THE UNIVERSITY *of York*

CHE
Centre For Health Economics

Does cost-effectiveness analysis discriminate against patients with short life expectancy? Matters of logic and matters of context

CHE Research Paper 55

Does cost-effectiveness analysis discriminate against patients with short life expectancy? Matters of logic and matters of context

Mike Paulden[1]
Anthony J Culyer[2,3,4]

[1]Toronto Health Economics and Technology Assessment (THETA) Collaborative, University of Toronto, Canada
[2]Department of Health Policy, Management and Evaluation, University of Toronto, Canada
[3]Department of Economics and Related Studies, University of York, UK
[4]Centre for Health Economics, University of York, UK

February 2010

Background

CHE Discussion Papers (DPs) began publication in 1983 as a means of making current research material more widely available to health economists and other potential users. So as to speed up the dissemination process, papers were originally published by CHE and distributed by post to a worldwide readership.

The CHE Research Paper series takes over that function and provides access to current research output via web-based publication, although hard copy will continue to be available (but subject to charge).

Disclaimer

Papers published in the CHE Research Paper (RP) series are intended as a contribution to current research. Work and ideas reported in RPs may not always represent the final position and as such may sometimes need to be treated as work in progress. The material and views expressed in RPs are solely those of the authors and should not be interpreted as representing the collective views of CHE research staff or their research funders.

Further copies

Copies of this paper are freely available to download from the CHE website www.york.ac.uk/inst/che/pubs. Access to downloaded material is provided on the understanding that it is intended for personal use. Copies of downloaded papers may be distributed to third-parties subject to the proviso that the CHE publication source is properly acknowledged and that such distribution is not subject to any payment.

Printed copies are available on request at a charge of £5.00 per copy. Please contact the CHE Publications Office, email che-pub@york.ac.uk, telephone 01904 321458 for further details.

Centre for Health Economics
Alcuin College
University of York
York, UK
www.york.ac.uk/inst/che

Abstract

The aim of this paper is to explore the claim of ageism made against the National Institute for Health & Clinical Excellence and like organisations, and to identify circumstances under which ageist discrimination might arise. We adopt a broad definition of ageism as representing any discrimination against individuals or groups of individuals solely on the basis that they have shorter life expectancy than others. A simple model of NICE's decision making process is developed which demonstrates that NICE's recommendations do not inherently discriminate on the basis of life expectancy *per se* but that scope for discrimination may arise in the case of specific technologies having identifiable characteristics. Such discrimination may favour patients with either longer or shorter life expectancy. It is shown that NICE's policies, procedures and the context in which NICE makes its decisions not only reduce the scope for discriminatory recommendations but also – in the case of "end of life" treatments – increase the likelihood that NICE's recommendations favour those with shorter, rather than longer, life expectancy.

Introduction

In a recent series of articles and editorials, John Harris[1] has accused the most prominent public agency espousing the practice of cost-effectiveness analysis (CEA) in health care decision making, the UK's National Institute for Health and Clinical Excellence (NICE), of "age discrimination" for embracing the use of quality-adjusted life years (QALYs) (Harris, 2005a; 2005b)[2]. According to Harris, "in every case QALYs are indeed, and by definition, inherently ageist and also favour those with the greater life expectancy regardless of age" (Harris, 2005b). Harris's argument has superficial attraction. Since older patients tend to have a shorter life expectancy than younger patients – and often a persistently worse state of health – the QALYs associated with their remaining life expectancy are generally lower too. A life-saving procedure therefore tends to yield a larger QALY benefit when carried out on a younger person than on an older person. Since NICE takes QALYs into account when assessing whether or not to recommend technologies (NICE, 2008), it seems to follow that NICE is inherently ageist in its decision making.

This paper demonstrates that ageism in NICE's decision making is neither inherent nor inevitable. It will be shown that, for a charge of ageism against any of NICE's guidance to stick, a number of conditions must hold for the technology in question. NICE's policies, procedures and the context in which it makes its recommendations all reduce the likelihood of such conditions being satisfied. We further show that, under some conditions, NICE will tend to discriminate not against the old but against the young. Moreover, NICE's recent guidance on "end of life" treatments increases the possibility of this type of discrimination while reducing still further the scope for discrimination against the old.

How we define ageism

The term "ageism" was coined by R. N. Butler as "… a process of systematic stereotyping of and discrimination against people because they are old" (Butler 1975, p.12). In our context, ageism has had two related meanings that tend to result in discriminatory decisions about health care resource allocation, each of which is rooted in a stylized, but broadly true, factual account. The first is that 'old people' may indeed be stereotyped in many of the ways vividly described by Butler as "myths" (pp. 7-11)[3]; the second is that, being old, their expectation of remaining life with or without a particular technology is generally lower than that of a younger person. However, as Harris (2005a) accepts, there are also young people with short expectations of remaining life. We therefore adopt a broad definition of ageism as representing any discrimination against individuals or groups of individuals solely on the basis that they have shorter life expectancy than others (irrespective of their age).

