Health policy

Adjusting for treatment refusal in rationing decisions

Richard Lilford, Alan Girling, Andrew Stevens, Abdullah Almasri, Mohammed A Mohammed, David Braunholtz

Assessments of cost effectiveness are increasingly used to get the most value from limited health resources. Could adjusting for people who wouldn't want the treatment improve the process?

Department of Public Health and Epidemiology, University of Birmingham, Birmingham B15 2TT

Richard Lilford
professor of clinical epidemiology
Alan Girling
senior research fellow
Andrew Stevens
professor of public health

Mohammed A Mohammed
senior research fellow

Department of Primary Care and General Practice, University of Birmingham, Abdullah Almasri
visiting statistician

Bill and Melinda Gates Foundation, University of Aberdeen, Aberdeen David Braunholtz
senior statistician

Correspondence to: R Lilford
R.J.Lilford@bham.ac.uk

BMJ 2006;332:542-4

Many treatments improve outcomes with few material side effects. They may be expensive, but if they are made available to patients almost all would accept. However, some treatments have more serious side effects, such that a substantial minority of patients would decline. Rationing bodies such as the National Institute for Health and Clinical Excellence treat both scenarios in the same way, calculating an average quality adjusted life year (QALY) for all candidates for the new treatment. Yet people who decline do so because they have different preferences and hence different expectations about the prospect of gain from the treatment. By declining the treatment, they cannot benefit from it. In these situations we argue that considering the QALYs of only patients likely to accept the treatment will lead to a less biased assessment of cost effectiveness.

Split choice decisions

QALYs are the main measure of health gain used in cost utility assessments to determine whether a treatment or intervention should be funded. Although other social dimensions (such as the pursuit of equity) may be included in the appraisal, our discussion assumes that the health gain is an important factor. Our argument applies to any health technology that a patient may decline or accept and hence is not relevant to public health measures, such as fluoridation of the water or public advertising, that are delivered to all individuals in a population, whether they want them or not.

A group of patients with similar clinical characteristics (which we refer to as a clinical group) carry the same probability in terms of outcome of treatment. However, they may differ in their capacity to gain personal utility.¹ This is because of differences between individual valuations of the possible treatment outcomes. For example, two 65 year old men with moderately differentiated localised prostate cancer may value the side effects of radical surgery (incontinence or impotence) differently, and two pregnant women with a risk of a Down's syndrome baby of 1 in 200 may trade-off procedure related miscarriage and Down's syndrome differently. In either case, we must suppose that each individual is choosing the treatment that offers the greatest prospective personal utility.

This idea can be operationalised by conceptualising the prospective utility of a treatment as its expected utility, which is an average of utility values weighted by the probabilities of these values occurring—that is, by the probabilities of the possible treatment outcomes. For the purposes of our argument, we suppose that personal valuations of

Frequency diagrams showing expected gain in QALYs from treatment under two scenarios. In both cases the expected QALY gain follows a normal distribution in the population, with a mean of 1 QALY. For curve A the standard deviation of the expected gain is 0.29 QALYs and nearly all are prospective gainers. In curve B the standard deviation is 1 and only 84% are prospective gainers. The conditional group mean (calculated from those who gain QALYs) is 1.29

treatment outcomes can be expressed on a personal utility scale, on which immediate death is 0 and a single year of healthy life is 1.

Treatments can be divided into two types. The first type is treatments for which the expected personal QALY gain is greater than that from the best available alternative treatment for all members of a clinical group (curve A in the figure). For example, chemotherapy is indicated in all children with acute lymphoblastic leukaemia, even though it may cause some secondary tumours in later life. The second type is treatments that are a matter of personal choice—that is, where members of the clinical group individually choose to accept or decline the treatment because, and only because, they place different values on the possible treatment outcomes (curve B in the figure). We refer to these as split choice treatments.

Treatments in the split choice category divide the clinical group into two parts: people who gain the expected QALYs from the treatment (prospective gainers) and those for whom the treatment comes with an expected loss of QALYs (prospective losers) and who should therefore decline (see shaded area of curve B).² In the Down's syndrome scenario above, suppose, for simplicity, that half the group is averse to prenatal diagnosis and would have an expected personal utility loss of 2 QALYs if it were carried out. The other half favours the procedure to the extent that they would expect to gain 2 QALYs. Conventional analysis would assign an average societal valuation to the diagnostic procedure of zero, disregarding the fact that it would only be accepted by those patients who expected to gain utility from it.¹⁻³ We contend that it is more

reasonable to value the procedure with reference only to those likely to accept it, in which case a valuation of 2 QALYs seems appropriate.

