Utilitas

Date of delivery:

Journal and vol/article ref: UTI 1600019

Number of pages (not including this page): 23

This proof is sent to you on behalf of Cambridge University Press. Please print out the file and check the proofs carefully. Please ensure you answer all queries.

Please EMAIL your corrections within 3 days of receipt to:

Nigel Hope <nrhope@aol.com>

Authors are strongly advised to read these proofs thoroughly because any errors missed may appear in the final published paper. This will be your ONLY chance to correct your proof. Once published, either online or in print, no further changes can be made.

NOTE: If you have no corrections to make, please also email to authorise publication.

• The proof is sent to you for correction of typographical errors only. Revision of the substance of the text is not permitted, unless discussed with the editor of the journal. Only one set of corrections are permitted.

• Please answer carefully any author queries.

• Corrections which do NOT follow journal style will not be accepted.

• A new copy of a figure must be provided if correction of anything other than a typographical error introduced by the typesetter is required.

Copyright Form: http://journals.cambridge.org/images/fileUpload/documents/UTI_ctf.pdf

• If you have problems with the file please email jharman@cambridge.org

Please note that this pdf is for proof checking purposes only. It should not be distributed to third parties and may not represent the final published version.

Important: you must return any forms included with your proof. We cannot publish your article if you have not returned your signed copyright form

Please do not reply to this email

NOTE - for further information about Journals Production please consult our FAQs at http://journals.cambridge.org/production_faqs
Author queries:
Q1: The distinction between surnames can be ambiguous, therefore to ensure accurate tagging for indexing purposes online (eg for PubMed entries), please check that the highlighted surnames have been correctly identified, that all names are in the correct order and spelt correctly.

Typesetter queries:

Non-printed material:
Comparative valuation of different policy interventions often requires interpersonal comparability of benefit. In the field of health economics, the metric commonly used for such comparison, quality adjusted life years (QALYs) gained, has been criticized for failing to respect the equality of all persons' intrinsic worth, including particularly those with disabilities. A methodology is proposed that interprets ‘full quality of life’ as the best health prospect that is achievable for the particular individual within the relevant budget constraint. This calibration is challenging both conceptually and operationally as it shifts dramatically when technology or budget developments alter what can be achieved for incapacitated individuals. The proposal nevertheless ensures that the maximal achievable satisfaction of one person’s preferences can carry no more intrinsic value than that of another. This approach, which can be applied to other domains of social valuation, thus prevents implicit discrimination against the elderly and those with irremediable incapacities.

INTRODUCTION

Current practice in health economics in measuring intervention effectiveness compromises equality by using an absolute standard of full health to which many cannot aspire. The challenge, set out in section I, is to define an alternative metric that is flexible enough that it can be used to compare treatment outcomes across individuals with very different capacities to benefit. Section II proposes a solution involving calibration of gains to individuals’ own maximal potential health prospect. In section III, the adequacy of the QALY under the proposed calibration to its role in social decision-making is assessed, in particular the requirement of homogeneity of intrinsic social value. Section IV considers the practical implications and applicability of this approach to the assessment of health treatments in a way that respects equality.

I. THE EQUITY CHALLENGE

There is widespread consensus in the practice of health economics that allocative decisions can usefully be supported by cost-utility analysis, and that the quality adjusted life year (QALY) is an appropriate unit of health-related utility. To the extent that it represents an accurate measure of social value, the QALY enables fair comparison of the effectiveness of different interventions in a range of contexts, including the promotion of health, the prevention of illness and accident, the
cure and the palliation of ill-health. Its use acknowledges that health benefits can involve either extending life expectancy or enhancing quality of life or both, and that beneficiaries rationally trade these different aspects of health benefit.

In this methodology, the concept of quality of life is implicitly defined by the method used to elicit substitution rates between quality of life and its duration. The QALY gain or loss from an intervention for a representative patient is estimated through two distinct steps:

i. Health Prospect Description. The patient’s health prospect in the absence of intervention is described, where a health prospect is the array of prospective health states with their probabilities of occurrence, and health states are characterized using a standard tool, such as the EuroQoL-5D. The EuroQoL-5D describes health states using five dimensions of health related quality of life: Mobility, Self-care capability, Ability to carry out usual activities, Pain/discomfort, Anxiety/depression. A health state is described by assessing whether it involves ‘no’, ‘mild/moderate’ or ‘severe/extreme’ problems on each of the five dimensions. (See <www.euroqol.org/>). Similarly, the patient’s health prospect with the intervention is described. An intervention’s effect is to move the patient from one health prospect to another.

ii. Health State Scoring. Each relevant health state is given a Health Related Quality of Life (HRQoL) score relative to full health, where full health is assigned a value of 1, and states equivalent to being dead are assigned a score of 0.

Taking account of duration and probability of each state in the array of states in the health prospects faced by an individual respectively with and without the intervention, the number of QALYs promised by an intervention can be derived. (Thus, a sure life extension of one year in full health generates 1 QALY, as does a certainty of raising someone in a health state with HRQoL = 0.5 to full health for two years.)

The assignment of different health states to HRQoL levels, step ii, is generally derived from a representative sample of members of the general public. It is here that the challenge to equality of human worth arises.

To assign a HRQoL to a specific health state, the sample of the general public is asked to respond as if they were in that state either to the question

What risk of death they would accept to achieve full health (this being the Standard Gamble, SG, elicitation technique)

or to the question
Calibrating QALYs to Respect Equality

What proportion of their imagined expected life-term (usually standardized for the exercise at ten years) would they forego in order to achieve full health (this being the Time Trade Off, TTO, elicitation technique).

The problem lies with the implicit interpretation of ‘full health’. With a sample drawn from the general public, ‘full health’ is likely to be interpreted as full psychophysical function. This is problematic when applied to interventions that might benefit those with permanent irremediable disabilities. So measured, the maximum possible health gains for such patients, from interventions that necessarily only partially mitigate disability or that extend life without addressing disability, will fall short of what can be gained by those without disability. Their health is capped at a level that will receive a HRQoL score of less than 1; for thus are such disabled states scored by the general public. Hence life extensions or best-possible cures for illnesses for the disabled will be assigned fewer QALYs and thus less social value than corresponding gains for the able-bodied.