Policy context

Under this definition of ageism, the charge against NICE is that it discriminates in its guidance against individuals with shorter life expectancy, whether old or otherwise. To investigate this claim requires a thorough consideration of exactly how and under what basis NICE makes its recommendations. These are made by one of four Appraisal Committees, whose members are drawn from the NHS, patient and carer organisations, academia and the pharmaceutical and medical devices industries.[4] NICE is charged by the Secretary of State with some specific duties, of which the first is "… to appraise the clinical benefits and the costs of such health care interventions… and to make recommendations" (NHS, 2005). NICE has interpreted "costs" as representing all costs associated with the intervention in question (whether direct or indirect) which are incurred by the NHS or Personal

[1] Harris has previously made a number of false statements about the use of QALYs in practice, which shall not detain us here. For example, the QALY is: [1] "supposedly a criterion of beneficial health care", [2] "a measure of efficiency in health care', and [3] "a method of settling priorities in health care" (Harris 1988, p. 77). Statement 1 is true so long as one remembers the singularity of the criterion – the QALY is but one of many possible indicators of beneficial care; statement 2 is false: more QALYs are an indicator neither of efficiency nor inefficiency (no one claims, of course, that they are an indicator of fairness); statement 3 is false: QALYs may be helpful in understanding in a quantitative way one element in a balancing of pros and cons when one is "settling" priorities.

[2] See also Quigley (2007) and the responses of Claxton and Culyer (2006, 2007, 2008).

[3] This is to be distinguished from the fact that old people often have multiple morbidities and a general health state that makes them less responsive to health care. Less ability to benefit from health care for such reasons may indeed be a ground for not prioritising someone but it is not inherently ageist.

[4] http://www.nice.org.uk/aboutnice/howwework/devnicetech/technologyappraisalcommittee/technology_appraisal_committee.jsp

Social Services (NICE, 2008). It has interpreted “clinical benefits” as the expected gain in health by the recipients of care and has recommended the EQ-5D version of the QALY as its preferred quantitative indicator of this gain (NICE, 2008). NICE has from its inception acknowledged that QALY estimates do not exist for many interventions and that, even when they do exist, they need to be interpreted with care – both regarding the well-known assumptions on which they are calculated and in the specific context of the patients and circumstances being considered at the time. QALY estimates are checked against patient experience through the deliberative processes of the relevant Appraisals Committee and, where it is deemed necessary, supplemented by other measures of health benefit.

NICE’s Appraisal Committee does not consider these QALY or cost estimates in isolation. Rather, it considers both the expected QALY benefit and the expected costs associated with the technology in question and any relevant comparator technologies and then calculates the expected ‘incremental’ QALY benefit and expected ‘incremental’ costs associated with each technology relative to the next less-effective or less-costly comparator technology (NICE, 2008). For example, if a particular technology provides each patient an expected benefit of 8 QALYs at an expected cost of £100,000, while the next less-effective comparator provides an expected benefit of 3 QALYs at an expected cost of £25,000, then the technology in question provides an expected incremental benefit of 5 QALYs at an expected incremental cost of £75,000. These expected incremental QALYs and costs are then rearranged into an ‘incremental cost-effectiveness ratio’ (ICER) by dividing the expected incremental costs by the expected incremental QALYs – in this case the ICER of the technology will be £75,000 / 5 QALYs = £15,000 per QALY. An important (but by no means only) part of NICE’s decision making process is to compare this ICER to the adopted ‘cost-effectiveness threshold’, which in NICE’s case is £20,000-30,000 per QALY – technologies with ICERs below the cost-effectiveness threshold are likely to be accepted, while technologies with ICERs above the cost-effectiveness threshold are likely to be rejected[5] (NICE, 2008). With an ICER of £15,000 per QALY, the example technology considered here is very likely to be accepted[6].

In practice, NICE will often consider subgroups of patients with different characteristics – for example patients of different age, sex, severity of disease, etc. If the QALYs and costs associated with the technology and its comparators differ between subgroups then the ICERs will tend to differ between subgroups too. NICE then considers each of these different ICERs, and in some cases the technology will be recommended for one subgroup of patients but not for another. In assessing whether NICE discriminates on the basis of life expectancy, a crucial question to ask is therefore: to what extent does life expectancy affect the ICER associated with any particular technology? More specifically, do differences in life expectancy between subgroups of patients result in sufficient differences between the ICERs for discrimination to result? The aim of this paper is to establish in a rigorous way the circumstances under which such discrimination may occur.

A formal model

The NICE decision setting can be modelled formally with some simple algebra. Suppose that NICE must recommend whether to accept a particular technology or its comparator and is provided with QALY and cost data for two different subgroups of patients: subgroup S and subgroup L. Following treatment with the technology, patients in subgroup S have a relatively short life expectancy of p years, while those in subgroup L have a longer life expectancy of q years, where $q > p$. The QALY benefits of treatment are denoted by Δh^S for subgroup S and Δh^L for subgroup L, while the costs of treatment are Δc^S for subgroup S and Δc^L for subgroup L. In each case Δ signals a difference of either outcome or cost between the technology and its comparator. For the moment assume no discounting. The ICERs for the technology are therefore $\Delta c^S/\Delta h^S$ for subgroup S and $\Delta c^L/\Delta h^L$ for subgroup L.

