

Research article

The equivalence of numbers: The social value of avoiding health decline: An experimental web-based study

David LB Schwappach

Address: Department of Health Policy and Management, Faculty of Medicine University Witten/Herdecke, Witten, Germany

E-mail: dlb@healthecon.de

Published: 5 March 2002

Received: 11 November 2001

BMC Medical Informatics and Decision Making 2002, 2:3

Accepted: 5 March 2002

This article is available from: <http://www.biomedcentral.com/1472-6947/2/3>

© 2002 Schwappach; licensee BioMed Central Ltd. Verbatim copying and redistribution of this article are permitted in any medium for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: Health economic analysis aimed at informing policy makers and supporting resource allocation decisions has to evaluate not only improvements in health but also avoided decline. Little is known however, whether the "direction" in which changes in health are experienced is important for the public in prioritizing among patients. This experimental study investigates the social value people place on avoiding (further) health decline when directly compared to curative treatments in resource allocation decisions.

Methods: 127 individuals completed an interactive survey that was published in the World Wide Web. They were confronted with a standard gamble (SG) and three person trade-off tasks, either comparing improvements in health (PTO-Up), avoided decline (PTO-Down), or both, contrasting health changes of equal magnitude differing in the direction in which they are experienced (PTO-WAD). Finally, a direct priority ranking of various interventions was obtained.

Results: Participants strongly prioritized improving patients' health rather than avoiding decline. The mean substitution rate between health improvements and avoided decline (WAD) ranged between 0.47 and 0.64 dependent on the intervention. Weighting PTO values according to the direction in which changes in health are experienced improved their accuracy in predicting a direct prioritization ranking. Health state utilities obtained by the standard gamble method seem not to reflect social values in resource allocation contexts.

Conclusion: Results suggest that the utility of being cured of a given health state might not be a good approximation for the societal value of avoiding this health state, especially in cases of competition between preventive and curative interventions.

Background

In the era of aging populations and rising incidence of progressive diseases in industrialized countries, health politicians in these countries are increasingly faced with the decision what fraction of their budgets to spend on treatment and cure or on avoiding (further) health decline. Therefore, health economic analysis aimed at informing policy makers and supporting resource allocation

decisions has to evaluate not only improvements in health but also prevented decline. The benefit of these "upward-movements" on a health scale compared to avoided "downward-movements" is often calculated as the potential gain in QALYs ("quality-adjusted life years"). Traditionally the standard gamble (SG) or the time trade-off (TTO) method is used to determine the quality of life weights needed to estimate the benefit of

healthcare interventions in terms of QALYs. The use of these techniques in resource allocation decisions has been criticized as of two concerns, among others [1,2]. First, because the benefit of an intervention is calculated as the difference in preferences for two health states initially valued *isolated* from each other [3,4]. The second concern is that the use of SG and TTO in policy decisions assumes that preferences an individual has regarding his own (hypothetical) health equal the preferences this individual has for the health of others, therefore being unaffected by distributional concerns [5,6]. Consequently, the person trade-off (PTO) method has been proposed because it forces respondents to take a social rather than a personal perspective in evaluating rationing frameworks [7]. As in the PTO *movements* on a health scale rather than distinct health *states* are presented, and it thus explicitly focuses on changes in health, the PTO seems to capture some of the factors that researchers have found to be important for the public in distributive decisions. Two of these are the severity of the initial health state and the potential for health [8–10]. The former describes the tendency to give priority to the worst off even if the associated health gain is smaller. The latter describes the aversion to discriminate against those with a limited potential for health as of chronic conditions or disabilities.

This experimental study addresses a different factor influencing distributional concerns of the public, that is, the *direction* of a change in health. In health economic evaluative studies the benefit of avoiding a health state is traditionally calculated exactly as the inverse equivalent of illness, and therefore – on a conceptual level – as the same as cure or treatment. This may be of special relevance to the evaluation of prevention programs, where benefit is derived from reducing the likelihood of a specific health state thought to occur in the future. However, despite the effect of the direction of changes in health itself, i.e., avoided losses versus improvements, 'prevention', as the term is commonly being used, also involves preferences towards time and uncertainty. This study focuses only on the equivalence of equally sized health effects achieved either by curative or preventive treatments, ignoring risk aversion and discounting of time effects. Imagine, for example, illness X with an associated health state x (which is worse than the top anchor 'perfect health'). If there is a cure for X that will return diseased persons to perfect health, the benefit of cure is $1-x$, assuming the conventional 0 to 1 utility scale and not taking the duration of illness into account. In case there is also the opportunity to avert X in people healthy at present, the benefit of avoiding health state x is also given by $1-x$. The value people attached to a move from a health state worse than perfect health upwards is used interchangeably with the value of avoiding a move from perfect health downwards this health state. The underlying hypothesis is that the *distance*

between two health states is the main carrier of value irrespective of the direction of the movement and the mode by which it is achieved. This common method of benefit estimation has been applied to various prophylactic treatments and cost-per-QALY-estimates of avoiding health decline are compared not only to treatment of the disease under study but also to treatments of other diseases in cost-utility-ratio league tables [11].

