



Report on NICE Citizens Council meeting

Quality Adjusted Life Years (QALYs) and the severity of illness

31 January – 2 February 2008

Contents

| | |
|---|----|
| Foreword | 3 |
| What NICE asked us to consider | 4 |
| The conclusions we reached | 4 |
| How we worked | 4 |
| What we heard and what we did | 5 |
| Appendix 1 - Tracking questionnaire results | 15 |
| Appendix 2 - NICE's question and briefing paper | 17 |
| Appendix 3 - The agenda | 29 |
| Appendix 4 - Speaker biographies | 32 |
| Appendix 5 - The Council members | 35 |

Foreword

The National Institute for Health and Clinical Excellence (NICE) exists to advise clinical and public health professionals on promoting good health and preventing and treating ill health. The Institute and its advisory bodies base their conclusions on the best available evidence. In doing so, however, they also have to make scientific and social value judgments.

The Institute and its advisory bodies are well qualified to make scientific judgments but have no special legitimacy to impose their own social values on the National Health Service (NHS) and its patients. These, NICE believe, should broadly reflect the values of the population who both use the service (as patients) and who ultimately provide it (as taxpayers). NICE therefore established a Citizens Council, in 2002, to help provide advice about the social values that should underpin the Institute's guidance. Their views are incorporated into a guideline for NICE's advisory bodies: *Social Value Judgements*.

The members of the Council reflect the demography of the English and Welsh populations. They serve for three years with one third retiring annually. They do not represent any particular section or group in society; rather they are a cross-section of the population with their own individual experiences, attitudes, preferences and beliefs. The Council meets twice a year.

At the January 2008 meeting of the Council, the key question for the Council was as follows:

Should NICE and its advisory bodies take into account the severity of a disease when making decisions?

If yes, should the advisory committees:

- take severity "into consideration" alongside the cost and clinical effectiveness evidence;
- or should severity be included in the calculation of the QALY?

The Council's report will be available for public comment before it is presented to the Institute's Board in September 2008.

Once again, the Institute is extremely grateful to the Council for its continuing help in developing NICE's social values.

Professor Sir Michael Rawlins

Chairman

What NICE asked us to consider

Should NICE and its advisory bodies take into account the severity of a disease when making decisions?

If yes, should the advisory committees:

- take severity “into consideration” alongside the cost and clinical effectiveness evidence;
- or should severity be included in the calculation of the QALY?

The conclusions we reached

The Citizens’ Council concluded, by 24 to 2, that NICE and its advisory bodies should indeed take the severity of a disease into account when making decisions. Among the 24 of us who took this view there was unanimity that rather than do so by including severity in the calculation of the QALY, it should be taken “into consideration” alongside the cost and clinical effectiveness evidence.

We reached this conclusion mainly because the process of QALY calculation already takes some account of severity, and because any changes intended to weight QALYs further in this respect will inevitably make them more complicated and harder to understand, and may also distort the model. This in turn could lessen their transparency, thereby making any attempt to understand a committee’s decision correspondingly more difficult.

The alternative - taking severity “into consideration” - would give appraisal committees more flexibility. We would not wish to see a mathematical or other formulaic approach to this task; such a step might simply recreate the rigidity of the QALY component of the decision. One possible course of action would be to add something about severity to NICE’s statement on social value judgements.

Transparency is vital to the acceptance of NICE decisions – not only in the way that committees reach them, but in how they are subsequently reported and explained. We must all be able to understand a committee’s reasoning.

We feel there is a problem with the EQ-5D questionnaire which we think is too blunt to capture all the factors relevant to the definition of a good or bad quality of life. We would prefer an approach that incorporates more of the social as well as the medical model of health and disease, and might - though we appreciate this could be difficult - take more account of the views of those who have first hand experience of the circumstances being rated.

We are not calling for the questionnaire or the QALY to be abandoned; rather we are suggesting that, in the light of experience so far, it is time they were subjected to a thoughtful and penetrating review.

How we worked

The Citizens’ Council met on Jan 31 and Feb 1 at the NICE headquarters in London, and then for a further half day at a nearby hotel. Twenty six of our members were able to attend the meeting. As usual we heard presentations from NICE staff and from outside experts, we questioned them, and we discussed what we had heard – collectively and in smaller groups. At intervals we also filled in short tracking questionnaires to assess how our thoughts were developing.

This particular meeting of the Citizens' Council was slightly unusual - indeed privileged - in that we had an opportunity to hear some of the preliminary findings of research commissioned by NICE to find out what the general public thinks about QALYs in relation to the severity of disease.

What we heard and what we did

In his brief welcoming address, the NICE chief executive Andrew Dillon described this topic as "the big one"! As he pointed out, the QALY lies at the heart of every decision NICE has to make, and remains the most consistently controversial aspect of its work. Although QALYs offer the best way of calculating value for money, many of those involved in making NICE's decisions - and in living with them - have at least a few reservations about the system. It is under constant challenge, we were told. So even if we were eventually to decide that modifications or improvements are unnecessary or undesirable, our verdict would at least permit NICE to feel more confident in defending its own position. In the event we did recommend changes.

Professor Peter Littlejohns, NICE's clinical and public health director, introduced the question that we'd been set with the promise that answering it was going to be hard work. He could have added - though he didn't! - that we would be in no position to complain about this. In past meetings a number of Citizens' Council members have expressed doubts and reservations about the QALY system - and the Council has even suggested that we might at some point formally discuss it. Our wish was being granted.

At this point one of us asked Peter Littlejohns to define the meaning of severity. He put the issue back to us, pointing out that this was something we would have to discuss. NICE itself has no agreed definition, and different people have different ideas. The initial vote we were about to take suddenly looked slightly less clear cut than some of us might have expected.

In response to the first part of the question we had been set ("Should NICE and its advisory bodies take into account the severity of a disease when making decisions?") 20 of us responded "yes". Three voted "no", and there were three "don't knows".

Those who had voted "yes" then had to choose between the two options we had been given: include severity in the QALY calculation; or instead have appraisal committees take severity "into consideration" alongside the other evidence. Only six of us, at this stage, felt prepared to commit to one or the other option - and the vote was split three and three.

At this point it became clear that not all of us were entirely clear about the difference between the two options. Peter Littlejohns explained that including severity in the calculation - weighting the QALY, in other words - was a more formal approach to the matter. Taking it "into consideration" was less formal and more flexible. In the light of this further insight we voted again. This time 11 of us felt able to take a position: five of us backed the modified QALY; six the alternative position. Still while more of us felt able to form an early opinion, our views on what to do were still split.

Our own quality of life

Dr Joanne Lord, a technical adviser to NICE, gave us the opportunity to experience a quality of life (QoL) assessment first hand. She gave each of us a copy of the EQ-5D [included in Appendix 2: The Question], the health questionnaire that underpins many of NICE's assessments. Each of us, anonymously, filled one in. Doing this provoked a flurry of doubts, including unease at giving numerical values to such complicated

issues as life quality, and uncertainty that all our answers would have been the same yesterday as they were today.

One of our number, a wheelchair user, felt that some of the questions - on mobility, for example - made no sense to someone in her circumstances. She, and indeed the Council as a whole, were later dismayed to discover that she had achieved a negative QoL score – meaning, to put it brutally, she would be better off dead. As she was there among us, manifestly cheerful and not seeking her own demise, her score was to say the least anomalous.

Joanne Lord countered these criticisms by pointing out that, for NICE's purposes, the aim is to acquire data relevant to whole societies rather than to single individuals. On the questions of day-to-day variations she said that, on average, one person's bad day will be offset by another's good one. And because the questionnaire relies on broad-brush generalisations, it won't satisfactorily describe every single person who uses it – not least, in this instance, our own wheelchair-using member. Even so, this incident did nothing to boost sympathy for the QALY among those Council members who were already doubtful about it.