Suppose for now that NICE considers *only* the ICERs of the technology in question when deciding whether to recommend a technology for any particular subgroup of patients, and that it uses a common cost-effectiveness threshold across all subgroups. The cost-effectiveness threshold is denoted by λ. NICE will recommend the technology for those patients with longer life expectancy if $\Delta c^L/\Delta h^L < \lambda$ and will not recommend the treatment for patients with shorter life expectancy if

[5] For the sake of simplicity, it will be assumed throughout this paper that the estimated incremental costs and QALYs are both positive (resulting in an ICER that is also positive). Our conclusions are unaffected by these assumptions.
[6] Hereafter ‘QALYs’ and ‘costs’ refer to the estimated incremental QALY benefit and estimated incremental costs respectively.

$\Delta c^S/\Delta h^S > \lambda$. For NICE to discriminate in its guidance against patients with shorter life expectancy therefore requires an ICER for patients with longer life expectancy sufficiently lower than the ICER for patients with shorter life expectancy that the cost-effectiveness threshold lies between the two:

$$\Delta c^L/\Delta h^L < \lambda < \Delta c^S/\Delta h^S. \quad (1)$$

This is shown graphically in Figure 1. The slopes of the two rays from the origin, *Oa* and *Ob*, represent the ICERs for the two subgroups, while the slope of the dotted line represents λ. The region between the two ICERs will hereafter be referred to as the “region of differential cost-effectiveness”, or RDCE. Discrimination results only if λ lies within the RDCE, as drawn. Where λ lies above or below the RDCE then NICE will make the same recommendation for both subgroups, so no discrimination will occur.[7]

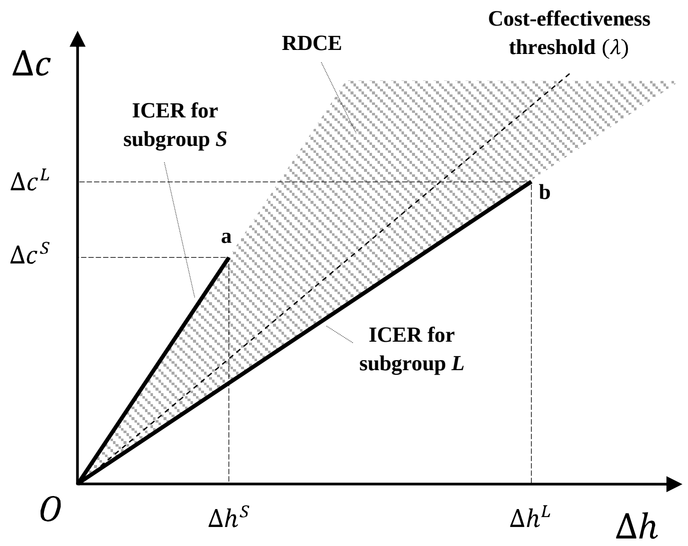

Figure 1

Unfortunately, the characteristics of technologies more or less likely to satisfy condition (1) are not immediately obvious. The model will now be used to derive a more intuitive condition for decisions to be discriminatory.

A more intuitive condition

Our first step is to extend the model to incorporate time. In general, the QALYs and costs of each technology will not be incurred immediately or even during the first year following the start of treatment but will instead be incurred over a much longer time horizon, in principle over a patient’s lifetime[8]. It is therefore usual practice to disaggregate the QALYs and costs for each subgroup by the year in which they will be incurred[9]. Denoting the current year as 1, the QALYs and costs for each subgroup may be disaggregated as follows:

$$\Delta h^S = \sum_{t=1}^{p} \Delta h_t^S = \Delta h_1^S + \Delta h_2^S + \cdots + \Delta h_p^S \quad (2)$$

$$\Delta c^S = \sum_{t=1}^{p} \Delta c_t^S = \Delta c_1^S + \Delta c_2^S + \cdots + \Delta c_p^S \quad (3)$$

$$\Delta h^L = \sum_{t=1}^{q} \Delta h_t^L = \Delta h_1^L + \Delta h_2^L + \cdots + \Delta h_p^L + \cdots + \Delta h_q^L \quad (4)$$

[7] The greater the difference between the ICERs (and hence the cost-effectiveness of the technology) for each subgroup, the larger the region of differential cost-effectiveness (RDCE). Where the technology is equally cost-effective in each subgroup, the RDCE is undefined.
[8] This is true even if the treatment itself is short: where the treatment is life-saving – or provides any lasting improvement in a patient’s health state, however small – QALY benefits may be realised long after the end of treatment.
[9] This is particularly useful for discounting, which is considered later.

$$\Delta c^L = \sum_{t=1}^{q} \Delta c_t^L = \Delta c_1^L + \Delta c_2^L + \cdots + \Delta c_p^L + \cdots + \Delta c_q^L, \quad (5)$$

where Δh_t^x and Δc_t^x represent the QALYs and costs of treatment for each subgroup, x, in each year, t. From (2-5), the ICERs for the treatment may be rewritten as:

$$\Delta c^S / \Delta h^S = \sum_{t=1}^{p} \Delta c_t^S / \sum_{t=1}^{p} \Delta h_t^S \quad (6)$$

$$\Delta c^L / \Delta h^L = \sum_{t=1}^{q} \Delta c_t^L / \sum_{t=1}^{q} \Delta h_t^L, \quad (7)$$

for subgroups S and L respectively. Condition (1) may then be rewritten as:

$$\sum_{t=1}^{q} \Delta c_t^L / \sum_{t=1}^{q} \Delta h_t^L < \lambda < \sum_{t=1}^{p} \Delta c_t^S / \sum_{t=1}^{p} \Delta h_t^S. \quad (8)$$