However, there is very limited evidence on the value people attach to treatment and avoided decline when involved gains and losses are directly compared to each other in resource allocation contexts. If it was the case that the distance between two health states is perceived as being of equal value no matter whether this distance is prevented downwards or gained upwards, one would expect that people reveal answers close to "100" when faced with a PTO question like:

"In how many people a decline to health state X had to be prevented to make you indifferent to treating 100 persons in health state X and returning them to complete health"?

The available evidence and a number of rationales cast doubts on the hypothesis that curative and protective interventions would be valued as equivalents. Ubel et al. surveyed 289 prospective jurors in the US how to allocate funds on different healthcare projects for nursing home residents with varying levels of disability [12]. Healthcare activities were described as having the same magnitude of benefit either by improving the level of functioning or preventing further decline. While there was a general attitude towards prevention, when the strength of preference, measured on a five-point scale, was taken into account, there was no statistically significant preference in favor of prevention. In a survey conducted among the Swedish general public, participants were asked to trade-off the value of different programs saving varying numbers of lives either by acute or by preventive care [13]. Again, responders favored prevention over acute care with a mean of 1.2–1.4 lives saved by acute care being judged equivalent to 1 life saved by preventive care. Overall, there seems to be, if at all, a slight trend to prioritize towards preventive health care. Both surveys presented choices among certain outcomes close in time, thereby focusing mode of intervention and the direction, in which changes in health are achieved.

There are a number of rationales for people expressing differing values for cure and avoidance of decline. As Johannesson and Johansson point out, giving priority to those in a state worse than perfect health already suffering from disease, may be explained by equity concerns [13]: Refusing treatment to individuals in a more severe initial health state and prioritizing those fortune to be healthy might be

judged as unfair and people might therefore place more value on cure. Instead, one could argue that persons healthy at the moment experience a *loss* when they fall sick, while those presently suffering already have gone through that loss. The disutility of the process of loss itself might be considered more important than the process of gain itself, i.e., recurring to health. One could see parallels to the observation that patients often attach higher values to the health state they are experiencing than the general public [14]. If subjects in impersonal allocation tasks would simultaneously adopt both perspectives, this could explain that a greater value would be attached to avoiding decline as compared to treatment since healthy subjects obviously "lose more" than patients win. In addition, subjects saved from developing a disease or decline might be expected to live longer or in overall better quality of life, and interventions for avoidance of disease might be perceived as less harmful or more effective.

There are also a number of psychological explanations why people would prioritize preventing decline compared to cure, for example, as of an impersonal equivalent to "status quo"- and "endowment"-effects [15]. Simply speaking, opting for protective treatments would keep the current situation unchanged for both the ill and the healthy. The main purpose of this paper is to investigate the value people place on equidistant health changes, either achieved by preventive or curative care, when directly compared in resource allocation decisions. A valuation method that captures this relationship best is to be identified. In addition to the common PTO technique, two new PTO frames will be applied which explicitly ask for the value of avoided health decline. A new idea for the PTO will also be introduced, which is intended to combine preferences for positive, and avoided negative health changes. This will be called the "weight for avoided decline" (WAD). As "the coefficient of loss aversion" in Prospect Theory, the WAD is interpreted as the "currency" by which effects in outcome are weighted and transformed when perceived as either a loss or a gain [16]. This conversion factor will be used to incorporate people's preferences in allocating resources among diseased patients awaiting treatment and healthy persons that could be saved from decline. Finally, the accuracy of the different valuation techniques will be assessed by comparing a prioritization ranking of health interventions directly obtained in a hypothetical allocation scenario with rankings predicted by the different valuation techniques.

A second purpose of the study is to test whether utilities, elicited by the traditional standard gamble technique (SG) can be used to appropriately reflect social valuations of movements between health states and the direction in which these are experienced, as assumed in common health economic practice.

