



ELSEVIER

Health Policy 51 (2000) 19–30

HEALTH policy

www.elsevier.com/locate/healthpol

A qualitative study of the extent to which health gain matters when choosing between groups of patients

Paul Dolan ^{a,*}, Richard Cookson ^b

^a *Sheffield Health Economics Group and Department of Economics, University of Sheffield, Sheffield S1 4DA, UK*

^b *LSE Health, London School of Economics, London WC2A 2AE, UK*

Received 27 March 1999; accepted 17 November 1999

Abstract

There is considerable debate about the appropriateness of allocating health care resources on the basis of the size of the health improvement that they generate. The aim of this study was to elicit the general public's views about the extent to which health gain matters vis-à-vis other considerations. A total of 60 respondents took part in group discussions designed to enable them to raise, discuss, and reflect upon, different arguments. The qualitative data showed that many responses were being generated by factors that were not directly included in the questions, and so it is difficult to meaningfully interpret the results of other studies which have asked similar questions but which have not looked at the reasons underlying the responses. However, a clear message did come through from the data; namely, that equality of access should prevail over the maximisation of benefits. However, this was subject to the outcome constraint that treatments are sufficiently effective. An important question for future research, then, is 'how effective do treatments have to be for the principle of equal access to apply?' © 2000 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: Health gain; Access; Equity; Justice

* Corresponding author. Tel.: +44-114-2220670.

E-mail address: p.dolan@sheffield.ac.uk (P. Dolan)

0168-8510/00/\$ - see front matter © 2000 Elsevier Science Ireland Ltd. All rights reserved.

PII: S0168-8510(99)00079-2

1. Introduction

There has been considerable debate in the literature regarding the appropriateness of allocating health care resources on the basis of the size of the health improvement that they generate. Health economists have generally argued that resources should be distributed so as to improve aggregate health as much as possible, whereas others have argued that this would lead to unacceptable discrimination against the elderly, the infirm, and other vulnerable groups in society with lower-than-average capacity to benefit from treatment [1,2].

As this debate continues, an increasing number of empirical studies are being undertaken by health economists and others to measure the extent to which people are willing to trade-off health gain for other considerations, such as the initial severity of illness or the final distribution of health. These studies have often asked people to adopt the perspective of a decision-maker [3–5] but some have placed them in the role of a potential patient behind a ‘veil of ignorance’ where they do not know which patient they would be [6].

The empirical studies cited above have asked respondents to give precise quantitative trade-offs to tightly defined questionnaires, often with little or no time for deliberation. Given that it is now well-established that seemingly subtle changes in question framing can change the stated preferences of respondents [7,8], it is not surprising that there is great variation in the quantitative results that these studies have generated. Rather than try to quantify the precise trade-offs of respondents, the aim of this study was to elicit more general (and arguably more meaningful) qualitative information from people regarding the extent to which health gain matters vis-à-vis other considerations. By asking some people to adopt the role of a decision-maker and by placing others behind a veil of ignorance, it was also an attempt to see whether people’s views are a function of the perspective they are asked to adopt.

2. Methods

2.1. *Study design*

To mitigate against the possibility of drawing erroneous conclusions about people’s views, it is important (i) to explore how respondents interpret questions that are put to them (particularly the extent to which they find them plausible), and (ii) to give them time to think about what is being asked of them and opportunity to reflect upon their responses. To achieve this, focus groups were convened to enable respondents to raise, discuss, and reflect upon, different arguments. All group discussions were tape recorded and transcribed so that insights could be gained into the cognitive processes that respondents used in order to arrive at their responses.

The aim was to recruit 10 groups of six people who would meet for about two hours on two separate occasions with a fortnight between each meeting. It was felt

that a relatively small number of people in each group (which is increasingly becoming the norm [9]) would provide a better environment for discussing the issues. Letters of invitation were sent out to 1000 people who were randomly chosen from two GP practice lists in York. They were given a list of times and venues and asked to indicate which pair(s) of dates they were willing to attend. Each was told that they would be paid £30 (at the end of the second meeting) for attending. Potential respondents were asked to indicate their gender and which of three age groups they belonged to (18–34, 35–55, or 55–70 years). There were 207 (21%) positive replies from which 72 people were invited to participate (8 groups of 7 and 2 groups of 8). Twelve people from each gender-age category were chosen and it was possible to invite at least one person from each category to attend each group.