We now make an important assumption. In practice, subgroups of patients with different life expectancies tend to differ in other attributes as well – patients with shorter life expectancy tend to be older, are more likely to have co-morbidities and generally have worse overall health. Such patients may respond differently to treatment and may incur different costs, particularly if complications are more likely. This will inevitably lead to variations between subgroups in terms of the QALY benefits and costs of treatment in each year. However, in order to tease out whether NICE's decision making discriminates on the basis of life expectancy *per se* we shall assume that all other attributes are equal – that is, patients in subgroup S are identical to those in subgroup L in every way other than in life expectancy[10]. Where this is the case, the QALYs and costs of treatment will be identical for both subgroups until year p, when those patients in subgroup S die. The QALYs and costs will then generally differ between the subgroups from year $p+1$ until year q, when those patients in subgroup L die. After year q the QALYs and costs for both subgroups will be zero and need not be considered. Algebraically, $\Delta h_1^S = \Delta h_1^L$, $\Delta h_2^S = \Delta h_2^L$, … , $\Delta h_p^S = \Delta h_p^L$ and $\Delta c_1^S = \Delta c_1^L, \Delta c_2^S = \Delta c_2^L$, … , $\Delta c_p^S = \Delta c_p^L$. The ICER for subgroup S therefore represents a 'common ratio' of costs to QALYs for both subgroups up to year p:

$$\Delta c^S / \Delta h^S = \sum_{t=1}^{p} \Delta c_t^S / \sum_{t=1}^{p} \Delta h_t^S = \sum_{t=1}^{p} \Delta c_t^L / \sum_{t=1}^{p} \Delta h_t^L \;. \quad (9)$$

For ease of notation, we will refer to the ratio of costs to QALYs for subgroup L over years $p+1$ to q as the 'subsequent ratio', since these costs and QALYs are incurred in the years subsequent to the death of subgroup S. This subsequent ratio may be written algebraically as $\sum_{t=p+1}^{q} \Delta c_t^L / \sum_{t=p+1}^{q} \Delta h_t^L$. While the ICER for subgroup S is determined solely by the common ratio, the ICER for subgroup L is determined by both the common ratio and the subsequent ratio. If these ratios are equal then both subgroups have the same ICER[11]. Any difference between the ICERs is therefore driven by the (generally larger) difference between the ratios. Importantly for our purposes, the ICER for subgroup L is lower than the ICER for subgroup S only if the subsequent ratio is lower than the common ratio.

We can use this observation to derive a weaker but more intuitive condition from (8). A necessary (but not sufficient) condition for NICE to discriminate on the basis of life expectancy is that the ICER for subgroup L be lower than that for subgroup S:

$$\sum_{t=1}^{q} \Delta c_t^L / \sum_{t=1}^{q} \Delta h_t^L < \sum_{t=1}^{p} \Delta c_t^S / \sum_{t=1}^{p} \Delta h_t^S. \quad (10)$$

If both subgroups are identical, except with regard to life expectancy, this is equivalent to stating that a necessary (but not sufficient) condition for (8) to hold is that the subsequent ratio be lower than the common ratio:

$$\sum_{t=p+1}^{q} \Delta c_t^L / \sum_{t=p+1}^{q} \Delta h_t^L < \sum_{t=1}^{p} \Delta c_t^L / \sum_{t=1}^{p} \Delta h_t^L. \quad (11)$$

[10] We may weaken this assumption if the attributes in which the subgroups differ have no bearing on the QALYs and costs associated with treatment.
[11] This can be determined from Figure 2 and is also a standard result from mathematics – if a ratio is the same over years 1 to p as over years $p+1$ to q then it holds over years 1 to q.

This has a clear prescription: a necessary condition for NICE to be discriminatory in its guidance against patients with shorter life expectancy is that the ratio of costs to QALYs for the technology in question be lower over the long term (following the death of patients with shorter life expectancy) than over the short term. Whether this holds depends on the intertemporal distribution of QALYs and costs for the technology. It is more likely to hold for technologies with largely upfront costs and long term QALY benefits (such as some preventive programmes or surgery) and less likely to hold for technologies with continuous long term costs and constant or declining QALY benefits (such as long-term care for diabetes or rheumatoid arthritis). In the case of the former, the upfront costs increase the common ratio while having no effect on the subsequent ratio, while in the latter case the flat or declining QALY benefits and continuous long-term costs result in a common ratio equal to or lower than the subsequent ratio.

Critically, if condition (11) does not hold – and NICE considers only the ICERs when making its decisions – then there is no possibility of NICE discriminating against patients with shorter life expectancy: if it accepts the technology for those in subgroup L then it must accept the technology for those in subgroup S. Even if (11) holds, discrimination is not inevitable. To satisfy condition (8), the ICER for subgroup S must lie above λ, and, furthermore, the subsequent ratio must be sufficiently small to pull the ICER for subgroup L below λ.