Methods

Health states and scaling

Six health states labeled A-F (with increasing severity) were prepared, which are described in table 3. In a pre-test with 25 students from the University of Witten/Herdecke, detailed descriptions of health states covering various symptoms were used. Extreme variance in the subjective ranking of health states was observed, e.g. between sexes. Therefore, it was decided to use very simple and mostly additive health problems implying an obvious ranking of health states. The worst health state (F) had to be different from the conventional absorbing "death"-state. As values for the same distance between two health states had to be elicited either experienced as an improvement or as avoided decline, all health states had to be constructed so that hypothetical patients could both move into and out of each health state. Therefore, an irreversible health state could not be used. In addition to this methodological reason, there is a deeper rationale: Because life and death differ in their absolute meaning, preferences towards health care resource allocation involving both categories are likely to follow and mix distinctive ethical rules as if only one category was concerned. Though health state F might be perceived as worse than death by respondents, it differs from the traditional "dead"-state bottom anchor in its qualitative category, since "worse than death" describes a point on the quality-of-life continuum, different to the irreversible, "unmodifiable" and final character of death. For example, the rule of rescue in its narrower sense, namely the societal duty to save lives in fatal situations, applies per se only in cases where the avoidable death of subjects is involved.

Readers should note this special scaling using a reversible lower end point because it makes the comparison between the PTO and SG scores reported herein with those found using the conventional scale (with death as the bottom anchor) inappropriate.

Questionnaire instrument and valuation procedures

An interactive survey was developed which was published in the World Wide Web. The survey was written in HTML and JavaScript. The questionnaire consisted of seven sections. They are described in some detail to make the reader familiar with the analytical steps and valuation procedures:

- 1) The first part of the survey introduced all health states to respondents. Subjects were asked to rank health states according to their severity. This was intended to make them familiar with all health states.
- 2) The conventional standard gamble method (SG) was presented for health states B to E after explaining the task. Instead of the usual definition of risk as death from treat-

Table 1: Health states as described in the questionnaire

Health state	Description
A)	People in this health state are in complete health.
B)	People in this health state have problems with the knee joints.
C)	People in this health state have problems with the knee joints and suffer from Asthma.
D)	People in this health state have problems with the knee joints, suffer from Asthma and are partially sighted.
E)	People in this health state suffer from Asthma, are weak sighted and need a wheelchair.
F)	People in this health state are completely dependent on others. They suffer extreme pain and are unconscious at times.

ment, risk in this study was defined as the probability of moving to health state F. The standard gamble was used as of two reasons: First, it was expected that the assessment of health states from a personal perspective would increase reflection and seriousness of participants. Second, results of the SG task were used to investigate the difference between personal (SG), and social (PTO) evaluation of health states. Following common health economic practice, SG utilities were also used to estimate the value of avoiding a health state by taking their inverse on the utility scale. For example, the benefit of avoiding a move from health state B downwards to health state C would be calculated as

$$U(B \rightarrow C) = U(B) - U(C) \quad (1.1)$$

where $U(B \rightarrow C)$ is the utility of not moving from B to C and $U(B)$ is the utility of being in health state B. Subsequently, the accordance of these values with those elicited by the other techniques was examined.

3) After completion of the four SG questions, respondents were asked to imagine that they were health authorities whose task it is to decide which healthcare interventions to offer to the public. They were explicitly told that their budget was not sufficient to offer every treatment that people could benefit from and that they consequently have to compare and evaluate the benefit of the different options. It was stated that patients were equal in all other respects except the described health conditions. The section started with a PTO like choice that was only intended to make subjects familiar with the task and avoid start point bias [17]. These results were not used in data analysis. The introduction was followed by three PTO sections: The first contained four pair-wise comparisons between curative treatments. In the following, this is called PTO-UP, because choices had to be made between two treatments that offer *improvements* in health. All patients were portrayed as suffering from the respective disease and

their health status would not change unless they were provided the treatment, which would certainly return them to complete health (state A). The frame used was originally presented by Pinto Prades [18]: Starting with the PTO task applied to the worst health state compared to the next best health state, subsequent comparisons were made between a health state and the bordering, next best health state. As illustrated in Figure 1, first the improvement from health state F to A (arrow 1) was compared to the improvement from E to A (arrow 2), followed by the comparison between the gain from health state E to A and D to A (arrow 3) and so on. The advantage of this incremental approach is that it might be cognitively and morally easier to handle health states that are not too different from each other. However, within this approach, additivity of benefits is assumed. Again, to calculate the value of not moving from B to C, first the value of moving from B to A as compared to a move from C to A was assessed, which is given by

$$100 U(C \rightarrow A) = x U(B \rightarrow A) \quad (2.1)$$

(The value of $U(C \rightarrow A)$ is known from the previous comparisons.) In the next step, the results of (2.1) were then used to calculate

$$U(B \rightarrow C) = U(C \rightarrow A) - U(B \rightarrow A) \quad (2.2)$$

4) The next section of the survey continued with the same procedure except that now four PTO choices between protective interventions for people healthy at present were offered. Respondents were told that people would certainly fall ill in the near future, if they did not receive the preventive treatment. That is, two avoided downward-movements were presented. For example, first an avoided decline from health state A to health state F (Fig. 1, arrow 4) was compared to an avoided change from health state A to health state E (Fig. 1, arrow 5), and this was continued with adjacent health states. This method is called PTO-DOWN. The same eliciting procedure as in PTO-UP was used. The benefit of avoiding a move from B to C was