Our first formal presentation was by Ann Bowling, professor of health services research at University College, London. The idea of using a QoL score, she said, is to broaden the measurement of health beyond more conventional yardsticks such as biochemical measurements, and into territory more directly relevant to patients themselves: broadly, their well-being, comfort and satisfaction. The EQ-5D and others questionnaires like it are a way of reflecting the subjective impact of health, illness and its treatment. But none can capture the full complexity of individual lives. Because they're subjective, QoL scores can change without the person being assessed having experienced any alteration in his or her actual circumstances.

The QALY, Professor Bowling conceded, can be "ageist": a reflection of the public attitudes on which, by definition, it is based. It is possible to tailor questionnaires to particular illnesses or even to individuals. This is a developing field in which new ideas are still to be explored. Formally incorporating severity into the QALY system, she speculated, might reduce age bias.

Our subsequent questioning and discussion raised a number of issues including the expense of individualised questionnaires, the extent to which our attitudes to matters such as youth and death are culturally influenced, age-related changes in QoL, and whether the EQ-5D could be improved simply by adding more dimensions.

A lesson on QALYs

Joanne Lord then returned to give us a tutorial on the nature of the QALY, how it has been derived, and the necessity to have it or something like it when demand for healthcare outstrips resources, and questions of cost-effectiveness cannot be avoided. The QALY makes it possible to compare different treatments for different diseases, and judges their impact over time. This is because a QALY is a QALY, irrespective of the patient, the disease, or its severity. Moreover quality of life can be traded against time: one QALY equals one year of perfect health, or two years of 50 per cent perfect health, or four years of 25 per cent perfect health, and so on.

The original assessment of life quality was based on surveys of public attitudes to various health conditions. For QALY and NICE purposes, these public attitudes are what count. This is because, through the NHS, we choose to pool our resources to pay for health care, and so we are all entitled to have a say in how those resources

should be allocated, including to the management of conditions which we do not (and may never) ourselves suffer. Using clinical data and/or estimates, and knowing the cost of an intervention, NICE is able to calculate the cost per QALY of offering this or that intervention for a particular condition. If the cost per QALY falls below a pre-determined threshold (currently £20,000) it is judged acceptable. If above £20,000 NICE will need persuasion to give it the go-ahead. If above £30,000 the verdict will normally be a thumbs-down – though NICE committees do have a degree of discretion over the rigidity of the threshold. They can take account of NICE's statement on social value judgements – though this does not, as yet, make specific mention of severity.

In the subsequent discussion one member pointed out that people enduring a particular condition can become accustomed to it. Their measured QoL may not change, but their own perception of it can. It was also pointed out that data on the length of life to be expected following a still newish intervention may not be known. But NICE can't wait for what could take years while the necessary research is carried out; it may have to make a decision on whatever data are currently available. Most of NICE's decisions - as Council members are well aware – do involve the exercise of judgment as well as the mere application of a simple formula.

We began to think about the implications of trying to build a severity weighting into the QALY. Joanne Lord reminded us that, at least to a degree, the QALY already takes account of severity through the public's estimations of the relative undesirability of certain states of health. A bigger health gain is needed to move upward at the top of the scale than at the bottom; going from 0.2 to 0.3 is already effectively weighted as worth more than going from 0.8 to 0.9. So further weighting could invite the accusation of double-counting. Against that it could also be argued that the extent to which severity is currently taken into account is insufficient. Maybe the QALY *does* need extra weighting. Someone else pointed out that you could make a case in favour of weighting for several other factors, and wondered where it might end. Another comment was that you have to be wary of making anything (including the QALY) more complicated than it need be. There is the consequent danger that weighting QALYs would reduce their transparency. The more complicated the QALY gets, the less clear it may become what is actually being measured. And the less that people understand the QALY, the more likely they are to challenge it.

We also wondered about the population survey that underlay the EQ-5D questionnaire. It was done some years ago, and may not have been entirely representative. Maybe it needs to be repeated? (We learned that it is indeed under review.)

The problems with QALYs

Our next session was a discussion with Joanne Lord, Sarah Garner who is NICE's associate director for R&D, and Susan Bennett, a patient representative on one of NICE's committees and a wheelchair user. Susan told us that her own QoL score is only just above zero: an anomaly similar to that which our wheelchair-using member had just discovered. She talked of the definitions of disability in the Disability Discrimination Act and also of the distinction between the social as opposed to the medical model of her condition. How many of the 3,000 people whose views had led to the EQ-5D health state evaluations had been disabled, she wondered? And does health-related QoL put too much emphasis on mobility impairments? As she pointed out, the scale represents a view of life and its quality from the perspective of people whose daily experience is very different from hers. Incontinence, she suggested, can be a far bigger problem to live with than loss of mobility. But how many people who had experienced neither would actually know that?

Issues raised in our discussion with the three panellists included: the extent to which some of the perceived problems are less to do with QALY as a concept than with the way in which the information on which it's based is collected; and whether changing the QALY would do anything to change society's view of disability. One member pointed out that dealing with one of Susan Bennett's complaints - that her QoL score created an unrealistically pessimistic view of her life - could act against her own interests. A higher rating might actually reduce the likelihood that treatments of relevance to her would be judged cost-effective.

We asked Susan how she herself would answer NICE's question on whether severity should be taken into account when making decisions. It depends what you means by severity, she responded – which of course was one of the problems that we had been grappling with! Either way, she didn't think it could easily be built into the QALY measurement.

Our final presentation was given by Dr Marcia Kelson, NICE's associate director for patient and public involvement. She re-iterated the distinction between patients and the general public, and confirmed that their respective views are not always the same. Patient organisations apparently claim that NICE places greater weight on survival than on QoL. In metastatic breast cancer, for example, a treatment may not lengthen life but might allow the patient to spend more time with friends and family: a valuable gain to which the QALY system would be insufficiently responsive. She also used her own mother's motor neurone disease to illustrate the subjectivity of assessments. Her mother's progressive impairment left her with a near zero QoL score even when she herself had continued to value her own life.

Patient groups also find the QALY system difficult to understand; they say it makes decision-making too opaque, and emphasise that it doesn't always capture the patients' own perception of their quality of life. But wouldn't trying to incorporate severity into the QALY calculation make it even more opaque?

There is also the problem of individual difference. A heart attack can affect the lifestyles of a couch potato and a fitness freak quite differently.

The topics covered in our subsequent discussion included: the hazards of reducing everything to a mathematical formula; the manner in which small improvements in health can sometimes lead to dramatic changes in life quality; whether the tension between the public and patient views of QoL can be in any way beneficial; the difficulty of rating the severity of totally different kinds of symptoms; and the fact that in averaging the views of large numbers of people you inevitably lose the distinctive views of minorities.

At this point we broke into small groups to discuss what, at that stage, we were clear about, and what we were not clear about. After 20 minutes or so we reconvened to report back.

The "not clear" category featured (inevitably) the definition of severity, how one can weight something that can't easily be defined, how somebody with a good quality of life can score a minus on the QoL scale, the benefit to NICE of taking severity into account, the implications of not making any change to the system, whose view should count most when assessing the impact of an illness (ie patients or public), how it's possible cope with the fact that the same illness can impact on different people in different ways, and how severity could satisfactorily be built into the QALY.

In the “clear” category were: that the EQ-5D is a blunt instrument, one not very good at measuring the impact – incontinence and sensory impairments, for example – or certain illnesses; that up to a point, severity does figure in the QALY measurement; that without “fiddling” with the QALY there are already precedents for stretching NICE’s cost-effectiveness threshold; and that while the QALY is widely accepted and generally understood, it does not work as a satisfactory decision-making tool in every case. Each of these propositions, of course, had its dissenters.

In a closing discussion with Professor Littlejohns we wondered why NICE had chosen this particularly moment to discuss the issue of severity. He conceded that questions about it had been raised almost from the outset. But to change the system NICE had so far been using would be quite an undertaking, and if this is the case the sooner NICE start planning it the better. Research had been commissioned and its results were now becoming available. Indeed we were to be given as sight of the preliminary findings on the second day of our meeting. With the completion of the first tracking questionnaire, the day came to an end.