Normative model for rationing under split choice

When there are both prospective losers and gainers from a treatment, we suggest using the mean expected benefit of the prospective gainers in calculating the societal value of a new treatment. Since this is the mean of a section of the clinical group, we call it the conditional mean—the mean conditional on being one of the people for whom the treatment comes with an expected gain in QALYs. By contrast, QALYs based on the overall group mean are biased, and we refer to this as split choice bias.

As an illustration, the distribution of expected QALY gains across the clinical group in curve B in the figure includes negative values (16%) as well as positive values (84%). If the curve applied to the clinical group with moderately differentiated localised prostate cancer it would imply that 84% of men might gain expected QALYs from radical treatment while 16% are prospective losers. All patients in the group have the same probabilities of benefit and harm, but the prospective losers place a particular value on avoiding the side effects of treatment. The mean expected utility increases by nearly 30% if the conditional mean (mean of the 84% who are prospective gainers) is used instead of the mean for the clinical group.

The effect on cost utility analysis would be considerable. For instance, if the cost of the treatment were £32 000 for each patient it would not be considered cost effective in England as the mean population benefit is only 1 QALY. However, if the conditional mean of 1.3 QALYs were used, the treatment would become cost effective, assuming a threshold of £25 000 per QALY for cost effectiveness.

Split choice bias increases in an accelerating non-linear fashion as the proportion of the population that lose expected QALYs rises. Thus, the QALY gain would be underestimated by a half if just 30% of a clinical group lose expected QALYs and decline treatment. Situations where more than 30% of an eligible patient population are prospective losers and will decline treatment include prenatal diagnosis for sickle cell disease,⁴ hormone replacement therapy for prevention of osteoporosis and colon cancer,⁵ bilateral prophylactic mastectomy for autosomal dominant breast cancer,⁶ and screening for prostate cancer.^{7 8}

The decision model based on expected utility is normative—it tells us who should (rather than does) accept or decline a treatment, in a rational world. It is also important to consider factors in the real world that will deflect practice from the ideal. This enables various assumptions or findings about real world behaviour to be factored into models. Furthermore, it may show how progress can be made to move real world practice in a normative direction.

Limitations of the split choice model

People may make a suboptimal choice about treatment because of poor communication, cognitive limitations, over-reliance on advisers, compliant behaviour,⁹ or

other unspecified reasons. The result is that some people may refuse treatment even though they are prospective gainers, while some prospective losers may accept. However, provided decision making is better than random, the overall mean will produce an estimate of expected QALY gain that is too low when there is split choice. Indeed, prospective gainers who refuse may actually increase the difference between the conditional and the overall means, thereby increasing the bias we have identified. This is because they are likely to come preferentially from people who gain less than the mean gain. Different assumptions about the proportion of people who wrongly refuse or wrongly accept could be incorporated into the model to provide an estimate of expected QALYs among those who actually accept. Alternatively, treatment preferences could be elicited alongside personal utilities to provide an empirical estimate of the subgroup whose gain ought to be included in any cost utility analysis.

We have not considered aspects such as regret adding to the disutility of an outcome after a decision not to have a treatment¹⁰ and anxiety from being offered a choice. However, these exist irrespective of whether an overall mean or a conditional mean is used in cost utility calculations and bear only tangentially on the issue of which of the means is more appropriate.

In practice it is unusual to construct a distribution of expected utilities. The mean expected utility (embedded in confidence limits and associated with sensitivity analysis) is usually fit for purpose. But split decisions will result in biased measurements that become rapidly increasingly biased as the proportion of those who refuse increases, unless a conditional mean is used. Calculating the conditional mean requires a full distribution of expected utilities or QALYs. Such a distribution could be derived from the full distribution of values for the various outcomes of care obtained by one of the established mechanisms recently reviewed by Ryan and colleagues.¹¹ The variation in this distribution is at best a proxy for the variation between patients. It does not have to be a perfect proxy for our argument to hold, but the more the respondents are placed in the shoes of real world patients, the better the approximation between the expected utilities in rationing decisions and those encountered by such patients in practice.

Nevertheless, in some circumstances a concern for the preferences that drive personal choice may be misplaced, especially if the effect of a treatment extends beyond the individual who receives it. For example, the decision to make vaccine available may be predicated on the perceived need to induce herd immunity rather than the wish to maximise individual expected utility or QALYs, and its societal valuation should reflect this.

We thank Professor Stirling Bryan for helpful advice and the reviewer, Joanna Coast, for comments that helped clarify our argument. We also thank Ian Young, Robert Silman, John and Ghislaine Young, Jerry Marsden, Michael Foster, Simon Rummens, and Jayne Parry for helpful comments on the article.

The study was supported by the EPSRC as part of the Multidisciplinary Assessment of Technology Healthcare Consortium (MATCH) and the NHS Research and Development Methodology programme.