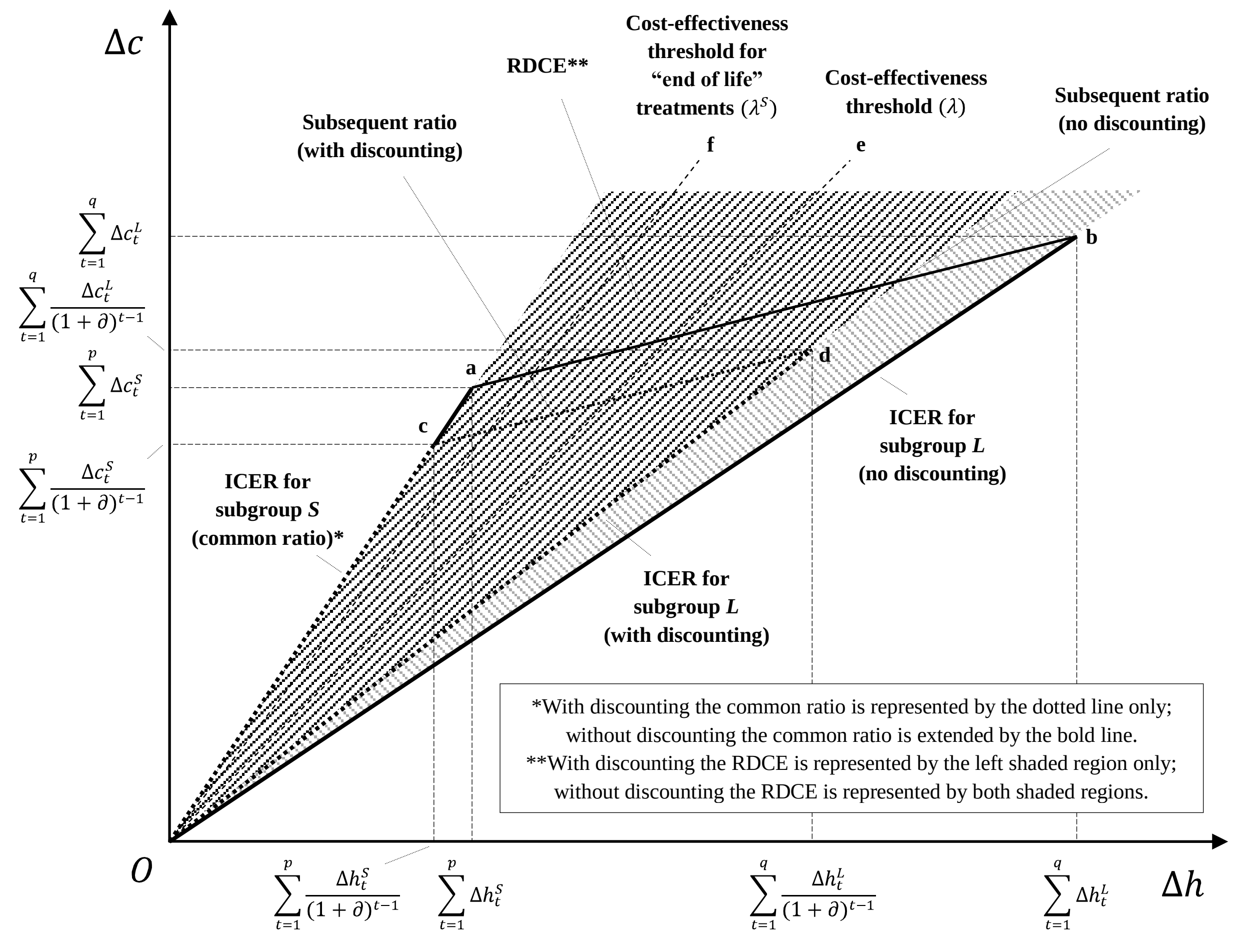

Figure 2

This is demonstrated in Figure 2. The slope of the ray *Oa* is the ICER for subgroup S (the common ratio), the slope of *Ob* is the ICER for subgroup L and the subsequent ratio is the slope of *cd*. As before, the region between the two ICERs represents the RDCE. The cost-effectiveness threshold, λ, is represented by the slope of *Oe*. Discrimination requires that *Oe* falls within the RDCE; for such discrimination to favour patients with shorter life expectancy requires that *Oa* lies above *Oe* and that *ab* be shallower than *Oa* to such an extent that *Ob* lies below *Oe*. If *ab* is not shallower than *Oa* then there is no scope for ageist discrimination. However, if *ab* is steeper than *Oa* then the RDCE will lie above the ICER for subgroup S; if, in addition, *Oe* lies within the RDCE then NICE will in fact discriminate against patients with *longer* life expectancy.