Day two: weighting for insight

We began the second day by discussing how far we had advanced in trying tease out the nature of severity, and deciding if it could be weighted. For some people severity implies a threat to life; to others it’s more about the extent to which they can no longer carry on living life as normal. Perhaps the fact that severity can be defined in many different ways argues against trying to build it firmly into the QALY. If dealing with severity is a matter of instinct, perhaps it would be better to leave it to NICE committees to take it “into consideration”.

Our next presentation was intended to help us get to grips with this issue. It was given by John Cairns, professor of health economics at the London School of Hygiene and Tropical Medicine. As he reminded us, the existing convention is that all QALYs are equal. So he set out to question whether the size of any one individual health gain should be of consequence, and whether it matters who is getting the QALYs.

Rather than “severity” he favours the term “poorer health” – an alternative that many of us found helpful in trying to clarify our thoughts. Should NICE value health gains going to people in poorer health more highly than those going to people in better health? EQ-5D, he reminded us, does already take account of poor health; so would such a valuation be double counting? Not necessarily. There’s a distinction to be made between the QALY as a measure of health gain, and the value that society places on QALY gains. Some changes may be - almost certainly are - valued by us more highly than others.

He then went on to ask if is feasible to weight QALYs. Doing so would increase the quantity and complexity of evidence that had to be collected. Would this be desirable?

Professor Cairns then spoke about alternatives to directly weighting the QALY. NICE appraisal committees, he reminded us, are not without flexibility; they already have the discretion to take account of clinical need. They could be mandated to give more weight to social value judgements, and proportionately less to cost-effectiveness. Of course it could be argued that this approach is a bit loose, and lacking in transparency. But it would sidestep the difficult technical issues of developing a weighted QALY system. Indeed, when we pressed him to offer his personal preference, he said he’d leave the QALY intact and look for ways of being more

systematic about arguments that went beyond cost-effectiveness: issues of the kind encompassed in NICE's social value judgements category.

Subsequent discussions with Professor Cairns further underlined the complexity that would follow from the introduction of a weighted QALY, the inevitable fact that weighting (in the absence of extra resources) creates losers as well as winners, and how palliative care and prevention might be affected. One of our members even speculated that the long term effects of shifting health gains in favour of those in poorer health could adversely affect the well-being of the working and tax-paying population and begin to undermine the financial basis of the NHS. We also wondered what guarantee there could be that NICE committees, even if instructed to do so, really would devote more attention to social value judgements in reaching decisions. On this last point Professor Cairns suggested that firm reminders from the chair would be important. And having the relevant principles written down in black and white should make a difference.

The next two speakers gave us the initial results of research commissioned by NICE to find out how the public believes that the severity of a condition should influence the allocation of resources. Professor Paul Dolan of Imperial College, London and Richard Edlin of Leeds University kicked off with their study of several possible influences including not only the severity of the illness but the age of those experiencing it, whether it was their "own fault", and how common it was. Their bottom line conclusion was that people do care about the severity of the illness - but not overwhelmingly. They also suggested that it's difficult if not impossible to study severity in isolation from other factors; they are mutually influential.

Newcastle University health economist Dr Rachel Baker, who'd also been commissioned to do research by NICE, then presented the findings of her survey. She had used two different methods which had reached slightly differing conclusions about the importance of severity: an indication of the difficulty of research in this field. In so far as severity does matter, it seems that we value health gains most in those aged 20 to 40 years, and most when the people concerned have a QoL of 0.2 – 0.4. Her tentative conclusion (followed, it should be said by a question mark) was that trying to weight the QALY at this stage of our knowledge and experience would be premature.

What is severity?

After we had filled in a second tracking questionnaire, all three researchers were joined by two more experts for a panel discussion on QALYs and severity. One of the two new witnesses was Rachel Rowson of the charity Cancerbackup. Describing the QALY as a "blunt instrument" she suggested that adding QALY modifiers or weightings might make an already wobbly base wobblier still. Peter Mansell of the National Patient Safety Agency argued that the weakness of the QALY is that it masks individual differences. The same intervention, he told us, can have a different impact on different individuals. Small gains certainly have a disproportionate impact on people with severe problems; but he didn't believe that a severity correction should be applied to the QALY – which he saw as an essentially political instrument masquerading as scientific one.

The panel discussion revealed that sympathy for the QALY was decidedly limited - but little consensus on what, if anything, should replace it. Paul Dolan was involved in some of the original work that led to the development of QoL scales and the QALY, and is therefore well-placed to appreciate its weaknesses as well as its strengths. The problem, he pointed out, is that the EQ-5D measures what people *imagine* the

experience of various health conditions to be like. Clearly most of us never have experienced most of them, and never will.

The discussion ranged back and forth, but with ever-lessening enthusiasm for attempts to weight the QALY in an effort to make it perform more “fairly”. There might even be the danger that a weighted QALY would leave NICE committees feeling that they need pay no particular attention to severity because the issue had satisfactorily dealt with.

More of the discussion began to focus on whether NICE should be paying less attention to QALY evidence, and more to “intuitive” evidence. Although this might create difficulties for appraisal committees when they had to justify their decisions, maybe it was time for NICE to start behaving more - as someone put it - as if it had a heart. Or, to express it less emotively, to pay more attention to the social value judgement part of the remit.

This prompted another train of thought: if we chose the “take into consideration” option of the two that NICE had given us, could we realistically suggest taking more account of severity without also taking more account of other factors that contribute to social value judgments? The panel were inclined to agree that disentangling any one factor from all the others is difficult. More thought and more research, said one of them. Indeed; but we had to make up our minds by the following day!

Someone else suggested introducing the term “special circumstances” instead of severity. This has the merit of covering severity and other things too. But on reflection it became apparent that everyone making a case to a NICE appraisal committee would find a special circumstance of some kind. We dropped the idea.

We returned yet again to NICE’s need for transparency when explaining and justifying its decisions. Would committees really have such an impossible task if obliged to use words rather than QALY numbers to account for their conclusions? Would they really find it such a problem to explain themselves? Professor Cairns was not entirely reassuring. By the time a committee has reached what is often a difficult decision, he said, a lot of its collective energy has been used up. There may not be much left over for then trying to explain and justify the decision.

Following a tea break we had one further attempt at brainstorming on the nature of severity, and trying to decide how it might be incorporated into the thinking of NICE committees. We broke into small groups to discuss it, and to draw up lists of what severity is, and what it is not. Our formulations were many and various. Among the suggestions in the “what it is” category were short life expectancy, risk of death, extent of suffering and the presence of intolerable symptoms, illness on top of an altered impaired state of health, and illness of long duration. We had more difficulty saying what severity isn’t. Our list included: what the “well” public, in surveys, imagine it to be; what the media claim it be; what companies claim it to be for the purposes of getting an intervention accepted by NICE; whatever a patients’ group, without evidence, says it is; and the fact, in isolation, that it happens to afflict disadvantaged groups whether defined by age, gender, ethnicity or class.

We finished the day by completing a third tracking questionnaire.

Reaching a decision

Notwithstanding the plethora of sometimes contradictory ideas we had mulled over the previous afternoon, we began our last day confident that our thoughts were coming together and that we could reach a conclusion by the end of the morning.

One of the issues with which we'd wrestled was a working definition of severity. NICE itself, we'd been told, has no agreed definition – and one of our members suggested that perhaps there really is no need for it. On a vote, 23 out of the 26 of us felt that even if NICE did adopt a definition it shouldn't be set in stone, but subject to periodic revision. In the end we moved away from precise definitions and shifted back to the exercise we'd attempted the previous afternoon: listing some of the factors that NICE might include in the term "severity".

Before doing so we tried to pin down a few other issues. For example, is it time for NICE to review the issue of the QALY more generally? On another vote, 18 said "yes", 6 said "no", and the remaining two of us were undecided.

Someone suggested that instead of completely damning the QALY - as a few of us were inclined to do - we should perhaps be asking NICE to refine it. Given that all sorts of people with quite different interests in NICE and its work are not entirely happy with QALYs, someone else expressed bewilderment that this hadn't already been done.