The practice of health economics often appears to assume that in the QALY we have a measure of a social good – health gain – that is clearly of uniform intrinsic social value. (Intrinsic in contradistinction to any extrinsic value that the health of an individual may bear, positive or negative, for example in enhancing tax receipts, or in mitigating inequality.) However, as currently calculated, the QALY is akin to a physical measure of health gain; to turn it into one appropriate for social valuation requires consideration of what contribution different health improvements make to what is of intrinsic social value. This issue is considered further in section III; it plausibly makes reference to individuals’ own rational aspirations for their lives (their welfare). If that is right, use of average valuation is problematic.

The problem lies in the transfer of the valuation of a gain elicited from one population group (the representative sample) to other groups (the individual patients who will be the beneficiaries of the intervention). Whilst benefit transfer, a common and very useful technique in normative economics, inevitably ignores much heterogeneity, systematic bias should be avoided. Transferring valuations from a representative cross-section of the population for use in valuing benefits for an irremediably disabled sub-section, involves systematic bias.

That use of standardized valuations to assess interventions for the permanently disabled in health technology appraisal is problematic has been widely recognized. And this critique has been used to discredit cost-utility appraisals. An article in The Lancet put it thus:

The QALY outcome measure has problems. Even if a life-year in which a person has impaired mobility is worse than a healthy life-year, someone adapted to
wheelchair use might reasonably value an additional life-year in a wheelchair as much as a non-disabled person would value an additional life-year without disability. Allocators have struggled with this issue.1

Within the health economics literature, the problem has generated a strand of analysis and various proposals. No such proposal has yet found its way into general practice of health technology appraisal, which therefore remains vulnerable to the claim of systematic bias against the disabled.

The most straightforward proposal was offered by Eric Nord with colleagues:

[F]or states of chronic illness or disability that are preferred to death, all saved life years count as one.2

To resolve the issue, we have suggested that all life years gained by disabled people should count as 1...3

This proposal, which is termed by its authors the 'equal value of life approach' (EVL), is however only a partial solution, as it does not address the valuation of an intervention somewhat to raise the quality but not to extend the years of someone with a permanent disability. Once we decide to accord equal value to the life of the disabled, it becomes appropriate to assign full value to the maximal functional state of someone with a permanent disability and to value functional recovery to that state following an illness as full recovery.

Hence, to make the use of a QALY metric acceptable, an alternative understanding of Health Related Quality of Life is required, one that does not implicitly rate the lives of those with less than perfect psychophysical functional ability as inferior.

The challenge is complicated by the fact that those with disabilities are often stuck in health states from which others can gain from treatments; indeed, it is for the appraisal of such treatments that the HRQoL and QALY methodology is designed. Further, even those with disabilities that are currently untreatable could gain from conceivable treatments. From this perspective the apparent full recovery for the permanently disabled is not after all a truly full recovery.

This creates a dilemma in assignment of states to HRQoL: to respect equality, we must accord all lives full worth; yet to motivate cure for those with temporary disability, and to motivate the search for cure for those with apparently permanent disability, we must recognize that some health states are less good than others.

Before developing a technical solution to this problem in section II—on which admittedly involves a degree of conceptual acrobatics—we should address three alternative approaches that might be thought to dissolve the problem: Dworkin’s proposal to select the basket of health treatments through a hypothetical health insurance optimization; use of equity weights to increase the valuation of treatments that benefit the disabled; and aversion to elicited societal preferences regarding valuation of interventions benefiting the disabled. Consider each in turn.

Suppose that health treatment appraisal is conceived from the perspective of a healthy individual maximizing her potential welfare in the face of a set of risks, and wishes to select the optimum basket of health services so to do. Use of an absolute scale of psychophysical function to appraise different treatments is in such a case appropriate as no interpersonal comparisons are involved. Compare for example a treatment that would give fourteen healthy years following a heart attack and another treatment that would yield fifteen traumatic years under treatment for cancer.\(^4\) If the risk of heart attack and the risk of cancer are similar, and the pain and suffering attending the cancer cure are unpleasant enough, it would be rational for someone to choose to include the heart treatment rather than the cancer treatment in the portfolio of treatments (assuming that she cannot afford both). There is no required judgement of the relative value of the life years of two different persons, respectively with cancer and heart disease: for the comparison here is of the possible outcomes for one individual.

In his ‘Justice in the Distribution of Health Care’, Ronald Dworkin proposes the use of a thought experiment of this kind, in which a society with an equitable distribution of resources chooses what basket of health care services it should fund, as a guide to the basket of services that should now be funded publicly (or mandated for social insurance funding). He argues that

\(^4\) I am grateful to Richard Cookson for the challenge presented by this example, which is picked up again at the end of section III.
that people in the community would make for themselves, one by one, in the
appropriate circumstances, if they were looking from youth down the course of
their lives and trying to decide what risks were worth running in return for
not running other kinds of risks.\textsuperscript{5}

Dworkin stipulates that for the experiment to work in guiding
community choices, you must

\texttt{[I]magine that no one in your community – including insurance companies – has
any information available about the antecedent probability of any particular
person contracting any particular disease or infirmity that he or she does
not evidently already have. No one would be in a position to say, of himself
or anyone else, that that person is more or less likely to contract sickle-cell
anemia, or diabetes, or to be the victim of violence in the street, than any other
person.}\textsuperscript{6}

It might be argued that the selection of treatments within a budget
constraint so as to maximize expected QALYs calibrated
relative to full psychophysical function is precisely what one would
expect of someone ignorant of their particular set of health needs.
Dworkin argues that nothing that that hypothetical society would do,
‘by way of health care arrangements, is open to objection on grounds
of justice’ and ‘what they would do through independent decisions can
serve as a guide to what we should do, in whatever way we can, to
improve justice in our own circumstances’.\textsuperscript{7}

However, Dworkin himself excludes from the ignorance stipulation
to which his hypothetical purchaser of health insurance is subject only
risk of ‘disease or infirmity that he or she does not evidently already
have’. Hence, in drawing lessons from the hypothetical choices, we will
still have to confront the challenge of designing a basket of services
that would fit the choices of those with existing permanent disabilities
as well as those who have none.