All group discussions were moderated by both authors. The first group meeting consisted of two main parts. In the first part, respondents were initially asked to discuss the question, ‘How would you set priorities in health care?’ and were then presented with a questionnaire which they were asked to fill in without discussion. In the second part of the first meeting, respondents were asked questions about allocating resources on the basis of health gain, and it is these questions that are the focus of the present paper.

Respondents were asked to imagine that there are two groups of patients who will both benefit from treatment (either in terms of length of life or quality of life) but by differing amounts or from different start-points. The tightly defined questions which followed were designed to focus attention squarely on the distribution of health outcomes between different patient groups, and so respondents were asked to assume that this was the only respect in which patients in the two groups differed. In common with many other preference elicitation studies, including those cited above, the questions deliberately made simplifying assumptions which set aside many of the uncertainties and complications of real-world situations in order to focus precisely on the issue of interest.

Respondents were told that only half of the patients could be treated. After a discussion, they were asked to decide whether they would choose to give the same priority to both groups or to give priority to the group that can gain the most from treatment (or, in the questions where the health gain was identical, to one group according to their start-point). Based on the answer to this question, the health gain to the group who gained least was either increased or reduced until the decision became most difficult. That is, respondents were asked by the moderator to state the point at which they would ‘draw the line’ between giving the same priority to both groups, and giving priority to the group who would benefit most from treatment. There were six questions in total: three focused on gains in life years and three focused on gains in quality of life, presented in terms of percentages of full health. Table 1 presents the initial choice faced by respondents in each question.

Five of the groups were assigned to a ‘social decision-maker’ condition, in which respondents were asked to imagine that they are a group of decision-makers who have to decide between the two groups of patients in the knowledge that they will not personally be in either of the groups. And five groups were assigned to a ‘veil

Table 1
The questions

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

of ignorance' condition, in which respondents were told to imagine that they will personally be in one of the groups, but they do not know which one.

3. Qualitative analysis

The aim of this analysis was to build up a picture of what respondents were thinking about when answering the questions they were presented with. There are many ways in which this can be achieved. In this study, word groups were coded within the text using a classification scheme, and then the number of times each idea was mentioned was counted. The classification scheme was built up through an iterative process based on the grounded theoretic approach [10], which starts with an *a priori* framework (in this case developed by the two principal investigators who were present at all group meetings). The framework was then modified to incorporate new ideas as the analysis proceeded. One of the principal researchers coded all the transcripts, the other then checked all the codings, and any disagreements were resolved through discussion. The unit of analysis was the whole sample and no attempt was made to separate out differences in ideas between individuals.

Ideas mentioned by respondents were divided into three categories: (i) reasons for decisions; (ii) problems with the particular questions used in the study; and (iii) general comments. Reasons for decisions were divided into principles and factors. Principles are generalised rules for priority-setting, such as 'give priority to those who gain the most' or 'give priority to those in most urgent need', which have been partially or fully articulated by the respondent. Factors are specific aspects of the situation, such as the age or health status of a patient, which have not been further articulated into any form of general principle.

For the purpose of counting, it was necessary to take various decisions about what does and does not count as a ‘mention’ of a reason. In general, a conservative view was taken: a word group was never given more than one coding; if the same person repeated the same idea this did not count as a second mention if it occurred during the same speech (although it did count as a second mention if someone else interjected in the meantime); and mentioning both general and specific versions of the same idea counted as a single mention of the specific idea.

If responses to the context-free questions asked in this study are seen to reflect respondents’ values about the particular equity issues being addressed, then the results will be generalisable across a wide range of different decision contexts. But if the results are being generated more by extraneous factors that the abstract questions have sought to eliminate, then the results cannot be interpreted as meaningful responses to the precise questions being asked. The qualitative data gathered in this study meant that it was possible to distinguish reasons according to whether they were ‘relevant’ or ‘irrelevant’. Distinguishing relevant from irrelevant in this way does not presuppose any particular theory (e.g. standard economic theory), although it does presuppose the researcher’s own particular interpretation of the question put to respondents.

4. Results

Of the 72 people invited to the meetings, 60 (83%) attended. Table 2 gives a breakdown of the characteristics of respondents and shows in very general terms that the sample was relatively heterogeneous. The sample was slightly older than would have been expected if it had been truly random (in which case, there would have been approximately one-third of respondents in each of the three age groups) but the proportion of people leaving school at the minimum age was similar to the proportion of the UK population with no qualifications (although, of course, these two figures are not directly comparable) [11]. In addition, the percentage of respondents consulting their GP in the previous year was also broadly similar to what would have been expected from a representative sample [12].