We voted on the future of the QALY. Six of us favoured keeping the QALY as it is; two wanted to throw it out entirely; 19 of us - a substantial majority - favoured keeping the QALY but reviewing and perhaps refining it.

Another view expressed was that the problem lies with the questionnaire, the EQ-5D. When we voted on the question "Is it the EQ-5D that needs fixing?", 23 responded "yes". One of the three abstentions was prompted by the fears that trying to refine it might simply create further complications.

Overnight, a Council member had been reading about the actual use of QALYs made by one of NICE's technology appraisal committees. In the particular issue before that committee, any reliance on QALYs had seemed to our member rather suspect. But, he added, the committee hadn't relied on QALYs alone. Well aware of the weakness of relying too much on QALY evidence they had gone on to exercise their own judgement. While not ignoring the QALY score, they had nonetheless delivered a judgment that our member described as "compassionate".

Someone else took this thought one step further. Maybe we should be leaving more of the weighting for severity to appraisal committee members, and not always have them rely so much on the QALY calculation. Fears had already been expressed about the loss of transparency that might go with decisions relying more on human judgement than the application of a mathematical formula. Against that it was pointed out that NICE intends to make its committee meetings open to the public so that their reasoning can always be scrutinised.

This seemed to be the appropriate moment for a final vote on the questions at the heart of our discussion: the questions that NICE had set us. In response to the first one ("Should NICE and its advisory bodies take account of the severity of the disease when making decisions"), 23 out of 26 voted "yes". We asked the three dissenters to outline their reasoning. Among their objections were:

- that severity is already taken into account in the QALY calculation;
- that severity is subjective;
- that inequality can only be addressed satisfactorily if doing so doesn't generate further inequality;
- and that prioritising health care on the basis of QoL measurements could actually disadvantage the disabled.

They also suggested that the debate over the EQ-5D is about its wording, and that adding extra dimensions wouldn't necessarily improve it.

We discussed these views but reached no clear accommodation; the majority and the minority agreed to differ.

Returning to the NICE questions, the majority - the 23 of us who did favour NICE taking specific account of severity - then had to choose between the two options we had been given: to include severity in the calculation of the QALY; or have it taken "into consideration" alongside cost and clinical effectiveness evidence. There was one abstention; otherwise all votes went to the latter course of action. Most of us do not favour tackling the issue of severity by trying to modify the QALY.

Things to tell NICE

In addition to answering the question that NICE had set us, we have a number of comments and suggestions to make. When NICE is making appraisals, we would like to see more emphasis on the social value judgment element of the decision-making process and, within it, attention paid to the issue of severity.

We are still concerned about transparency. We feel that committees should make a determined effort to be explicit about how and why they make up their minds, and what weight they give to factors such as severity. We think they should also ensure that their reports explain the reasons for their decisions so that anyone not present to witness them being made could nonetheless grasp the thinking that shaped them. We acknowledge that this is not easy to do, but urge committees to try. Going beyond QALY calculations might make it easier to deal with certain issues – palliative care is in this category – which may face an uncertain fate at the hands of an orthodox QALY analysis.

In addition to producing decisions which, we hope, would be fairer and more just, we think appraisal committees might also seem less mechanistic and more humane. This would be good for the public image of NICE – which is sometimes seen as too cold-hearted, or concerned only with the economics of decision-making.

We recognise that, in this context, severity is a difficult concept. One way of viewing it is as "how far away you are from your final destination: perfect health." Another concept we found helpful was the "health thermometer" in which poor health places you at the low (cold) end and good health at the top (warm) end. This is really another way of expressing the QoL scale, with zero (dead) at the bottom and one (perfect health) at the top. Several times during the meeting we expressed interest in the idea of weighting health gains that are achieved at the lower end of the scale, and so benefit those in an already poorer state of health. A move from, say, 0.2 to 0.3 would be given more weight than an equal numerical shift from 0.8 to 0.9.

One of our members was in favour of an approach to severity which recognised that the underprivileged are likely to be disproportionately affected by illness, and argued the case for a system of weighting in their favour as a group. But there was little support for the suggestion, most of the rest of us feeling this was both undesirable and unfeasible. As someone pointed out, weighting increases at the lower end of the QoL scale will in any case do proportionately more for the most disadvantaged groups.

NICE must recognise that, for some people or under some circumstances, length of life is by itself of value, whatever its quality. Surviving to see something important

happen, for example. Severity judgments are currently dominated by a medical model of disease. We would like to see NICE take more account of social factors.

Because we are keen to see NICE doing this, we wondered if committees might find their task easier if they could call upon a numerical scoring system. Having floated this idea, however, a good many of us began to doubt it. Like “trying to nail down jelly”, as someone commented. The criteria for a scoring system would, in any case, need to be set out in words. So words should be sufficient to describe what’s been decided and – most important – why.

We also tried to make an agreed list of what severity is and what it is not. Severity, we think, must take in account of:

- life expectancy;
- how far away you are from perfect health;
- your state of health prior to, during and after diagnosis and treatment;
- the nature of the treatment and its impact (side effects as well as benefits) on health;
- and health states that incur social stigma such as incontinence.

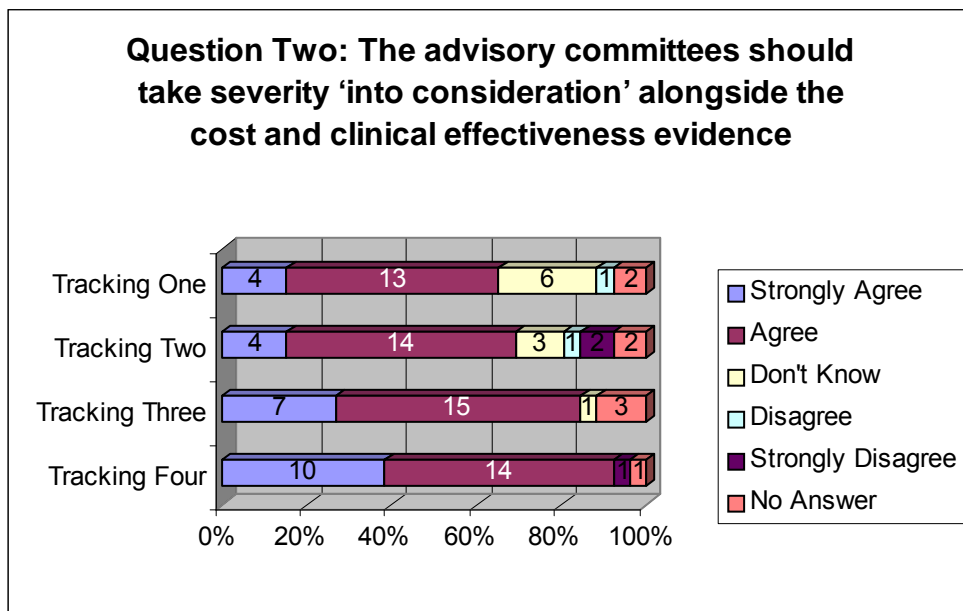
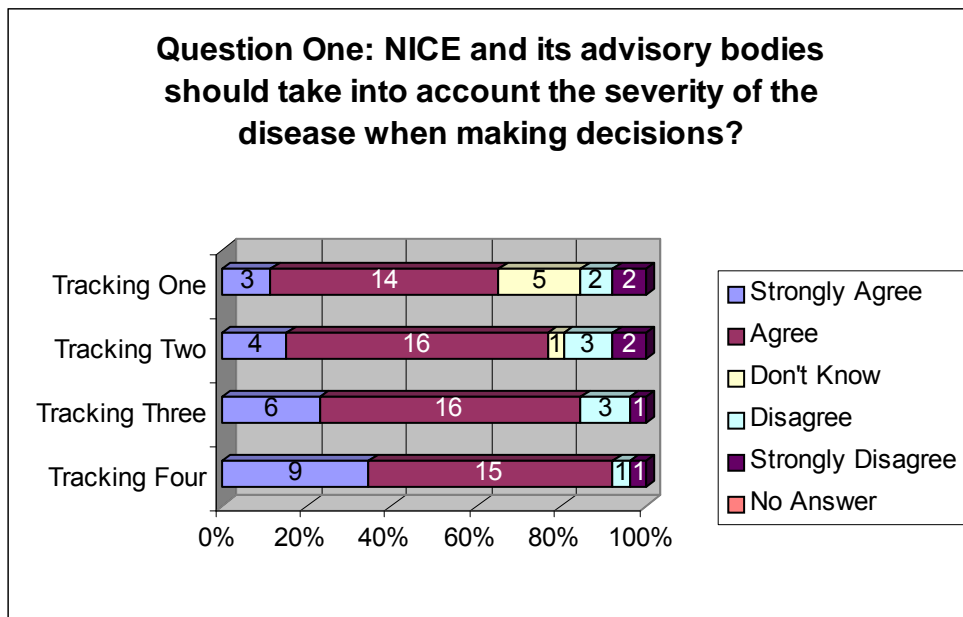
Severity, someone said, may be hard to define; but you generally know it when you see it. One member toyed with defining it as “A condition I really, *really* don’t want to have” – until it was pointed out that at various times and in various circumstances, that could apply to almost any disorder!