Dworkin does not spell out the rationale for thus suggesting that
the disabled should choose with knowledge of their disability, but
presumably it is that those who are actually disabled could otherwise
deny the relevance of the thought experiment: they are not in a position
to choose a basket of services that would have mitigated the risk of or
the consequences of the disability from which they suffer, so it is unjust
to insist that they respect the decisions that they might have made had
they been in such a circumstance. If the community is to act on their
behalf it must recognize that their conception of full health differs from
that of others.


\textsuperscript{6} Dworkin, ‘Justice in the Distribution of Health Care’, p. 889.

\textsuperscript{7} Dworkin, ‘Justice in the Distribution of Health Care’, p. 890.
Another response to the challenge is to stick with psychophysical function as the interpretation of ‘full health’ but to attach additional value to treatments that benefit those who are disadvantaged through disability.

Use of such weights would be appropriate were the problem one of a failure to move towards equality of outcome.

For egalitarian reasons it may be appropriate to compromise total health gain in order to achieve a fairer distribution of health. But that egalitarian question cannot even be posed without a prior assessment of what health gain has and could be achieved for different individuals using interpersonally comparable units. Calibration of health states to respect equality of human worth is required to generate units of uniform intrinsic social value. Hence, the problem is a failure appropriately to calibrate health gain in a way that respects the equal value of lives lived with disability.

A third possible route to addressing the challenge is to advert to elicited societal preferences, using such preferences to overlay QALY measurement with societal valuation of QALY gains in different circumstances. Peter Ubel and colleagues attempt to address the confrontation between the standard approach to QALY measurement (described above) with ‘a preference . . . for avoiding discrimination against people who have limited treatment potential’ as follows:

For example, suppose the public thinks that saving the lives of people with paraplegia is equally as important as saving the lives of people who can be returned to full health . . . saving the life of either group of patients [would then] bring 1 QALY per patient [per year of life extension]. However, this rescaling would also force us to conclude that people with paraplegia have the same quality of life as people without paraplegia and that curing paraplegia would not improve HRQoL. (No HRQoL is gained by ‘improving’ patients from an HRQoL of 1.0 to an HRQoL of 1.0.)

Here is the attempted solution:

Suppose the societal value of program A [which cures 100 people of a life threatening illness returning them to full health] is given an arbitrary value of 1.0 [per patient]. Now suppose people think that program B, which saves the lives of 100 people [who remain] with paraplegia, should receive the same priority for funding as program A . . . Now suppose the same people think that curing 600 patients of paraplegia (program C) is equally as important as saving 100 otherwise healthy people’s lives (program A). Program C therefore has . . . a societal value of 0.16 [per patient] . . .

What is the benefit of separating societal value from HRQoL measurement? . . . the HRQoL brought by programs B and C (in conventional CEA [Cost Effectiveness Analysis]) must sum to the number of QALYs brought by program A (1 QALY [per person year]). However, . . . the societal value of programs B and C add to 1.16 . . . We can now say that saving the lives of people with paraplegia
is equally as valuable as saving other people’s lives while still acknowledging that it is beneficial to cure people of paraplegia.\textsuperscript{8}

The problem with this proposal is that it creates a paradox, which the authors half-acknowledge:

Some may worry that the societal value approach to QALYs is inconsistent because it allows the value of programs like B and C to sum to a value $>1$. However, if society places the same value on saving the life of a paraplegic and saving the life of a non-paraplegic and if society also values the cure of paraplegia, then our suggestion is consistent with societal values.

The defence does not answer the worry that the advocated approach is inconsistent. Consider how it would value a programme D, one that was able to save the lives of those with paraplegia threatened by some fatal illness and cure their paraplegia at the same time. Programme $D = B + C$ in its effect, so should be valued at 1.16 per life year – apparently making the programme more valuable than programme A (which saved the lives of non-paraplegics), notwithstanding that both programmes A and D leave their beneficiaries, who would all otherwise die, with exactly the same HRQoL.

If the Ubel \textit{et al.} scheme is inconsistent, then that it reflects society’s values merely shows society’s values also to be inconsistent. Nord \textit{et al.} suggest that ‘we may be forced to accept [such] inconsistencies rather than impose a framework that clearly violates social preferences’.\textsuperscript{9} However, it is better to strive for an interpretation of citizens’ expressed preferences that is both consistent and free of ethical errors.

The paradox is dissolved by systematically distinguishing between those whose incapacity is irremediable and those whose condition is curable, in other words by calibrating health assessment to beneficiaries’ varying capacity to benefit.

\section*{II. CALIBRATING TO BEST ATTAINABLE HEALTH STATE}

This section develops ‘A Proposal to Solve the Comparability Problem in Cost-Utility Analysis’, by Bleichrodt, Herro and Pinto. They recognize that valuation of health states should vary systematically with the ‘attainable health state’ of each individual:

The optimal solution . . . is to determine for each individual his set of attainable health states and to elicit his health utility function, scaled such that the utility of death is equal to zero and the utility of his best attainable health state is equal to one. This scaling ensures that a year in the best attainable health

\begin{itemize}
\item \textsuperscript{9} Nord \textit{et al.}, ‘The Value of Life’, p. 875.
\end{itemize}
state gets the same weight across individuals and avoids the possibility of discrimination due to differences in capacity to benefit.  