Further considerations

Discounting

In common with other public sector institutions (HM Treasury, 2003), NICE requires that estimates of future QALYs and costs be discounted - currently at a compound rate of 3.5% per annum - in order to account for time preference (NICE, 2008). So an estimated benefit of 10 QALYs five years from now would be discounted to a present value of 8.42 QALYs, while an estimated cost of £1m ten years from now would be discounted to a present value of £708,919. Algebraically, the discounted ICERs are now:

$$\Delta c^S/\Delta h^S = \sum_{t=1}^{p} \frac{\Delta c_t^S}{(1+\partial)^{t-1}} \Big/ \sum_{t=1}^{p} \frac{\Delta h_t^S}{(1+\partial)^{t-1}} \qquad (12)$$

$$\Delta c^L/\Delta h^L = \sum_{t=1}^{q} \frac{\Delta c_t^L}{(1+\partial)^{t-1}} \Big/ \sum_{t=1}^{q} \frac{\Delta h_t^L}{(1+\partial)^{t-1}} \qquad (13)$$

where ∂ denotes the discount rate. NICE's decisions therefore discriminate against those with shorter life expectancy where (from 8, 12 and 13):

$$\sum_{t=1}^{q} \frac{\Delta c_t^L}{(1+\partial)^{t-1}} \Big/ \sum_{t=1}^{q} \frac{\Delta h_t^L}{(1+\partial)^{t-1}} < \lambda < \sum_{t=1}^{p} \frac{\Delta c_t^S}{(1+\partial)^{t-1}} \Big/ \sum_{t=1}^{p} \frac{\Delta h_t^S}{(1+\partial)^{t-1}}. \qquad (14)$$

Since the discount rate is compounded over time, the long term costs and QALYs which comprise the subsequent ratio are discounted more heavily than those which comprise the common ratio. This has an ambiguous impact on the size of each ratio[12] but will in all cases diminish the influence of the subsequent ratio on the ICER for subgroup L. Technologies with largely up-front costs and long term benefits will therefore tend to have higher ICERs following discounting, while technologies with relatively flat long term costs and declining QALY benefits will tend to have lower ICERs[13].

This is shown in Figure 2. In view of the ambiguous impact of discounting upon the common and subsequent ratios, these ratios are assumed to remain unchanged. However, the rays representing the common and subsequent ratios become shorter under discounting, shifting from *Oa* to *Od* and from *ab* to *cd* respectively. Critically, because the costs and QALYs which comprise the subsequent ratio are discounted more heavily than those which comprise the common ratio, the ray representing the subsequent ratio is shortened proportionally more. As a result, the subsequent ratio is relatively less influential in determining the ICER for subgroup L. This loss of influence pulls the ICER for subgroup L closer to that for subgroup S, narrowing the RDCE[14]. The higher the discount rate, the further the RDCE is narrowed, reducing the scope for *Oe* to lie within the RDCE.

Discounting therefore reduces the scope for discrimination on the basis of life expectancy, with this scope diminishing with increases in the discount rate. Since the 3.5% discount rate used by NICE has been argued by many authors to be too high (Brouwer et al., 2005; Claxton et al., 2006; Gravelle et al., 2007; Paulden & Claxton, 2009), a strong case can be made that NICE's current discounting policy actively (if unintentionally) reduces the scope for discrimination.

[12] Each ratio will rise or fall depending on the specific distribution of incremental costs and incremental QALYs within the ratio – where the costs tend to be incurred later than the QALYs the ratio will likely fall, and vice versa.

[13] Brouwer et al. (2005) give some examples of the impact that discounting can have on the ICERs for treatments with long term costs or QALY benefits.

[14] This is also true in cases where the subsequent ratio is lower than the common ratio, raising the possibility of discrimination against those with longer life expectancy. In such cases, the undiscounted ICER for subgroup *L* (represented by *Ob*) would lie above that for subgroup *S* (*Oa*), but *Oc* would lie below *Ob* (rather than above, as drawn in Figure 2). Discounting would still narrow the gap between the ICERs, reducing the scope for discrimination.

"End of life" treatments

In January 2009, following the establishment of a National End of Life Care Programme by the Department of Health, recommendations from its Citizens Council (NICE Citizens Council, 2009) and a public consultation, NICE issued supplementary guidance on evaluating treatments which satisfy all of the following criteria: (a) the treatment is indicated for patients with a short life expectancy, normally less than 24 months; (b) there is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared to current NHS treatment; (c) no alternative treatment with comparable benefits is available through the NHS; and (d) the treatment is licensed, or otherwise indicated, for small patient populations (NICE, 2009a). Where a treatment satisfies these "end of life" criteria and the ICER for those patients with short life expectancy is above the cost-effectiveness threshold, the Appraisal Committee now considers "the magnitude of the additional weight that would need to be assigned to the QALY benefits in this patient group for the cost-effectiveness of the technology to fall within the current threshold range" (NICE, 2009a), in other words the weight which would need to be applied to the (discounted) QALYs of patients with short life expectancy for the ICER for such patients to fall below £30,000 per QALY. If the committee regards this weight as acceptable then the treatment may now be approved for those patients even though the ICER is above the cost-effectiveness threshold. For example, in NICE's recent appraisal of Lenalidomide for the treatment of multiple myeloma, the committee identified the weight for the relevant subgroup of patients as being approximately 1.4 (corresponding to an ICER of £43,800 per QALY) and deemed such a weight to be acceptable (Wishart, 2009). Lenalidomide was subsequently approved for those patients despite the ICER lying well above NICE's cost-effectiveness threshold (NICE, 2009b).