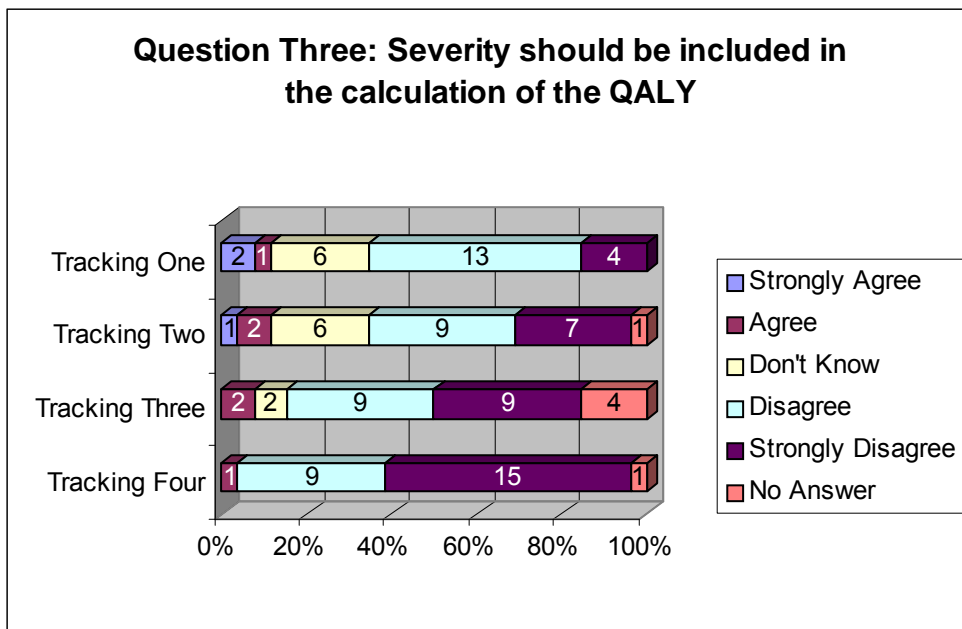
In the light of a discussion we’d had earlier, we also felt it important to add that severity is not about discrimination on social grounds, and is not a concept to be reserved for particular ethnic, gender or age groups in the absence of clear evidence of clinical cost-effectiveness and capacity to benefit.

A final vote on the question brought us to 24 in favour of NICE taking severity into account, and two against. Of the 24, all favoured taking severity into consideration alongside cost and clinical effectiveness evidence; none were in favour of modifying the QALY.

Not quite unanimity. But not too far away from it.

Appendix 1: Tracking questionnaire results





The results of the tracking questionnaires illustrate the progressive shift in our opinion took place during the course of the meeting. The response to question one shows how our view that NICE should take severity into account when making decisions received steadily increased backing, both in numbers and in the strength of the agreement.

The response to question three showed initial doubts about the value of tackling severity through the use of a modified QALY. As the meeting progressed, opinion hardened against the QALY; by then end there was overwhelming disagreement with this approach.

Movement on question two – again in numbers and in strength of feeling – was in the opposite direction. A majority of us, right from the outset, were inclined to think that advisory committees should tackle severity by taking it “into consideration” during their deliberations. As we learned more about the alternative approaches, this was the one that gathered progressively more favour.

NATIONAL INSTITUTE FOR HEALTH

AND CLINICAL EXCELLENCE

Citizens Council

Should NICE and its advisory bodies take into account the severity of the disease when making decisions?

If yes,

should the advisory committees take severity 'into consideration' alongside the cost and clinical effectiveness evidence

or

should severity be included in the calculation of the QALY?

Background

No healthcare system in the world is able to meet the entire clinical needs of the people it expects to help because budgets are finite. How decisions about clinical priorities are made varies from country to country but the need to do so is clear. There just isn't enough money to go round!

In Britain, until recently, these decisions were largely made behind closed doors. Although, in some circumstances, formal economic assessments were made these were rarely exposed (or explained) to the public. More often, decisions about how the NHS's money was used were based on other factors. These included assumptions about where (and how) additional investment might most appropriately be made, pressure from special interest groups, political lobbying and perceptions about public preferences.

The establishment of NICE, in 1999, was a 'sea change' for the NHS. For the first time a public body was to be charged – in effect – with making decisions about the availability of new and established treatments, and pathways of care, that formally took cost-effectiveness (or value-for-money) into account. Few healthcare systems had tried to do this before; and in those that did, such as in Australia, it was mainly limited to new drugs.

NICE's approach to decision making

In making its decisions on whether a treatment (e.g. a drug or a device) should be available to the NHS, NICE compares treatments by considering how much of an impact each treatment has on overall health and how much it costs. To do this, NICE uses an economic approach called 'cost-utility analysis' to compare healthcare interventions.

Cost-utility analysis uses a measuring instrument or 'yardstick' called the Quality Adjusted Life Year (QALY). The QALY approach takes into account that both 'quality of life (QoL)' and length of life are important, and allows treatments for different conditions to be directly compared. For more explanation of economic analysis and QALYs please see the accompanying document 'Health economics: the basics'.

The aim of the cost-utility approach used by NICE is to 'purchase' the greatest number of QALYs possible i.e. to maximize the amount of health gained for the money available. The limit that should be paid for each QALY, is termed the

'threshold'. This approach also assumes that the QALYs from one group of patients are identical to those from another (a QALY is a QALY is a QALY).

NICE however recognizes that society may have preferences for how this money and resulting health or welfare benefits should be distributed and that these preferences should be taken into account. This would mean that society may be prepared to sacrifice some of the health (QALYs) it could potentially gain in order to secure what it considers to be a 'fair' distribution of health.

Examples of such preferences that have already been discussed by the Citizens Council are the 'rule of rescue', where those lives in imminent danger are saved first and; the age of the individual; and using resources to target the most disadvantaged members of society. In all cases the Citizens Council concluded that some allowance should be made in some circumstances.

The focus of this Citizens Council meeting is severity of disease. Research indicates that people in general consider that greater emphasis should be placed on allocating resources to the most severely ill. The Citizens Council is therefore being asked whether NICE and its advisory bodies should take into account the severity of the disease when making decisions?

In the methods that NICE uses, there are two ways in which the severity of the underlying illness might be taken into account. The one currently used by NICE is that the Committees takes the issues 'into consideration' alongside the results of the analysis of clinical and cost-effectiveness. The second way would be to adjust, mathematically (termed 'weighting'), the QALY by a factor that takes into account the QoL from which patients were starting.