But there is work to be done. First: how do we conceptualize this scaling in the presence of uncertainty regarding attainable health state? Second, what is attainable depends not only on technology but also on budget. Third, taking this approach to avoid discrimination of the disabled seems paradoxically to inhibit us from attributing a value to possible mitigations of incapacity. And finally, what if those whose prospects are very bleak – is individual scaling still appropriate?

My proposal involves four steps to address these issues:

(1) calibration to best health prospect (rather than health state),
(2) respecting the budget constraint of the particular budget holder when assessing best health prospect,
(3) recalibration when changes in technology or budget expand (or constrict) the prospects achievable,
(4) exempting from calibration to best health prospect lives of little quality and all consideration of duration.

(1) **Calibrating to personal best health prospect**

When estimating the value of health interventions that shift individuals from one prospect to another, calibration to best attainable health permits sensitivity to the varying capacity to benefit of different individuals. However, as treatment decisions are essentially forward-looking, and as health outcomes are probabilistic, it is important that calibration is to a best attainable health prospect rather than a best attainable health state (a departure from the Bleichrodt et al. proposal).

We adopt:

**Principle One: calibration to personal best attainable health prospect.** In the assignment of cardinal values to individuals’ health prospects (Health Related Quality of Life $HRQoL_{hp}$) for each relevant future period for the purpose of comparative appraisal of possible interventions, where a prospect is an array of possible health states $hs$ each associated with a probability $p_{hs}$ such that $\sum p_{hs} = 1$:

- the value zero is assigned to the health state of being dead,
- the value 1 is assigned to the best health prospect achievable by that person for that period, where health prospects are ordered according to the individuals’ own rational preferences,
- intermediate values are assigned to a health prospect $hp$ for a period, according to individuals’ rational indifference between $hp$ and a

---

prospect involving a probability HRQoLhp of best attainable health
prospect and a probability \((1 - \text{HRQoLhp})\) of death,

- HRQoLhp is then a coefficient applicable to the duration of the
affected temporal parts of individuals' health prospects with/without
an intervention, thus generating a QALY gain or loss attributable to
the intervention, e.g. a 0.5 gain in HRQoLhp for a period of a year
would represent a ½ QALY gain.

Hence, for somebody with both a remediable and an irremediable
condition, a best attainable health prospect for that person is defined
to remedy the remediable but to include the irremediable condition.

Consider first appraisal of a screening programme to mitigate the
risk of a life-threatening condition. To avoid discrimination against
people with disabilities, the assessment of quality of life used in
calculating incremental QALYs conferred by the screening programme
must be calibrated to the best health prospect that is available to
each person affected. The years that are in jeopardy from the life-
threatening disease are full value years even if they will be enjoyed
by someone whose maximum gain cannot include sight, no technology
being available that can restore sight.

Now, consider the position of a person whose blindness is curable.
Attainment of sight is attainable for this person, so we should assess
her options against this standard. Her blindness is not part of her
endowment but a condition for which she seeks treatment, and hence
we should calibrate other prospects against a health opportunity set
that includes a chance of full sight.

Suppose this second blind person suffers from cataracts that can
be removed, but the operation has not yet occurred. Suppose she is
also at risk of developing a life-threatening condition. Is screening her
mitigate this risk as valuable as screening a sighted person? Yes,
but it is less cost-effective, for two procedures are required to achieve
the same outcome (sight with mitigated risk). The fact that a given
health prospect for some people involves more cost than for others is
incontrovertible and should properly inform appraisal.

(2) Respecting the budget constraint

Suppose the same scenario arises with a blind person whose blindness
is technically curable but for whom the budget is not available to cure
her: perhaps the cure for this person’s type of blindness falls above the
cost-effectiveness threshold consistent with the budget constraint.

If the budget constraint is binding, it is just as real a constraint as the
technological limit. Hence, this person's life years with blindness should
be attributed intrinsic social value equal to the person whose blindness
is technically incurable. So, in our example, it should be considered as
Calibrating QALYs to Respect Equality

Cost-effective to screen her for a life-threatening condition as to screen a sighted person. Hence, we must accept:

**Principle Two: maximal individual health prospect is defined with reference to any binding budget constraint.** In the assessment of the maximum health prospect achievable by an individual, the maximum is defined not only by technological limits but also by the budget constraint, assuming that the budget constraint is truly binding.

Note that the principle that individuals' best attainable health prospects are determined by taking account of the budget constraint applies only if the budget constraint is actually binding. If there is resource allocated to some other activity which can be reallocated into health, then that assumption does not hold, and the allocation between that other good and health must be set using some metric of value that encompasses both health and whatever good is realized by the other budget.

If a thousand pounds becomes available to the health budget, it should be used to fund those interventions that would maximize the calibrated QALYs gained. Amongst interventions that only improve quality of life without affecting longevity, the best intervention for an individual is the one that minimizes the HRQoL that that individual would then have reason to assign to their current health prospect.

Budgetary resources should be displaced from elsewhere just in case the social value that they are realizing in that alternative use falls short of the social value of the QALYs that would be created by their best health use.

Whose budget constraint should be applied? Budgets are particular to budget holders, and budgets are governed by decision-making individuals and protocols. The best attainable health prospect for a beneficiary will therefore vary with the resources available to the decision maker for that use.

Compare, for example, the relative valuation of different interventions carried out by an aid agency to the valuation carried out by officers of the impecunious health system that the aid agency wishes to support. The latter is forced to take as fixed incapacities that from the perspective of a wealthier donor are properly seen as ailments requiring cure.

This gives us:

**Principle Three: The budget constraint that defines beneficiary's best attainable health prospect is that of the current decision-maker.** Therefore, each funder of health care allocating its budget so as to maximize health gain must regard any unaffordable improvement of
health condition as an incurable condition not detracting from the
g Value of the lives afflicted.

Decision-making in a publicly funded health service, however, is
c complex: budgets are distributed across many different levels, and the
hardness of these budget constraints is often not clear. This may require
a decision-maker locally to countenance two conflicting perspectives:
on the one hand prioritizing locally on the assumption that the budget
is fixed, on the other making the case for budget expansion from an
understanding of what health gains could be achieved were the budget
restraint relaxed.