The final classification scheme for the qualitative data contains five principles and five factors. These are shown in Tables 3 and 4 together with the number of times each principle or factor was mentioned. The results from all ten groups were pooled, rather than analysing the social decision-maker and veil of ignorance sub-groups separately. This is because the perspective respondents were asked to adopt was hardly ever articulated by group members, and had no discernible impact on the discussions or the choices that respondents made. Of course, not mentioning something does not necessarily mean that people were not thinking about it. However, the overwhelming impression was that people tended to adopt a simple and rather detached decision-making perspective, with little attention to any complicating details, for example, about real-life decision-making roles.

Table 3 shows that respondents were not willing straightforwardly to give higher priority to patient groups who will gain more health from treatment. The main

Table 2
Respondent characteristics ($n = 60$)

Characteristic	Category	<i>N</i>
Gender	Male	32
	Female	28
Age (years)	18–34	16
	35–54	21
	55–70	23
Annual household income	<£15 000	25
	>£15 000	35
Smoking status	Smoker	15
	Non-smoker	45
School leaving age	Minimum	26
	Stayed on	34
Political allegiance	Conservative	8
	Labour	25
	Neither/not saying	27
Private health insurance	Insured	10
	Uninsured	50
Visits to doctor in last year	None	20
	One or more	40

reason for this appears to be a principled ethical objection to the idea of discriminating against people on the basis of capacity to benefit. This was phrased in a number of ways: as the idea that a life is a life, everyone is equal, that we should not play God, or that we should not pass judgements on other people's lives. Some respondents contrasted the 'investment' or 'logical' point of view, which tries to gain as much health as possible from limited resources, with the 'humane' or 'moral' point of view, which tries to give everyone an equal chance of being treated. Others mentioned the idea that giving the same priority to everyone gives everyone a chance or a hope of being treated. Still others voiced the idea that giving the same priority to everyone is a way of avoiding hard decisions.

However, respondents were not totally unwilling to give higher priority to patient groups who will gain more health from treatment. They were generally willing to choose between groups once the difference in end-points brought about by differential health gains went beyond a certain threshold. The most common idea was that there comes a point at which the benefit provided does not enable the recipient to have a meaningful quality or length of life (where 'meaningful' in both contexts varied from person to person). Some respondents mentioned an absolute threshold i.e. they were thinking about whether the end-point of one group was meaningful or not independently of the end-point of the other group. Fewer respondents mentioned a relative threshold i.e. they were comparing the two groups to find the point at which the relative difference in end-points was sufficiently large.

Table 3
Principles mentioned

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

A very different kind of threshold was used by those respondents who were initially willing to give priority to the group that gained the most health. These respondents generally felt that there comes a point at which the (absolute or relative) difference between the two end-points is sufficiently small for both end-points to be valued equally.

It can be seen from Table 4 that the factors that people mentioned were largely ‘irrelevant’ ones, which the questions were deliberately designed to set aside and which the moderator repeatedly asked not to be taken into account. Respondents were especially likely to mention: (i) the age of patients; (ii) possible additional costs of ongoing treatment (nine mentions in seven groups); and (iii) possible additional health benefits. The most popular ‘additional health gain’ was the idea that people may live longer due to medical advancement. This was an ‘irrelevant’ consideration, since respondents were explicitly asked to set aside all uncertainties of this kind.

Table 5 presents the results of individual responses to the initial choice questions (although respondents were asked to consider where they would ‘draw the line’ between the two groups, this was to get respondents to articulate their views and therefore was principally designed to gather qualitative rather than quantitative data). This choice data reinforces the qualitative data. That is, many people were not immediately willing to give higher priority to the group that can gain most from treatment but were willing to do this when the difference between the end-points

Table 4
Factors mentioned

Factor	Number of mentions in:	
	Total	Group
<i>1. Gains in terms of:</i>		
a. length of life	1	1
b. quality of life	9	7
c. health in general	1	1
<i>2. Patients with:</i>		
a. poor health from now	2	1
b. disability in general	2	2
<i>3. Age:</i>		
a. children	2	2
b. the elderly	4	3
c. age in general	16	8
<i>4. Additional benefits:</i>		
a. costs of ongoing treatment	9	7
b. tax revenue from earnings	1	1
c. long-term gains from research	2	1
<i>5. Additional health benefits:</i>		
a. length of life	7	4
b. quality of life	7	6
c. health in general	4	4

Table 5
Individual results

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

^a See Table 1 for the question parameters.

^b Three respondents (5%) gave the same priority to both groups, irrespective of the difference in benefits they could receive, in all questions.