This new guidance can be incorporated into our model. Suppose that, following treatment, the life expectancy of those patients in subgroup S is less than 2 years, while the life expectancy of those in subgroup L is greater than 2 years ($p < 2 < q$). For treatments which satisfy the "end of life" criteria, NICE would now apply a weight to the QALYs for patients in subgroup S but not to those for patients in subgroup L. It would therefore discriminate against patients with shorter life expectancy if (from 14):

$$\sum_{t=1}^{q} \frac{\Delta c_t^L}{(1+\partial)^{t-1}} \Big/ \sum_{t=1}^{q} \frac{\Delta h_t^L}{(1+\partial)^{t-1}} < \lambda < \sum_{t=1}^{p} \frac{\Delta c_t^S}{(1+\partial)^{t-1}} \Big/ \Omega. \sum_{t=1}^{p} \frac{\Delta h_t^S}{(1+\partial)^{t-1}} \qquad (15)$$

where Ω denotes the highest acceptable weight that NICE is willing to assign to the QALYs of those with shorter life expectancy.

Equivalently (and perhaps more intuitively), NICE could instead be viewed as adopting a higher cost-effectiveness threshold for patients with shorter life expectancy, λ^S, rather than placing a greater weight on their QALYs. In this case NICE would discriminate against patients with shorter life expectancy if (from 15):

$$\sum_{t=1}^{q} \frac{\Delta c_t^L}{(1+\partial)^{t-1}} \Big/ \sum_{t=1}^{q} \frac{\Delta h_t^L}{(1+\partial)^{t-1}} < \lambda < \lambda^S < \sum_{t=1}^{p} \frac{\Delta c_t^S}{(1+\partial)^{t-1}} \Big/ \sum_{t=1}^{p} \frac{\Delta h_t^S}{(1+\partial)^{t-1}} \qquad (16)$$

where $\lambda = \lambda^S/\Omega$ and $\Omega > 1$.

In Figure 2 this higher cost-effectiveness threshold, λ^S, is represented by *Of*. For "end of life" treatments, where the subsequent ratio is lower than the common ratio (as in Figure 2), discrimination against those in subgroup S now requires that both *Oe* and *Of* (rather than only *Oe*) lie within the RDCE, narrowing the scope for ageist discrimination. Furthermore, if *Oe* lies below the RDCE and *Of* lies above the RDCE then NICE will now discriminate against patients with longer life expectancy, where previously it would not. Alternatively, where the subsequent ratio is higher than the common ratio (and so the RDCE lies above the ICER for subgroup S and below the ICER for subgroup L, unlike in Figure 2), NICE's supplementary guidance widens the scope for discrimination against those with longer life expectancy – such discrimination now occurs not only if *Oe* lies within the RDCE but also if *Oe* lies below the RDCE and *Of* lies within the RDCE.

Other considerations of the Appraisal Committee

Although useful for our modelling, the assumption that NICE only considers ICERs when making decisions is grossly over-simplistic. NICE has repeatedly stated that such evidence is used to aid and not replace the judgement of the Appraisal Committee, and NICE is acutely aware both that QALYs may not be plausible in specific contexts and that the models used to calculate the ICERs often incorporate uncertain data and occasionally dubious assumptions. NICE regards its decision making as a deliberative process and makes considerable use of patient and other stakeholder representation. As such, NICE is less likely to discriminate between subgroups on the basis of small differences between ICERs than would be implied by a strict interpretation of the conditions above. This further reduces the scope for ageist discrimination.

Discussion

We have sought to investigate a particular argument made against NICE: that its use of the QALY is "ageist" by discriminating against those with shorter life expectancy. Using a simple model of NICE's decision making setting, we have shown that ageist discrimination is neither inherent nor inevitable and requires the following conditions to hold for the technology in question: (a) the ratio of costs to QALYs must be lower over the long term than over the short term; (b) the ICER for patients with shorter life expectancy must lie above the cost-effectiveness threshold; and (c) the ICER for patients with longer life expectancy must fall below the cost-effectiveness threshold[15]. These conditions must hold after taking into account discounting and – where applicable – NICE's recent policy on appraising "end of life" treatments, both of which reduce the scope for ageist discrimination. Even where technologies satisfy these conditions, the Appraisal Committee has the authority to exercise discretion and to consider any possibility of ageism bias. It should also be noted that, where technologies fail to satisfy any of conditions (a) to (c), NICE's guidance may in fact discriminate against patients with *longer* life expectancy. As such, Harris's claim that NICE is ageist seems very difficult to substantiate. A fairer assessment of NICE would be that – except in the case of "end of life" treatments – it does not systematically discriminate on the basis of life expectancy. In the case of "end of life" treatments, NICE's recent supplementary guidance quite intentionally (and paradoxically) increases the scope for discrimination against patients with longer life expectancy by weighting the QALYs of patients with shorter life expectancy higher.

Harris states that "in every case QALYs are indeed, and by definition, inherently ageist and also favour those with the greater life expectancy regardless of age". While this would generally be true if the only QALYs considered were the remaining QALYs in each patient's life, this is not a basis on which NICE and similar organisations make recommendations. NICE instead considers the incremental QALY benefit resulting from treatment – where a technology results in a greater improvement in health-related quality of life for more seriously ill patients, or results in a greater extension to life in patients closer to death, the QALY benefit of the technology may in fact be greater for those patients who have fewer remaining QALYs. Furthermore, when making recommendations NICE also considers the costs of treatment and not just the QALY benefit. While saving the life of a patient with long life expectancy tends to yield a larger QALY benefit, the long term costs associated with any such treatment will generally also be greater. Indeed, the ICER associated with treating that patient may then be as high (or higher) than that for treating a patient with shorter life expectancy. It is these ICERs – not the incremental QALY benefits alone, and certainly not the remaining QALYs in a patient's life – that form an important (but not the only) consideration in NICE's decision making process.