(3) Recalibration to best health prospect when changes in technology
or budget expand (or constrict) the prospects achievable

Suppose a technology emerges that would allow a particular form of
blindness to be corrected cost-effectively, the attainable quality of life
of those suffering from that form of blindness shifts upwards, and
it is correct to attribute value to the improvement of their lives to
that higher level. Potential gains of quality of life are thenceforward
reckoned as a proportion of the attainable quality of life including the
innovation.

Yet if the new technology is assessed to be cost-ineffective, the current
lives of that group continue to be viewed (e.g. in assessing other
interventions from which they might benefit) as of full quality.

The appearance of paradox arises because we fail readily to grasp
the radical shift in the status of the persons affected that occurs if the
innovation is accepted – from one of disability-demanding-respect to
one of affliction-demanding-treatment. A member of such a group is
now unable to realize her potential without intervention, because her
potential in life has expanded. Existing treatments become inadequate
to her need.

We may wish to say that the innovation of an affordable cure was
there all the time – awaiting discovery, and we merely underestimated
the potential quality of life of those afflicted by the condition. With
adoption of the innovation, it turns out that that was the true maximum
prospective quality of life. Such an account would explain why we
attribute value to the discovery itself.

A similar shift occurs when a budget expansion renders an existing
treatment affordable – or, in reverse, if a budget contraction forces
out of the basket of interventions an existing marginally cost-effective
treatment.

(4) From calibration to best health prospect lives of little quality and
all consideration of duration
Two objections to the proposal to calibrate quality of life to maximum potential function should now be addressed:

- The implications become counter-intuitive when applied to those with very low potential quality of life.
- The proposal responds to alleged discrimination against the disabled. But the QALY approach to measuring benefit is open to a similar objection on behalf of those who have less than average life expectancy.

We will consider each issue in turn.

Suppose someone is in a permanent vegetative state and the possibility emerges to restore to them some minimum level of brain function still short of consciousness. To the extent that that represents the maximum function that can be attained for such a person, it might seem to follow from the equality of persons that such an intervention should be reckoned as if it were conferring full quality of life.

To avoid this outcome, we could stipulate a minimum level of mental function as delimiting the ambit of the principle of equality. Social value might be thought to inhere in the existence or the doings of ethical persons, and there is philosophical space to deny ethical personhood to those lacking basic mental function. Whilst there may be separate and powerful ethical reasons to attribute value to the lives of the comatose, the demented and to tiny infants in most contexts, conditionality upon personhood shields from absurdity the proposal to calibrate maximum HRQoL level to the individual’s best attainable health prospect when appraising interventions.

Conditionality upon personhood is in turn supported by a minimal substantive understanding of what gives life its equal potential value. To avoid conditioning valuation upon one particular theory of life’s value, calibration for equality aims to value improved psychophysical function according to the importance that individuals themselves would attribute to it. (Calibrating against an average citizens’ assessment, by contrast, would implicitly ignore the potential of someone with some disability to find a distinctive yet equally valuable use for the time that they have available to them.) There is nevertheless a minimum level of functioning beneath which neither experiences nor activities of value are conceivable. The permanently comatose are capable of neither experiences nor activities of intrinsic value.

A more demanding minimum is set by Nord et al. in qualifying the Equal Value of Life proposal mentioned in section I: ‘To resolve the issue, we have suggested that all life years gained by disabled people should count as 1 as long as the health state in question is preferred to...’
being dead by those concerned.\textsuperscript{11} In some communities where a suicide wish is reckoned reasonable (rather than the product of a distress that observers may hope will prove temporary), intervention to prevent it is not thought appropriate. It is no doubt with such situations in mind that Nord et al. proposed the limitation to their EVL rule. The important point here however is that some lower limit is plausibly consistent with the principle of equality.

In section III, intuitions regarding our proposal at this difficult borderline are tested against an example.

Regarding the second issue, are we bound by the principle of equality also to calibrate potential gain in years of life to the maximum potential years of life available to each person?

Though such calibration would be technically feasible, the results would be strikingly counterintuitive.

Suppose of two patients facing a small risk of death, small enough that this is not a risk from which rescue is demanded (given that rescue creates its own ethical demands), the life expectancy of one of the patients is dramatically shorter: she has only a day or two to live before she will die from a pre-existing condition. It is apparent that an intervention to reduce the small risk of immediate death is more valuable for the person with greater life expectancy. And the reason is this: the capacity for experiences and activities of value is broadly proportionate to time; the doctrine of equality thus lacks plausibility over this dimension. Whether there is strict proportionality between time granted and value of life is less obvious, as we will discuss in the next section.

III. THE ROLE OF QALYS IN SOCIAL DECISION-MAKING

QALYs play a particular role in social decision-making: they represent units of impact that bear constant intrinsic social value, such that when appraising different options, an option yielding twice as many QALYs will yield twice as much intrinsic social value. In this section I first focus upon the theoretical requirements for value-homogeneity of the QALY metric. I then test the proposal against intuitions regarding minimal quality of life and the displacement of treatments delivering higher relative quality of life.

As Sen points out, to allow interpersonal comparisons of levels and of increments, achievement of social value must be susceptible to measurement for each individual on a ratio scale, and the origin of the scale must be non-arbitrary.\textsuperscript{12}

\textsuperscript{11} Nord et al., ‘The Value of Life’, p. 873.

Sen suggests a possible zero: ‘The interpretation [of ‘the “origin” of a person’s welfare function’] may be to identify a distinguished point below which misery dominates . . . ‘. He notes however that ‘other interpretations are possible’. It is natural to identify this ‘origin’ with the zero point on the HRQoL scale described in section I. This zero is commonly associated with ‘death’ rather than with ‘misery’, but death is actually inappropriate as an interpretation, being an event rather than a state, and irreversible. Sen’s ‘misery’ suggestion may be more helpful, understood, under the SG methodology, as a level of pain or other affliction such that there is no risk of death a person would be unwilling to accept to recover from a prospect in which that state is endured indefinitely.