Contributors and sources: RL has a long interest in clinical decision analysis, arising in part from his experience of talking to families about prenatal diagnosis when he was an obstetrician. RL and AS are members of the NICE appraisal committee and

Summary points

Rationing decisions are currently based on the cost of the average gain from a treatment

Some patients may not want certain treatments because they weigh the side effects more than the gains

Patients who would decline treatment should not be included in assessment of average gain

Excluding these patients increases the cost effectiveness of a treatment

AS has published extensively on health needs assessment. AG, MM, AA, and DB have interests in statistics and modelling and AG is sponsored by MATCH to develop novel approaches to modelling. RL conceived the idea for this paper in discussion with MM and DB. RL, AG, and AS wrote the manuscript, AA assisted with some calculations. RL is the guarantor.

Competing interests: None declared.

- 1 Thornton JG, Hewison J, Lilford RJ, Vail A. A randomised trial of three methods of giving information about prenatal testing. *BMJ* 1995;311:1127-30.
- 2 Lilford RJ, Jackson J. Equipoise and the ethics of randomisation. *J R Soc Med* 1995;88:552-9.
- 3 US Congress, Office of Technology Assessment. *Costs and effectiveness of prostate cancer screening in elderly men*. Washington DC, USA: US Government Printing Office, 1995.
- 4 Davies SC, Cronin E, Greengross P, Hickman M, Normand C. Screening for sickle cell disease and thalassaemia: systematic review with supplementary research. *Health Technol Assess* 2000;4(3):34-5.
- 5 Hundrup AY, Thoning H, Obel EB, Rasmussen NK, Philip J. Lifestyle factors and choice of hormone replacement therapy among Danish nurses. *Scand J Public Health* 2002;30(1):47-53.
- 6 Metcalfe KA, Snyder C, Seidel J, Hanna D, Lynch HT, Narod S. The use of preventative measures among healthy women who carry a BRCA1 or BRCA2 mutation. *Fam Cancer* 2005;4:97-103.
- 7 Kirby RS, Kirby MG, Feneley MR, McNicholas T, McLean A, Webb JA. Screening for carcinoma of the prostate: a GP based study. *Br J Urol* 1994;74:64-71.
- 8 Chadwick DJ, Kemple T, Astley JP, MacIver G, Gillatt DA, Abrams P, et al. Pilot study of screening for prostate cancer in general practice. *Lancet* 1991;338:613-6.
- 9 Bekker H, Thornton J, Airey CM, Connelly JD, Hewison J, Robinson MB, et al. Informed decision making: an annotated bibliography and systematic review. *Health Technol Assess* 1999;3(1).
- 10 Zeelenberg M. Anticipated regret, expected feedback and behavioral decision making. *J Behav Decis Making* 1999;12:93-106.
- 11 Ryan M, Scott DA, Reeves C, Bate A, van Teijlingen ER, Russell EM, et al. Eliciting public preferences for healthcare: a systematic review of techniques. *Health Technol Assess* 2001;5(5).

(Accepted 1 December 2005)

Ethics

One-time general consent for research on biological samples

David Wendler

It is now recognised that people should give informed consent for use of their biological samples in research. The literature on individuals' views supports one-time general consent as the best approach for this purpose

Department of
Clinical Bioethics,
NIH Clinical
Center, Bethesda,
MD 20892, USA
David Wendler
staff scientist
dwendler@nih.gov

BMJ 2006;332:544-7

Clinicians and clinical researchers routinely obtain human biological samples and store them for future research. Previously, samples were often stored and used without informed consent. Most investigators thought that the protections in place for research with human participants were not needed for research with human samples.

Recent commentators recognise the importance of informed consent for research with biological samples, but they disagree about when it is required and what types of consent should be obtained. Some argue that people should provide consent for each new study, at the time the study is proposed. Others support prospective consent, but they endorse different sets of options regarding the storage of samples, which investigators can use the samples, and what types of research can be performed.

This profusion of guidance has led to divergent practices,^{1 2} with one study finding that institutional review boards in the United States recommend "various consent options, all of which are different."³ Such variation may undermine the scientific value of patients' contributed samples and greatly increase the costs of such research. Assessing whether people's views support one of the recommended options for consent could provide a solution to these problems.

Methods

I searched PubMed (see appendix 1 on bmj.com) for studies published in English that reported the views of individuals on consent for research with human biological samples. This search identified 30 eligible studies.

Results

Individual willingness

The 30 studies provide data on the views of more than 33 000 people (table).⁴⁻³² The studies assessed the views of patients, research participants, family members, religious leaders, and the public. The studies yield consistent findings, despite being conducted around the world, over a 10 year period, in different groups, using different methods. Most respondents want to decide whether their samples are used for research purposes.

Of the 20 studies that assessed willingness to donate, 17 found that at least 80% of respondents would donate a sample if asked. Respondents who

Appendices showing the search terms used and the consent process for research with human biological samples are on bmj.com