In short, it does not logically follow that NICE is ageist simply because it considers QALYs as part of its decision making. Ageism in NICE's decision making is neither inherent nor inevitable. NICE's only policy which actively addresses discrimination on the basis of life expectancy is its recent supplementary guidance on "end of life" treatments. This policy arose as a direct consequence of the context in which NICE operates – in response to explicit political imperatives from Parliament and following consultations with the public – and raises the scope for discrimination against those with longer, not shorter, life expectancy.

---

[15] While condition (a) is implied by (b) and (c), it deserves to be noted separately for its intuition.

Policy implications

Our analysis has important implications for bodies such as NICE who adopt a time-dependant measure of health (such as the QALY) in their decision making. The first is that the kind of ageism which can arise through the application of cost-effectiveness methods is discrimination against individuals or groups of individuals on the basis that they have shorter *life expectancy* than others, irrespective of their age. Ageism may thus arise in the case of young people and mere age is an unreliable indicator of the circumstances under which it may exist.

However, the main implication of our analysis is that such discrimination is neither inevitable nor inherent. It can only arise when the technology under consideration satisfies the following conditions:

a) The ICER for a subgroup of patients with relatively short life expectancy lies above the cost-effectiveness threshold; and
b) The ICER for a subgroup of patients with longer life expectancy falls below the cost-effectiveness threshold.

For these conditions to hold requires that the long term ratio of costs to QALYs associated with the technology be lower than the short term ratio of costs to QALYs. Such technologies are likely to display specific identifiable characteristics, such as high upfront costs and/or continuous long term health benefits. Where such characteristics are identified, a consideration of the possibility of ageism ought to inform the deliberations of the advisory committee(s) responsible for drafting guidance**.**

References

Butler, R.N. 1975. Why Survive? *Being Old in America*. Harper & Row, New York.

Brouwer, W.B.F., Niessen, L.W., Postma, M.J., Rutten, F.F.H. 2005. Need for differential discounting of costs and health effects in cost effectiveness analyses. *British Medical Journal* 331, 446-448.

Claxton, K., Culyer, A.J. 2006.Wickedness or folly? The ethics of NICE's decisions. *Journal of Medical Ethics* 37, 373-378.

Claxton, K., Culyer, A.J. 2007. Rights, responsibilities and NICE: A Rejoinder to Harris. *Journal of Medical Ethics* 33, 462-464.

Claxton, K., Culyer, A.J. 2008. Not a NICE fallacy: A reply to Dr Quigley. *Journal of Medical Ethics* 34, 598-601.

Claxton, K., Paulden, M., Gravelle, H., Brouwer, W.B.F., Culyer, A.J. 2010. Discounting and decision making in the economic evaluation of health care technologies. *Health Economics*, forthcoming.

Claxton, K., Sculpher, M., Culyer, A.J., McCabe, C., Briggs, A., Akehurst, R., Buxton, M., Brazier, J. 2006. Discounting and cost-effectiveness in NICE - stepping back to sort out a confusion. *Health Economics* 15, 1-4.

Gravelle, H., Brouwer, W.B.F., Niessen, L.W., Postma, M.J., Rutten, F.F.H. 2007. Discounting in economic evaluations: stepping forward towards optimal decision rules. *Health Economics* 16, 307-317.

Harris, J. 1988. More and better justice. in Bell, J.M., Mendus, S. (editors) *Philosophy and Medical Welfare*. Cambridge University Press, Cambridge 75-96.

Harris J. 2005a. It's not NICE to discriminate. *Journal of Medical Ethics* 31, 373.

Harris, J. 2005b. Nice and not so nice. *Journal of Medical Ethics* 31, 685.

Harris J. 2006. NICE is not cost effective. *Journal of Medical Ethics* 32, 378-380.

NHS. 2005. *Directions and Consolidating Directions to the National Institute for Health and Clinical Excellence*. Downloaded on 12/11/09 at http://www.nice.org.uk/aboutnice/whatwedo/niceandthenhs/directions_from_the_secretary_of_state.jsp.

HM Treasury. 2003. *Appraisal and Evaluation in Central Government*. HM Treasury, London.

NICE. 2008. *Guide to the Methods of Technology Appraisal*. NICE: London.

NICE. 2009a. Appraising life-extending, end of life treatments. NICE: London. http://www.nice.org.uk/media/88A/F2/*SupplementaryAdviceTACEoL*.pdf. NICE: London.

NICE. 2009b. Lenalidomide for the treatment of multiple myeloma in people who have received at least one prior therapy. *Technology appraisals TA171*. http://www.nice.org.uk/nicemedia/pdf/TA171Guidance.pdf

NICE Citizens Council. 2009. *Departing from the Threshold*, London: NICE.

Paulden M, Claxton K. 2009. Budget allocation and the revealed social rate of time preference for health. *CHE Research Paper 53*. Centre for Health Economics: University of York.

Quigley, M. 2007. A NICE fallacy. *Journal of Medical Ethics* 33,465-466.

Wishart, A. 2009. *Price of Life*. http://www.bbc.co.uk/programmes/b00l9dmw. BBC: London.