In order to discharge a commitment to the equality of human worth, the intrinsic social value attributed to any individual’s welfare should also be scaled to a common maximum value for any given duration. (Regarding a maximum, Sen refers briefly to the possibility of prohibiting a set of welfare functions for interpersonal comparisons ‘that “blows up” the welfare function of one person arbitrarily keeping those of others unchanged’.) This ceiling on the intrinsic social value of individual welfare can be identified with the ceiling on the HRQoL scale, \( HRQoL = 1 \), attributable to any individual’s best attainable life prospect.

It might be objected that even a person enjoying their best attainable health prospect may yet lack adequate resources to flourish. Nevertheless, we can call in aid the point established in section II that calibration to maximum potential is relative to what is achievable by the budget-holder conducting an appraisal, and is thus subject to the resources at their disposal. For the allocator of healthcare resources, the allocation of other resources is taken as a given, so for that budget holder there is an equivalence between HRQoL and Quality of Life \textit{tout court}.

Indeed, this approach is not limited to healthcare decision-making. The methodology outlined in sections I and II makes no essential reference to health: for decision-makers with other interventions in their portfolio, health states, health prospects and attainable health states can be displaced by more general concepts of welfare. Individuals’ conceptions of what lends value to their lives, whether health or education or other goods, are incorporated into the methodology for assigning quality of life levels to different life states and life prospects relative to best attainable life prospect.

Sen, in his championing of an ethically broad-based measure of success in social policy, focuses upon ‘individual capabilities to do things that a person has reason to value’.

formal restrictions suggested above, the capability approach is open to an elitist interpretation, one that values the societal sum of capability calculated employing absolute measures of functioning. To insist on a maximum level of capability that is attainable is to deny the validity of perfectionist accounts of human worth. Notwithstanding that we may be confident that Jack will flourish less than Jill upon any and every objective scale, we take it as axiomatic that the social value of enabling Jack to flourish as much as he possibly can for a year is of no less value than doing the same for Jill. This is interestingly consistent with Jack actually wasting what opportunity he is given. The principle of equality applies in prospect not in retrospect – but the decision contexts with which we are dealing do not require retrospective comparison.14

This proposal is not to be understood as a compromise to meet egalitarian or fairness concerns, for we lack an alternative ethically tolerable metric of interpersonal value comparison. Rather it is intrinsic to our understanding of value that the social value of any individual’s life prospect is of no more value, year for year, than any other’s. To respect this intuition regarding valuation we therefore insist on calibrating welfare against a scale that not only has a significant zero but also has a significant maximum, set at the individual’s maximum achievable and affordable welfare prospect.

The social value of welfare must be measured in units that have a time dimension as well as a magnitude at a time. Elsewhere I have argued that a year’s gain in life expectancy (a statistical life year) is an attractive interpersonally comparable measure of roughly constant intrinsic social value, notwithstanding that the lives vary in longevity.15 But to use the QALY as a measure of homogeneous intrinsic social value is to assign equal value to each of a person’s prospective life years at best attainable health, as well as equal value to the life years of different people.

Assessing aggregate value by integrating over duration is inconsistent with John Broome’s contention that time is not separable in the social value function.16 However, separability of times, like separability of persons (for which Broome argues), is a simplification that may be adequate to certain decision contexts, though not to others. Broome is correct that it is not adequate to a context in which decisions are being taken that might affect the age structure of a whole community. For other questions, perhaps dealing with....

15 Donald Franklin, ‘Valuing the Time of Your Life’ (under review).
sustaining communal institutions such as the family, separation of
persons will also be inadequate. Nevertheless, for social decision-
making in the contexts here under consideration (the allocation of
healthcare resources, and similar goods and services affecting the
quality of the lives of existing groups), a reasonable simplification is
to extend equal concern to each period of a person’s life.

Let us now test this proposal against challenging scenarios involving
lives whose maximum health prospect even with best treatment is of
low quality, and whose treatment will displace that of individuals with
better prospects.

Imagine a condition that leaves its sufferers fully conscious, but in
a very poor physical condition. Suppose that this condition is rated
as preferable to being dead (or that the choice of being dead in this
community is not available). Now a treatment becomes available that
would improve quality of life to a very small, but real, degree, but
only at such an expense that it would only be reckoned affordable
if the improvement were calibrated in the way discussed such that
the improved state is deemed full health for this group (HRQoL = 1).
To fund such treatment might displace treatments from others whose
quality of life would be much higher subsequent to treatment at only
slightly higher cost per year of benefit. Can funding the treatment
nonetheless be justified?

For example, suppose the condition leaves sufferers in almost
constant severe pain – but with regular remissions of average
duration of fifteen minutes per day. Suppose the proposed treatment
could increase the period of remission to one hour. Suppose that
a conventional assignment of HRQoL levels to these two health
prospects – respectively without and with the treatment to extend
remission periods – would generate assignments of 0.025 and 0.1
(assuming for simplicity proportionality of life value with time in
remission). That is to say that a representative sample of the general
public would be willing to take a 97.5 per cent risk of death to avoid the
untreated prospect if the alternative were full psychophysical health,
and that they would be willing to take a 90 per cent risk of death to
avoid even the with-treatment prospect (to use the Standard Gamble
interpretation of the assignment of health prospects to HRQoL levels
described in section 1).

Four years with this condition would conventionally generate only
one tenth of a QALY \((0.025 \times 4 = 0.1)\), and the treatment would raise the
tally to 0.4 QALY \((0.1 \times 4 = 0.4)\) – a gain of 0.3 QALY. With calibration,
however, we would scale to the maximum achievable quality of life for
this group – which is the prospect with the treatment. To those with
this condition, suppose we elicit a willingness to take a risk of death of
75 per cent to get the benefit of the treatment and to avoid the prospect of remissions remaining at only fifteen minutes per day. From which we infer that, from the perspective of this group, the treatment raises HRQoL from 0.25 to 1, generating \( \frac{1}{4} \) QALY each year, three QALYs over four years.