^c For these two questions, giving the same priority to both groups in the initial choice means that the question is finished.

(and hence benefit) of the groups became sufficiently large. In each of the four questions where the initial benefit to the two groups is different (i.e. questions 1, 2, 4 and 5), about 10 people were totally unwilling to discriminate between the groups, even if the benefit to one group became very small (for example, a few hours of life expectancy or a tiny percentage of quality of life). Only three people consistently maintained a rigid principle of equal priority to both groups in every one of these questions, but 20 people (i.e. one-third of the sample) were totally unwilling to discriminate between the two groups in at least one question.

That these 20 respondents were placed in five of the 10 groups suggests that the group process had some ‘group consensus effect’ on responses. In addition, two groups had a significantly higher proportion of initial ‘same priority’ responses and three groups had a significantly higher proportion of ‘priority to group B’ responses than would have been expected from the overall results presented in Table 5 ($\chi^2 > 6.63$, 1 d.f.). In one group, the initial choice of all respondents in every question was to give the same priority to both groups whereas in another group 28 out of 30 initial choices was to give priority to group B. However, the choices that respondents made were unrelated (according the χ^2 test) to the background variables listed in Table 1.

5. Discussion

Respondents in this study were asked to consider whether and how they would choose between two groups of patients who differ only in terms of their respective gains in length or quality of life. Many respondents appeared to start with a horizontal equity concern for treating everybody equally. For some, this was due to an unwillingness to use capacity to benefit as the only criterion by which to discriminate between the groups. Others appeared to be more concerned with procedural principles (i.e. a fair process of decision-making) than they were with distributive principles (i.e. a fair outcome) [13,14]. This was most often expressed by the idea that fairness requires everyone to have an equal chance of treatment, irrespective of the benefits they could derive from it.

Such findings lend support to one of the founding principles of the British National Health Service; namely, that there should be equal access to health care for all [15]. However, given the way in which the questions were formulated, the ‘default position’ would have been to treat both groups equally rather than to give priority to the group with the greatest capacity to benefit. Therefore, the results could partially be explained by people’s well-established susceptibility to ‘omission’ bias, whereby the disutility associated with an action is greater than that associated with inaction [16]. There was some evidence of this bias in that some respondents appeared to be reluctant to choose between the groups precisely because of the anxiety associated with doing so, and instead chose to ‘sit on the fence’.

However, although many respondents initially wanted to treat both groups equally, most were willing to trade this principle off for the vertical equity consideration of giving greater priority to those who gain most from treatment

when the size of the gain to one group was no longer considered large enough. Most based their decision on the end-point levels of health that the groups end up with after treatment, rather than on the gain in health provided by treatment (although, of course, the end-point and the benefit is the same in those questions where patients will die without treatment).

In the heated debate about whether equity in health care should be defined according to access or according to outcome [17,18], the message from many of the respondents in this study would seem to be something like this: 'let's maximise equality of access subject to the outcome constraint that treatments are sufficiently effective'. It is likely that this rule would be intuitively appealing to many people. If this is true, an important question for future research is 'how effective do treatments have to be for the principle of equal access to apply?' It would appear that this represents a more fruitful way of addressing the issue of equity in health care than continuing to have a rather polarised debate whether equality is defined solely in terms of access or solely in terms of outcome.

6. Concluding remarks

Overall, reaction from the public to the focus group methodology was favourable. The general feeling was that group discussions of this kind are worthwhile; but only if they end up having some impact on actual decisions. Only one respondent was strongly critical of the whole enterprise, expressing the view that the debate should focus on getting more funding for the NHS rather than on rationing. And although some respondents initially had difficulty with the questions discussed in this paper, all felt able to answer them after further explanation and clarification from the moderator.

However, some respondents clearly felt that the questions were too abstract and made assumptions which were hard to believe (for example, that the age profiles of the two groups were identical). Even when respondents were explicitly told to ignore certain factors that would be relevant to a real-world decision, they were often unable or unwilling to do so. Equally, respondents could not be 'forced' to adopt a strict social decision-making or veil of ignorance perspective; they all tended to adopt a simple decision-maker perspective.

As noted above, the questions, like those in many other studies, were designed to abstract from many real-world decision-making contexts. The qualitative results from this study have important implications for the interpretation of the results from surveys which ask tightly defined questions of a similar kind, but where qualitative evidence on the reasons for responses is not gathered. In many cases, it would seem that the results may have very little to do with the factors that researchers thought were responsible for them, and much more to do with extraneous factors that researchers were not aware of.