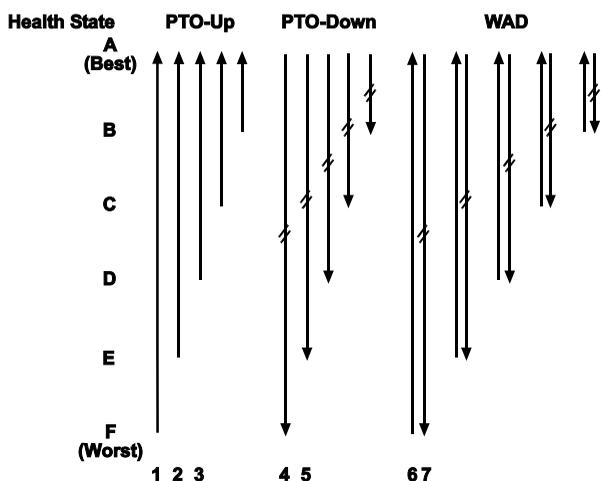


Figure 1
Illustration of the compared health state movements. Crossed arrows indicate avoided decline.

calculated as the difference between an avoided move from A to C and A to B, namely

$$U(B \rightarrow C) = U(A \rightarrow C) - U(A \rightarrow B) \quad (3.1)$$

5) Subsequently, five PTO choices between curative treatments for diseased persons and treatments saving persons from (further) health decline were presented. While the distance on the health scale was equal in both options, they differed in the direction of the potential health effect. For example, a choice between a treatment which would return 100 persons in health state F to complete health was compared with the opportunity to protect X persons in complete health (state A) from entering health state F (Fig. 1, arrows 6 and 7). The purpose of this procedure was to assess directly the rate at which health gains and avoided losses are traded against each other. This "weight for avoided decline" (WAD) for the distance between health states B and C is defined as:

$$WAD_{B,C} = \frac{100U(C \rightarrow B)}{xU(B \rightarrow C)} \quad (4.1)$$

If, for example, participants indicated by their answers, that they were indifferent between improving the health of 100 patients from health state C to B and avoiding decline from B to C in 80 persons, then $WAD_{B,C} = 1.25$. The size of the coefficient and whether it is constant, i.e., independent of the reference level, is the focus of interest. While a WAD higher than 1 indicates that people value

avoiding decline over health improvements, a WAD lower than 1 implies that respondents prefer the latter to the former. With a WAD equal to 1, curative and protective treatments carry the same value, as indirectly assumed when taking the inverse value of health improvements as the value of avoided decline. While the size of WAD reveals whether individuals favor avoiding decline over health improving treatments or the opposite, it cannot be used directly for the comparison of interventions differing in both, direction and size of a health effect. Whereas an individual's WAD implies, e.g., preferences for prevention over cure, these preferences need also be related to health effects of different magnitude. That is, preferences for health improvements or avoided decline need to be traded against their relative size. Such information would be needed, for example, if a decision maker faces a decision on whether to fund curative care returning patients in HS C to complete health or to fund preventive care that avoids decline to HS E in patients currently in HS B. This makes it necessary to accommodate individuals' preferences towards curative and preventive effects, measured in terms of WAD, as a special factor in the conventional measurement of value. One way to achieve this and make values for health improvements and avoided decline compatible, i.e. quantified on the same scale, is to adjust PTO-values for the direction in which health changes are experienced. Since the measurement of value attached to health improvements (PTO-UP) is more common than that of avoided health decline (PTO-Down) it was decided to use PTO-UP values as the "baseline", subject to adjustment. The individual-specific mean WAD was therefore applied as a weight to PTO-UP values. These weighted PTO-UP values may then be attached to *avoided health changes*, while the "raw" PTO-UP values would still be attached to health improvements. For example, the value of avoiding a move from health state A to C was calculated as the value of PTO-UP for the move from C to A, weighted by the individual-specific mean WAD. These combined values are indexed "PTO-WAD". To calibrate the resulting adjusted PTO-values to the 0-1 range, the PTO-UP values were weighted using a power transformation.

