistent with Rationing Choices: A Partial Explanation of Why Oregons Cost-Effectiveness List Failed." *Medical Decision Making* **16**(2): 108-116.

Ubel, P. A., J. Richardson and J. Baron (2002). "Exploring the role of order effects in person trade-off elicitations." *Health Policy* **61**(2): 189-199.

Ubel, P. A., M. D. Spranca, M. L. DeKay, J. C. Hershey and D. A. Asch (1998). "Public preferences for prevention versus cure: What if an ounce of prevention is worth only an ounce of cure?" *Medical Decision Making* **18**(2): 141-148.

Victorian Government Department of Human Services (1999). Victorian Burden of Disease Study: Morbidity. Melbourne, Public Health and Development Division.

Wikler, D. (1989). "Equity, efficiency, and the point system for transplant recipient selection." *Transplantation Proceedings* **21**(3): 3437-3439.

Williams, A. (1988). Ethics and efficiency in the provision of health care. *Philosophy and Medical Welfare.* S. Mendus. Cambridge, Cambridge University Press**:** 111-126.

Williams, A. (1997). "Intergenerational Equity: An Exploration of the Fair Innings Argument." *Health Economics* **6**(2): 117-132.

Wolf, C. (1997). Person-Affecting Utilitarianism and Population Policy; or, Sissy Jupe's Theory of Social Choice. *Contingent Future Persons.* N. Fotion and J. C. Heller. Dordrecht, Kluwer Academic Publishers**:** 99-122.

World Health Organization (1993). *World Development Report 1993: Investing in Health.* New York, Oxford University Press.

World Health Organization (2000). *The World Health Report 2000: Health Systems: Improving Performance.* Geneva, World Health Organization.

World Health Organization. (1993). *World Development Report 1993: Investing in Health.* New York, Oxford University Press.

Yaari, M. E. and M. Bar-Hillel (1984). "On Dividing Justly." *Social Choice and Welfare* **1**(1): 1-24.