Of course, there is always the danger that some respondent's 'true' preferences might be changed or shaped by more articulate or forceful group members. The fact that there was evidence of a 'group consensus effect' suggests that this could

have happened in this study. However, there was no clear evidence from the qualitative data that particular respondents with particular background characteristics were responsible for the greater (but by no means complete) consensus that emerged at the end of the discussions. Nor was there any evidence that the group effect was due to unthinking ‘herd’ behaviour rather than reasoned agreement.

A group environment may also have made some respondents more reluctant to discuss issues of discrimination on the grounds of capacity to benefit, since to do so may have made them feel ‘calculating’ in the eyes of their peers. It is difficult to mitigate against this effect except to say that respondents were encouraged to articulate their views irrespective of the extent to which these views sounded ‘silly’ or conformed with those of others.

Since it is recognised that individual preferences rarely come well-articulated (particularly regarding such complex and unfamiliar issues as priority-setting in health care) [19,20], there is increasing evidence that more accurate representations of these preferences emerge from discussion and debate than from without it [21,22]. Moreover, the methodology employed in this study allowed the implications of particular choices to be made clear to respondents, who could then indicate whether or not they were happy with those implications. All of this suggests that discussion groups would appear to represent a promising way of eliciting the views of the general public regarding distributional issues in health care.

References

- [1] Culyer A. The primary goal of the health care system should be to maximise health. In: New B, editor. *Rationing: Talk and Action in Health Care*. London: King’s Fund, 1998.
- [2] Harris J. What the primary goal of the health care system should really be. In: New B, editor. *Rationing: Talk and Action in Health Care*. London: King’s Fund, 1998.
- [3] Nord E. The trade-off between severity of illness and treatment effect in cost-value analysis of health care. *Health Policy* 1993;24:227–38.
- [4] Nord E, Richardson J, et al. Maximizing health benefits vs. egalitarianism — an Australian survey of health issues. *Social Sci Med* 1995;41(10):1429–37.
- [5] Johannesson M, Johannesson PO. Is the valuation of a QALY gained independent of age? Some empirical evidence. *J Health Econ* 1997;16:589–99.
- [6] Johannesson M, Gerdtham UG. A note on the estimation of the equity-efficiency trade-off for QALYs. *J Health Econ* 1996;15:359–68.
- [7] Dubourg WR, Jones-Lee MW, Loomes G. Imprecise preferences and survey design in contingent valuation. *Economica* 1998;64(256):681–702.
- [8] Baron J, Greene J. Determinants of insensitivity to quantity in valuation of public goods-contribution, warm glow, budget constraints, availability and prominence. *J Exp Psychol Appl* 1996;2(2):107–25.
- [9] Krueger RA. The future of focus groups. *Qual Health Res* 1995;5(4):524–30.
- [10] Patton M. *Qualitative Evaluation and Research Methods*. 1990.
- [11] OPCS, 1990. *Key Population and vital statistics, Series VS No. 17, PP1 No.13*, HMSO, London.
- [12] Carr-Hill R, Rice N, Roland M. Socio-economic determinants of rate of consultation in general practice based on fourth national morbidity survey of general practices. *Br Med J* 1996;312:1008–13.
- [13] Hughes T, Larson L. Patient involvement in health care: a procedural justice viewpoint. *Med Care* 1991;29(3):29.

- [14] Mooney G, Jan S. Vertical equity: weighing outcomes? or establishing procedures? *Health Pol* 1997;39:79–87.
- [15] Beveridge, W. *Social Insurance and Allied Services*, HMSO, London, cmd. 6404, 1942.
- [16] Ritov I, Baron J. Status quo and omission biases. *J Risk Uncertainty* 1992;5:49–61.
- [17] Mooney G, Hall J, Donaldson C, Gerard K. Utilisation as a measure of equity: weighting heat? *JHealth Econ* 1991;10:475–80.
- [18] Culyer A, Wagstaff A. Equity and equality in health and health care. *J Health Econ* 1993;12:431–57.
- [19] Slovic P. The construction of preferences. *Am Psychol* 1995;50(5):364–71.
- [20] Fischhoff B. Value elicitation: is there anything there? *Am Psychol* 1991;46:835–47.
- [21] Loomes G. *Valuing health and safety: some economic and psychological issues*, mimeo. University of York, York, 1994.
- [22] Dolan P, Cookson R, Ferguson B. Effect of discussion and deliberation on the public's views of priority setting in health care: a focus group study. *B Med J* 1999;318:916–9.