An example

To illustrate the issue of severity, consider the four health-care interventions A, B, C and D in table 1. The interventions could be drug treatments, surgery or rehabilitation programs in different patient groups for example individuals with heart disease, arthritis, asthma, and lower-back pain. The table shows the average QoL (described by the 'utility') of patients before and after treatment. The gain in QoL ('utility gain') has also been calculated by subtracting the before treatment value from that after treatment. For more

explanation of utilities please see the accompanying document ‘Health economics: the basics.’

Table 1: Comparison of different treatment programs

| Intervention | Utility 0 is dead and 1 is perfect health | | Utility Gain (After – before) |
|--------------|--|-----------------|----------------------------------|
| | Before treatment | After treatment | |
| A | 0.4 | 0.5 | 0.10 |
| B | 0.8 | 0.9 | 0.10 |
| C | 0.85 | 1.00 | 0.15 |
| D | 0.6 | 0.80 | 0.20 |

According to the existing methodology used by NICE, program D would be considered the most valuable because it gave the biggest utility gain (0.2). C is the second most valuable (0.15) and A and B share third place (0.10).

If you consider the arguments on severity, a different ranking may however apply whereby A is preferred to B or C because the individuals have greater severity of disease before treatment as indicated by the lower utility score. For the more severely ill individual treated with intervention A the utility increase from 0.4 to 0.5 represents a 25% increase in their total QoL $((0.1 \div 0.4) \times 100)$. For the less severely ill individual treated with intervention B it’s only a 12.5% increase in their total QoL because they already have a utility of 0.8 $((0.1 \div 0.8) \times 100)$.

The debate

There is controversy over whether severity should be taken into account and whether the QALY should be mathematically adjusted or ‘weighted’. This debate was summarized by the late Professor Alan Williams, a renowned health economist, as follows: ‘There is a danger that such weights become arbitrary and capricious and come to be used to fudge outcomes in ways that would not be acceptable if their basis was exposed.’

Furthermore, there is no agreement on whether disease severity should mean ill-health at the point of time the patient presents for treatment or ill-health over a life

time. There is also no agreement as to how the weights should be obtained or how they should be used. If comparisons are made across diseases where different methods have been used to calculate the numbers that are used to adjust the QALY, this may result in incorrect priority settings.

Alan Williams recommended having an 'underlying (or over-arching) general principle enunciated, which can be confronted with evidence so that its various implications can be explored in a quantitative way.'

The meeting

The Citizens Council is being asked to consider the overall principle; whether it might be more appropriate for NICE to favour interventions that treat patients with severe disease rather than those with less serious disease on the grounds that they are suffering more and are less able to function in everyday life.

The format for this meeting will be slightly different to previous ones. In order to understand what the general public think about the issue, NICE commissioned 2 research groups to undertake more extensive studies. Professor Paul Dolan and Professor Cam Donaldson are nearing completion of their research and they are now ready to present their results. They will present their findings to the Citizens Council who will be asked to consider the implications of the research for the decisions made by NICE and how it assesses QoL. You will also hear from experts who consider that there are limitations to this type of research and who may hold different views.

Sarah Garner

Peter Littlejohns

Michael Rawlins

22nd January 2008

Health Economics: the basics

Methods of health economic analysis

When trying to assess whether one treatment, or pathway of care, offers better 'value-for-money' health economists have developed a number of different approaches. Three of these are:

- cost–minimisation analysis
- cost–effectiveness analysis
- cost–utility analysis.

Cost–minimisation analysis: This is the term used to compare two (or more) treatments that produce the same end-result. You choose the one that costs the least!

Cost–effectiveness analysis: This is the cost per 'natural unit' of health improvement (or 'health gain' as health economists like to put it). Some simple examples are shown in table 1 below:

Table 1: Examples of Cost–effectiveness analysis

| Condition-treatment | Cost–effectiveness measure |
|---|---|
| Short stature in children treated with growth hormone | Cost (£) per increase in unit of height (cms) |
| High blood pressure treated with a blood pressure lowering drug | Cost (£) per reduction in units of blood pressure (mm mercury) |
| High blood cholesterol treated with a cholesterol-lowering drug | Cost (£) per reduction in units of blood cholesterol (mmols or mg/l) |
| Depression treated with a drug that improves mood | Cost (£) per reduction in units of a rating scale that scores depression |
| Heart failure treated with a drug that improves cardiac performance | Cost (£) per unit increase in the output of the heart |
| Breast cancer treated with a drug that prevents recurrence | Cost (£) per additional life year gained as a result of effective treatment |

Cost–effectiveness analysis is useful if you are comparing two or more treatments for the same condition and therefore can be measured using the same measure of outcome. It falls down, however, when you try to compare a treatment for growth hormone deficiency, with a treatment that prevents recurrence of breast cancer. In this instance the measurements of ‘value-for-money’ are expressed in different ways (i.e. cost per additional inches of height with growth hormone versus the cost per extra year of life that is gained from treating breast cancer). Not only can they not be compared; they can’t even be used to decide which represents the better value-for-money and is therefore the ‘better buy’.

Put another way, cost–effectiveness analysis is unsatisfactory at comparing one treatment for one condition, with another treatment for a completely different condition. NICE has, on occasions, been forced to use cost–effectiveness analysis when there has been no alternative. For example, in its assessment of human growth hormone, the Institute had to use the cost per cm increase in height (£ per cm) because there was no other reliable alternative.

Cost–utility analysis: In cost–utility analysis the impact of different treatments are compared on a common measuring instrument or ‘yardstick’. The yardstick used is

the quality adjusted life year (the QALY). This tries to express the benefits of treatment as an increase in both the quality and length of life.

How is Quality of Life measured?

Traditional methods of evaluating disease and response to treatment examined the process of health care from the point of the physician, focusing on symptoms, cure and survival/mortality. In the past few decades however the focus has moved towards the patient; and health assessment techniques have evolved to include mental and social aspects of health, which are broadly categorised under the term 'quality of life' (QoL).

QoL is affected by many factors and is referred to using many terms. In the medical context the focus is usually on aspects of personal experience of healthcare and illness, and is often referred to as 'Health-related QoL'.

The EQ-5D

To quantify the effects of technologies on the QoL for patients with different diseases, a questionnaire called the EuroQol-5D (or (EQ-5D) (<http://www.euroqol.org/>) is widely used. This is one of a number of QoL questionnaires but few are able to assess 'utilities'. The EQ-5D has been tested in many different patient populations. It comprises five dimensions (see table 2) and each has three possible levels; (1) no problem, (2) some problems, and (3) major problems.

Table 2: The EQ-5D

| |
|--|
| Mobility <ol style="list-style-type: none">1. I have no problems in walking about2. I have some problems in walking about3. I am confined to bed |
| Self-Care <ol style="list-style-type: none">1. I have no problems with self-care2. I have some problems washing or dressing myself3. I am unable to wash or dress myself |
| Usual activities (<i>e.g. work, study, housework, family or leisure activities</i>) <ol style="list-style-type: none">1. I have no problems with performing my usual activities2. I have some problems with performing my usual activities3. I am unable to perform my usual activities |
| Pain/Discomfort <ol style="list-style-type: none">1. I have no pain or discomfort2. I have moderate pain or discomfort3. I have extreme pain or discomfort |
| Anxiety/Depression <ol style="list-style-type: none">1. I am not anxious or depressed2. I am moderately anxious or depressed3. I am extremely anxious or depressed |

There are a total of 243 possible different combinations of responses. Each different combination is termed a 'health state'. There are an additional two possible health states: unconscious and dead.

Each health state is converted to a numerical value known as a 'health utility' (see table 2). These range from 0 to 1, where 0 is equivalent to being dead and 1 represents the best possible health state. However, some health states are regarded as being worse than 0 and are given a negative value. For example, some people may consider being in a permanent vegetative state worse than death and so would give such a health state a negative utility.