Under the calibration-for-equality proposal, therefore, this treatment would be judged affordable in preference to one of equal per capita cost that generated an expected additional thirty-five months in full psychophysical health to some other group, e.g. through mitigation of a lethal heart condition.

Is it reasonable to attribute such a high value to a treatment that leaves sufferers in such a miserable condition?

Note first, remembering Dworkin’s insurance proposal, that if the chronic pain condition is merely one to which the general population is at risk, alongside the risk of heart disease, and the choice is which treatment to fund, then calibration is not appropriate. In that case, each member of the population is trading between risks of different outcomes, and each would rationally maximize uncalibrated QALYs in so doing.

Rather, we are envisaging that known sufferers of the chronic pain condition (like those who are blind from birth in the earlier examples) are members of the population for whom the basket of services is being chosen. (Whether this is appropriate for a particular condition may depend upon the age of incidence – a issue touched upon in section IV.) Their perspective with the condition is the one that is pertinent to the community’s choice of service basket: there is no legitimate argument based upon what choices they would have made from a full health perspective had they been merely at risk of the condition that they now irremediably have. They cannot be bound by such hypothetical choices.

As this is a question of interpersonal comparisons rather than of risk optimization for a single set of people, the group with this condition can invoke the equal value of human life in defending their claim to have the health gain on offer recalibrated to their maximum health prospect. If the intuition of equality is challenged by this example – attributing as it does equal value to a year with chronic severe pain only remitted for an hour a day and that of a year with no disability – the burden of proof is on those who would salvage some coherent account of a commitment to the equality of human life. Alternatively, the objector would have to abandon that principle, and then presumably countenance differentiation of the intrinsic value of life across all dimensions of function (intelligence, strength of interpersonal relationships, happiness of disposition *inter alia*).
It is clear that the proposed calibration of comparative health impact on different people conflicts with our unreflective judgement of comparative value of different health states – as these will unreflectively be based on an intrapersonal comparison.

For example, recall from section I that in the discussion of the Dworkin insurance model we considered treatments respectively for heart and for cancer patients, the heart treatment yielding fourteen untroubled years, the cancer treatment delivering fifteen years marred by significant pain and disability. We saw that so long as we are ignorant as to which group cancer risks apply, and to which group heart risks apply, then appraisal is for the general population. In which case, interpersonal comparisons of value are not involved and calibration for equality is not required, so the heart treatment can be selected for the basket of treatments. If, however, the two groups are distinct, calibration for equality requires us to view the best health prospect for each group to be of equal intrinsic value year by year, and the cancer treatment must be selected (at least if they involve equal total cost per person). This outcome may seem counterintuitive, perhaps because we slide between the two cases.

To press the point, suppose there are two separate groups of individuals, A and B, who are at equal risk of contracting respectively condition A and condition B, and all who contract these conditions would die with no treatment. (Again, it is important that these are two separate groups; if all were at risk of either disease, then intrapersonal comparative valuation would yield answers without the need for calibration.) Suppose for those who contract condition A treatment A1, which gives health state hs1 for one year, is preferred to treatment A2, which gives health state hs2 for three years, given the difficulties implicit in the latter state. The only available treatment for condition B is treatment B, which gives health state hs1, for two years. Given that hsB is the maximal health state for condition B sufferers, it must be assigned a HRQoL score of 1. If resource is available only to treat those at risk of condition A or those at risk of condition B, but not both, and treatment costs and other relevant impacts are the same, treatment A2 would be socially preferred to treatment B, as it yields three years to every beneficiary whereas treatment B yields only two years. As A1 yields only one year, were it taken to displace treatment A2 from the social reckoning (given that it is preferred by A), treatment B would be chosen. Yet, the potential for three years of life for group A is what is normative for the social decision-maker comparing the two treatments, notwithstanding that the outturn (given that group A will actually opt for treatment A1) will be fewer years in the same health-state than would have been achieved for group B. Whilst longevity
may be sacrificed for quality of life by individuals in group A with respect to their own lives, social cost-benefit analysis is constrained not to make such sacrifices interpersonally. Where there are different population groups for whom outcomes are being compared, calibration to maximum health prospect is required.

Perhaps this corollary of the principle of equality remains counterintuitive. The analogy with dexterity or height or intelligence or pulchritude or happiness of disposition may help to render this approach natural. These attributes are generally valued as goods in contexts in which we assess how much human beings can contribute to others’ lives (for example in recruitment for employment or selecting members for a sports team or a musical ensemble), or when seeking to avoid hazards that might jeopardize our capabilities in these dimensions, or in the exceptional cases where a deficit can be remedied (by a growth hormone, or by therapy of some sort, or by surgery). But for the most part, we accept that we are endowed with different levels of these attributes, and that that is our lot, without these differences rendering us less or more valuable intrinsically as human beings. Hence, when assessing relative effectiveness of health interventions, these differences are irrelevant. Irremediable health conditions should be treated likewise.

IV. IMPLICATIONS FOR THE PRACTICE OF HEALTH ECONOMICS

What difference should this proposal make in the practice of health economics?

Calibration to maximum health prospect ought in principle to have wide application: given the high incidence of multiple morbidity, treatments will often benefit large numbers of individuals with a range of other pre-existing conditions or disabilities; without calibration, the benefit of treatments for such individuals would be reckoned to fall short of benefits to otherwise healthy and able-bodied individuals.