Table 3: EQ-5D health state valuations (utilities)

| Health State | Description | Utility |
|---------------------|---|----------------|
| 11111 | No problems | 1.000 |
| 11221 | No problems walking about; no problems with self-care; some problems with performing usual activities; some pain or discomfort; not anxious or depressed. | 0.76 |
| 22222 | Some problems walking about; some problems washing or dressing self; some problems with performing usual activities; moderate pain or discomfort; moderately anxious or depressed | 0.516 |
| 12321 | No problems walking about; some problems washing or dressing self; unable to perform usual activities; some pain or discomfort; not anxious or depressed. | 0.329 |
| 21123 | Some problems walking about; no problems with self-care; no problems with performing usual activities; moderate pain or discomfort; extremely anxious or depressed. | 0.222 |
| 23322 | Some problems walking about; unable to wash or dress self, unable to perform usual activities, moderate pain or discomfort, moderately anxious or depressed | 0.079 |
| 33332 | Confined to bed; unable to wash or dress self; unable to perform usual activities; extreme pain or discomfort; moderately anxious or depressed. | -0.429 |

The value of the health utility derived from each health state is based on a very intricate study involving the preferences of 3000 members of the UK general public. The health utility values obtained from this study indicate that the 5 dimensions are not considered equal. For example, rating level 3 in the 'mobility' dimension (see Table 1) is considered to be much more serious and to be a greater impediment to QoL (i.e. is associated with a greater loss in QoL) than rating a level 3 in 'usual activities'. This reflects the views, attitudes and preferences of the British public who consider 'being confined to bed' a much greater impediment to their QoL than 'being unable to perform my usual activities'.

It is important to appreciate, however, that although it is the general public's values that are used to convert the health states into utilities, the actual health states themselves are obtained from patients with the specific condition. It has been suggested that, rather than those of the general public, it would be preferable to use values from actual patients to convert the health state into a utility. However, this would be impractical as the experiment would need to be repeated each time. Moreover, there is also the argument that in a publicly funded healthcare system – such as the NHS – it is right that the public should determine the values given to individual elements of QoL because all may potentially benefit (or not) from a decision.

Quality-adjusted life years

NICE uses cost–utility analysis to compare healthcare interventions. This method uses the number of quality-adjusted life years (QALY) that are gained by a treatment measure of 'benefit'. Calculating the number of QALYs gained is reasonably straightforward. The utility value of a health state (derived, for example, using the EQ–5D) is multiplied by the length of time that the individual is in that health state. The 'cost per QALY' is the cost of the treatment divided by the total numbers of QALYs gained.

For example, let's say that a hip replacement improves the utility of the average arthritic person from 0.8 to 0.9 (gain of 0.1). Let us say that, again on average, people have their hip replaced at the age of 65 years; and that their average life expectancy, at this age, is 15 years. They therefore gain $0.1 \times 15 = 1.5$ QALYs from the operation. If the cost of a hip replacement is £5,000 then the cost per QALY is:

$$5,000 \div 1.5 = \text{£}3,333 \text{ per QALY}$$

QALYs provide a yardstick to compare two (or more) interventions in terms of their impact on both the length and quality of life. The number of QALYs gained with each intervention can then be directly compared. For example if intervention A increases average utility by 0.2, for 4 years, the gain is $4 \times 0.2 = 0.8$ QALYs. If intervention B results in four years in a health utility gain of 0.5 for 6 months (0.5 years), the QALY gain is $0.5 \times 0.5 = 0.25$ QALYs. Therefore intervention A generates an additional 0.55 QALYs compared to intervention B ($0.8 - 0.25$).

If intervention A costs £10,000 the cost per QALY will be £12,500 ($10,000 \div 0.8$). If intervention B costs £2,000 then the cost per QALY will be £8,000. ($2,000 \div 0.25$). Intervention B would therefore represent better 'value for money' than intervention A despite gaining fewer QALYs.

NICE uses the QALY as an outcome measure because it takes into account both the increase in life expectancy from an intervention as well as the quality of that increase in life. This reflects the value judgment that living longer, in itself, is an insufficient measure of success; and that the expected QoL, during that extra time also needs to be considered. Of course many treatments (such as hip replacement surgery) do not increase longevity but give a better QoL during a person's remaining years.

When health economists calculate QALYs they do not take any account of how many a person has already had in the past. For example if a woman is successfully treated for breast cancer and gains 30 QALYs, it makes no difference to the calculation of the number of QALYs she gets when her hip is replaced 15 years later.

Appendix 3: The Agenda

| NICE Citizens Council: QALYs | | |
|--|---|--|
| Time | Title | Who |
| Thursday 31st January – MidCity Place | | |
| 9.00am – 9.15am | Welcome from facilitators | Ela and Brendan |
| 9.15am – 9.30am | Update on second edition of <i>Social Value Judgments: Principles for the Development of NICE Guidance</i> | Prof Peter Littlejohns |
| 9.30am – 9.45am | Welcome from NICE and introduction to NICE’s question | Andrew Dillon and Prof Peter Littlejohns |
| 9:45am – 10:15 | Measuring the quality of life of the Citizens Council Health-Related Quality of Life (HRQoL) exercise (EQ – 5D) | Dr Joanne Lord , Technical Advisor, NICE |
| 10.15am – 11:00am | Quality of life and how it is measured Explaining HRQoL and HRQoL measurement | Prof Ann Bowling , Professor of Health Services Research, University College London |
| 11.00am – 11.15am | BREAK | |
| 11.15am – 12:00pm | What are QALYs? Understanding QALYs including cost per QALY | Dr Joanne Lord , Technical Advisor, NICE |
| 12.00pm – 1.00pm | What do you make of this so far? Citizens Council Discussion | |
| 1.00pm – 2.00pm | LUNCH | |

| NICE Citizens Council: QALYs | | |
|---|--|--|
| Time | Title | Who |
| 2.00pm – 3.00pm | How did the Citizens Council rate quality of life? Results of HRQoL exercise and discussion | Discussion with Dr Joanne Lord , Technical Advisor, NICE & Susan Bennett , patient representative. |
| 3.00pm – 3.15pm | BREAK | |
| 3.15pm – 4.15pm | Does severity matter? What do the Citizens council think so far? | Discussion with Dr Marcia Kelson , Associate Director, NICE - Patient and Public Involvement |
| 4.15pm - 5.00pm | Any questions? | Prof Peter Littlejohns |
| Friday 1st February – MidCity Place | | |
| 9.00am – 9.15am | What are the Citizens Council thinking so far? An opportunity to recap and clarify thoughts | Ela and Geoff |
| 9.15am – 10.15am | Setting the scene: QALYs and weighting What does this mean and how does this effect decision making? | Prof John Cairns , Professor of Health Economics, London School of Hygiene and Tropical Medicine |
| 10:15am – 10:30am | Discussion | |
| 10.30am – 10.45am | BREAK Please note that the MidCity Place fire alarm is tested at 10.30am every Friday. Unless the ringing continues, there is no need to vacate the building. | |
| 10.45am – 12.00pm | What does the general public think about QALYs and severity? Part I Looking at NICE's commissioned research | Prof Paul Dolan , Imperial College London and Richard Edlin , University of Leeds |

| NICE Citizens Council: QALYs | | |
|---|---|---|
| Time | Title | Who |
| 12.00pm – 12.45pm | LUNCH | |
| 12.45pm – 2.00pm | <p>What does the general public think about QALYs and severity? Part II</p> <p>Looking at NICE's commissioned research</p> | Dr Rachel Baker , Lecturer (Health Economics), Newcastle University. |
| 2.00pm – 3.30pm | <p>Should NICE and its advisory bodies take into account the severity of the disease when making decisions and should severity be included in the calculation of the QALY?</p> <p>A panel discussion</p> | <p>Prof Paul Dolan Richard Edlin Dr Rachel Baker Dr Marcia Kelson Prof John Cairns Peter Mansell, Director. for Patient Experience and Public Involvement, National Patient Safety Agency Rachel Rowson, Policy and Public Affairs Manager, Cancerbackup</p> |
| 3.30pm – 3.45pm | BREAK | |
| 3.45pm – 4.45pm | Discussion | |
| 4.45 – 5.00pm | Final thoughts for the day | Ela & Geoff |
| Saturday 2nd February – Holiday Inn, Bloomsbury | | |
| 9.30am – 9.45am | Recap from the day before | Ela & Geoff |
| 9.45am – 12.30pm | Drawing conclusions and deciding what goes into the report. | |
| 12.30pm – 12.40pm | Close and Thanks | Prof Peter Littlejohns |
| 12.40pm | LUNCH | |

Appendix 4: Speaker Biographies

Rachel Baker

Rachel Baker is a lecturer in health economics at the Institute of Health and Society, Newcastle University. She graduated with a BA in Economics and Social Policy in 1995 and with a PhD also from Newcastle University in 2003. Her doctoral work, funded by the Medical Research Council, involved the use of novel methods to explore economic theories of rational choice in the context of the health and lifestyle choices made by people with type 2 diabetes.

Her postdoctoral fellowship from the Economic and Social Research Council included three months as a visiting researcher at the University of Calgary, Canada. She has published academic papers on the use of qualitative methods and Q methodology in health economics and her work has particularly focussed on the rationales individuals give for their responses to preference elicitation techniques used by health economists, such as 'standard gamble' and 'willingness to pay' – and how those rationales fit with the theoretical foundations of such techniques.

For the past three years Rachel has been the main researcher on the 'Social value of a QALY' (SVQ) project with a team of academics from the universities of Newcastle, East Anglia and Aberdeen. She is now engaged in a European research project (EuroVaQ) funded by the European Commission and involving partners from ten European countries.

Sue Bennett

Sue Bennett was a healthcare professional in the NHS for 18 years and completed a master's degree in medical research in 1993. A spinal condition led to early retirement. In more recent years she was a member of the Public Reference Group for the NHS Information Authority from 2003 - 2005. From 2005-2007 she sat on the NICE Faecal Incontinence Guideline Development Group. In 2007 Sue was appointed to the Interventional Procedures Advisory Committee at NICE and to a pilot NICE/ NPSA Patient Safety Advisory Committee. Sue has been a Trustee of *Incontact*, an organisation for people with bladder and bowel problems since 2002. She is also a member of RADAR (Royal Association for Disability and Rehabilitation) and has a keen interest in health, disability and equality issues. Other interests include genealogy and tall ship sailing.

Professor Ann Bowling

Professor of Health Services Research, University College London.

Ann Bowling is a social scientist and is also seconded part time to the MRC Health Services Research Collaboration in Bristol, where she is exploring patients' preferences for treatment for angina. Her interests focus on social gerontology, in particular on quality of life in older age and on issues of equity of access to health care. Recently completed projects include a national survey of quality of life in older age and clinical decision-making in cardiology by age of patient (equity issues) (both funded by ESRC). She has an interest in methodology and has written three best selling texts on quality of life measurement and on research methods (all published by the Open University Press). She is currently a member of the Health Technology Assessment Board, the HTA Methodology Panel and the MRC's College of Experts assigned to the Health Services and Public Health Board.

John Cairns

John Cairns is Professor of Health Economics at the Department of Public Health and Policy, London School of Hygiene and Tropical Medicine. He has previously worked at the Universities of Aberdeen and York. Since 2003 he has been a member of the NICE Appraisal Committee and the Scottish Medicines Consortium. He is a member of the Health Technology Assessment Commissioning Board and of the recently formed advisory committee on the Safety of Blood, Tissues and Organs.

Paul Dolan

Paul Dolan is Professor of Economics at Tanaka Business School, Imperial College London. The general theme of his research activities is how individual and social well-being should be defined, measured and distributed for the purposes of informing public policy. He has held academic appointments at the Universities of York, Newcastle and Sheffield, where he became Professor of Economics in 2000. Professor Dolan was awarded a Philip Leverhulme Prize in Economics in 2002 for his contribution to health economics and was a visiting professor at Princeton University in 2005.

Richard Edlin

An economist by training, Richard came to the UK in 1998 to study for a PhD. His thesis considered how the views of the public could be used to inform public policy decisions, considering both the elicitation of such preferences and a possible interpretation within health economics. Richard worked as both a Research Associate and Lecturer at the University of Sheffield between 2002 and 2007, before joining AUHE in August 2007.

Richard's research interests are broadly in the area of equity, and particularly the elicitation, interpretation and use of the public's preferences in resource allocation decisions. Current research (nearing completion) includes a NICE and National Collaborating Centre for Research Methodology funded study looking at the societal value of health gains to different individuals. Previous research has considered the implicit trade-offs between cost-effectiveness, equity and access.

Marcia Kelson

Associate Director - Patient and Public Involvement NICE.

Dr Joanne Lord

Technical Adviser NICE.

Peter Mansell

Peter is the National Patient Safety Agency's Director for Patient Experience. His role is to provide the major lay contribution to the overall success of the NPSA through the development and implementation of effective lay participation and influence in NPSA's work.

Peter left school without any qualifications and worked in various manual jobs until becoming paraplegic at the age of 20, in 1978, through a road traffic accident. Peter then studied and gained a Degree in social policy and administration and a postgraduate Diploma in management.

Peter's injury led to an interest in health and disability and the social policy issues surrounding them, and to his working for many organisations in this field. Peter has been Chief Executive of both the Spinal Injuries Association and The Royal Association for Disability and Rehabilitation, as well as Chairman of the Royal National Orthopaedic Hospital Trust.

Peter has spent over three years as an inpatient in hospitals and has experienced medical error and harm.

Peter is 49 years old and lives in south London with a wife Karen and his youngest son Alex (17). His other son, Greg, is at University in Hull studying German.

Rachel Rowson

Rachel Rowson is Policy and Public Affairs Manager at Cancerbackup, Europe's leading cancer information charity. Prior to joining Cancerbackup, Rachel worked at a healthcare focussed political consultancy with a strong emphasis on advising clients about health technology appraisals, and in a pharmaceutical company working in public affairs and economic policy.

Appendix 5: The Council Members

Jonathan Barwick – is a lecturer and trainer in hospitality and travel at a Further & Higher Education college in Norfolk.

David Batchelor – an engineer from Leicester.

Michael Beecroft – a self-employed driving instructor from Lincolnshire.

Andrew Callaghan – a gardener from West Yorkshire.

James Cambourne – is a restorer and decorator from Wolverhampton.

Steven Coad – an industrial safety engineer who lives in County Durham.

Tim Duckworth – a courier from Bury in Lancashire.

Freda England – works for the Citizens Advice Bureau and is from Lymington in Hampshire.

Ron Findley – a database administrator from London.

Geraldine Fost - a retired careers guidance manager, who lives in Hungerford, Berkshire.

Alan Garvey – an auto engineer who lives in Manchester.

Lorna Girling - lives in Norfolk, and is a part time literature student and a housewife and mother of two.

Terry Hamer - lives in Southampton. He works on the cruise ships at the terminal.

Meryl Hobbs – a retired teacher and farmer's wife from Herefordshire.

Kelly Hughf – a veterinary nurse who comes from Bishop Auckland.

Susan Jackson – is a cabin crew member from Surrey.

Robert Jones - works as a warehouse operative and is a football referee in his spare time. He lives in Cwmbran, Wales.

Catherine Kaer-Jones – a student support leader working in a Bradford school.

Jack Kelley – is from Doncaster and worked in the construction industry but is now in security.

Claire Marshall – is a freelance writer from London.

Tina McDonnell – a trainer with a High Street bank from London.

Freda McEwan – a witness liaison officer for the Metropolitan Police, from London.

Christine Minton – a retired community service unit manager for the Probation Service, living in Essex.

Linda Moss - currently unemployed, trained as a TEFL teacher and now lives in Todmorden, West Yorkshire.

Patricia Roberts – an accounts assistant from Flintshire.

Heena Sabir - worked for a while in human resources, and has recently moved to Huddersfield, where she is looking for suitable work.

Mohammed Shakil – is from Rotherham and training to be an accountant.

Ian Smith – a musician from Manchester.

Rebecca Sparling – a full time university student living in Birmingham.

Paddy Storrie - a secondary school Deputy Headteacher, lives in Harpenden, Herts.