In practice, however, health economics valuations are usually valuations of treatments, and appraisals of treatments generally assume that the target group is homogeneous. If a treatment is approved, it will be approved for a population, and ethical and legal norms governing the clinical context will prevent any discrimination against the disabled in authorizing or funding the treatment, whilst clinicians determining to whom to apply these treatments are sheltered from the direct application of cost-utility analysis. (Co-morbidities are taken into account but only to the extent that they raise risk of adverse outcomes or reduce likelihood of recovery.)
Calibrating QALYs to Respect Equality

Calibration to maximum health prospect as here proposed is therefore achieved simply by not making any explicit adjustments for individuals with co-morbidities or disabilities notwithstanding their lower achievable health prospect. For example, where a treatment offers a full cure for a disease, it is normally implicitly assumed that all individuals who benefit from the treatment will be returned to full health, notwithstanding that some individuals who benefit from the treatment will have other conditions or disabilities that affect their health (so that ‘full health’ is effectively understood as ‘maximum health prospect’).

Hence, our proposal can be seen primarily as a principled justification of current practice. It implies that it would not only be a breach of clinical ethics to discriminate against the irremediably blind person in authorizing a risk-mitigating treatment, it would also be a mistake from a cost-effectiveness perspective, even if it were only marginally cost-effective.

The implications of the methodology are less straightforward for treatments whose principal beneficiaries are those who have fallen victim to a progressive disease that limits maximal psychophysical function, and for the elderly. The measured QALY impact, and hence the assessed value for money, of interventions to slow illness progression is systematically reduced if potential health gain is assessed relative to a general population conception of full health, as is standard practice.

Whether this is discriminatory depends upon the perspective employed: whether we seek to maximize outcomes for the currently-healthy at-risk population (as in a Dworkin insurance model) or whether those currently suffering from such conditions should be considered part of the population for whom the basket of services covered by health services is being chosen.

For example, suppose we do take the view that it is the valuation of diabetics rather than those merely at risk of diabetes that should be authoritative in assessing an intervention to reduce risk of adverse sequelae. In that case, the importance of avoiding a particular adverse outcome – like blindness – is assessed for its estimated particular impact proportioned to the realistic health aspirations of those with diabetes (rather than relative to some standard characterization of full health). Standard assignments of HRQoL level to health prospect, i.e. those based upon responses from surveys where respondents assume no fixed limit to achievable psychophysical function, would have

to be recalibrated for application to a population with maximum
health prospect involving irremediable disability. Thus if a treatment is
under consideration for a population whose maximum health prospect
would be assigned a HRQoL level of m for those without irremediable
disability, then all other prospects with a standard assignment of 1, less
than m, should be scaled up to l/m. (For example, a prospect standardly
assigned a level of 0.2 would be rated at 0.5 for a population whose
maximum possible health prospect would be rated for the population
without irremediable disability at only 0.4.)

The same dilemma regarding perspective arises with much more
general application in considering treatments that particularly benefit
the elderly. Health economics valuations of treatments are sometimes
reduced on account of the lower average self-reported health-related
quality of life scores of the elderly, notwithstanding that these lower
scores are unavoidable. (This discount is applied to the fewer life years
that are inevitably attributable to treatments benefiting older people.)

The methodology proposed could be used to avoid such diminished
valuation of additional years and increased health relative to maximum
health prospect for the elderly as for all those with incapacities
unrelated to the condition targeted by the intervention being appraised.

Whether it is appropriate thus to calibrate elder life years to
best attainable health prospect depends upon whether we can adopt
the Dworkin insurance model for this group. To do so would allow
consideration of disabilities associated with age to be weighed by each
potential beneficiary of health services ‘looking from youth down the
course of their lives and trying to decide what risks were worth running
in return for not running other kinds of risks’.

It is at least arguable that older citizens would hope that their
diminished average level of health would not be allowed to compromise
valuation of the prospective health treatments from which they might
benefit, and that they might deny the relevance of the insurance
argument. Evidently the timing of the selection of treatments for the
insurance basket relative to the life course of members of society, and
the temporal scope of decision, are relevant when determining which
infirmities are to be considered known. Perhaps the constituency whose
welfare is at issue should be defined by a planning horizon for the
decisions to be made. If any who might benefit are currently suffering,
then their perspective, alongside those merely at risk, would at least
be relevant.

It is a subtle point in political economy and in social contract theory
requiring debate and further work whether the participants in the
contract are to be considered all to be in their youth as Dworkin
seems to suggest; or whether we envisage some broader constituency,
and if so how to integrate competing perspectives regarding the same
Calibrating QALYs to Respect Equality

intervention. However, what is clear, or so I have argued, is that respecting equality of human worth requires calibration of welfare impacts for all those for whom social decisions are being taken to their best attainable welfare prospect however limited that may be; to do otherwise is to denigrate their intrinsic life-value.18
donald.franklin@dh.gsi.gov.uk

Queries

Page 1, line 26: ‘prospect’ OK, or ‘prospects’ (plural)?
Good point, but better change the first part of the sentence to the singular, shifting the apostrophe in "individuals" and adding "each" before "individual's", to read: "... gains to each individual’s own maxim al potential health prospect."

Page 10, line 355: ’(1-HRQoL_i)’: is the dash after ‘1’ a minus sign?
Yes it is supposed to be a minus sign.

Page 11, line 434: ‘beneficiary’s ...’: ‘a beneficiary’s’?
Yes, that would be better, thank you.

Page 13, line 509: ‘average citizens’ assessment’: OK as plural possessive, or should it be ‘citizen’s’ (singular possessive)?
The problem is actually that it is the assessment that is average not the citizen! So better to invert "average" and "citizens", and drop the "an": Calibrating against [any] citizens’ average assessment

Page 17, line 671: [Your instruction:] delete "To"; insert "From".
My comment: which 'To'? There are two 'to's in this line and both seem correct. Do you intend another line?
This edit should have been marked to line 679 – sorry! It should read:
[From those with]
680 this condition, suppose we elicit

18 I am grateful to the following for perceptive comments: Eric Nord, Tongtong Qian, the Editor and an anonymous reviewer for Utilitas, members respectively of the Golders Green Kreis, of the Health Economics Study Group of the United Kingdom, and of the London Economics Journal Club of the UK Department of Health. Views expressed are not necessarily those of the Department of Health.