$$PTO - WAD_{A \rightarrow C} = PTO - Up_{C \rightarrow A}^{\frac{1}{WAD_{CA,i}}} \quad (4.2)$$

Thus, a WAD higher than 1 increases the social value of an avoided health change compared to the "raw" PTO value for curative treatment effects. If for example, subject's responses to the PTO-Up procedure reveal that the distance between health states C and A gained upwards is 0.4, and this individual's mean WAD is 1.5, which implies a preference towards preventive care, avoiding a decline from A to C would be calculated as 0.54. Whether the accuracy of

the PTO-Up values in predicting preferences for prioritization among curative and preventive interventions can be improved by weighting them with the WAD, will be analyzed in the next step.

6) The sixth section presented the direct allocation task to the subjects. From all 30 possible health changes on the scale (in both directions), five were randomly chosen from those *not presented in the previous three PTO tasks*. These were three curative movements (1) $C \rightarrow B$, (2) $E \rightarrow C$, (3) $F \rightarrow B$, and two treatments saving from decline (4) $B \nrightarrow D$, (5) $C \nrightarrow F$. Subjects were told that they had a finite budget for the next year as health authorities, but that they did not yet know how sufficient it was and which of the presented interventions they could offer to the public. They were then asked to rank the interventions according to their preferences in giving priority to patient groups. Results were interpreted such that the intervention ranked first has the highest social value for respondents at an *ordinal* level. While the absolute size of the intervals between health states represented by the five scenarios cannot be derived from this ranking, several statements on their relative size can be made. For example, one would expect (3) $F \rightarrow B$ to be ranked higher than (1) $C \rightarrow B$ and (2) $E \rightarrow C$. Consequently, if irrelevance of the direction of health change on preferences is assumed, higher priority would be attached to curative treatment (3) $F \rightarrow B$ as compared to avoiding decline in (5) $C \nrightarrow F$ or (4) $B \nrightarrow D$, while the improvement (1) $C \rightarrow B$ would rank lower than avoiding the decline in (4) $B \nrightarrow D$, and (2) $E \rightarrow C$ would have lower priority than (5) $C \nrightarrow F$. To analyze the assumption of equidistances, the ranking obtained from each individual was investigated for the number of preference reversals between curative and preventive treatments. Using the rules on relative size given above, a reversal was defined as either an avoided decline that was ranked higher than an improvement of a greater change (reversal in favor of avoiding decline) or as an improvement that was attached higher priority than an

avoided decline of larger size (reversal in favor of curative treatment).

The ordering obtained by the direct ranking task was then compared with the orderings implied by individuals' responses to the four valuation techniques (PTO-UP, PTO-Down, PTO-WAD, SG). On the presumption that the direct ranking task reflects individuals' preferences to prioritize among the five scenarios best, the number of 'errors' found in each of the implied rankings was calculated. An 'erroneous' rank was defined as a mismatch between the rank a scenario obtained in the direct prioritization task as compared to the rank the same scenario obtained predicted by the particular valuation technique. The number of potential errors ranges from 2, in case the ranks of two scenarios are reversed in the prediction, to a maximum of 5 in case all ranks are predicted incorrectly. However, not only the number of errors but also their degree is relevant. For example, a reversal between two adjacent ranks is less serious than a reversal between the first and the fourth rank. Therefore, the degree of incorrect prediction measured as the absolute sum of deviation in ranks between direct prioritization ranking and predicted ranking, averaged over all individuals, was also analyzed. The potential absolute sum of deviations in ranks ranges from 2, in case the ranks of two scenarios are adjacently reversed, to 12, reflecting the maximum disagreement in orderings.

7) Demographic characteristics such as age, gender and employment status were asked in the last section of the questionnaire. In addition, subjects were asked to rate how difficult they found it to answer the survey's questions. They were also encouraged to give qualitative comments. A JavaScript function was programmed that calculated the time respondents needed to complete the survey starting. The recording of time started with the first standard gamble task (with the first 'click' performed by responders) and stopped after the provision of demographic information (before final qualitative comments).

Table 2: Examples of search algorithm used in the PTO-tasks

Choice		Vote	2. Choice		Vote	3. Choice		Vote	4. Choice		Vote	PTO calculation
A	B		A	B		A	B		A	B		
100	100	A	100	110	A	100	130	A	100	170	Indiff.	100 A = 170 B
100	100	B	100	90	B	100	70	A	100	80	A	100 A = 85 B
100	500*	A	100	700	B	100	600	B	100	550	Indiff.	100 A = 550 B

Note: Each row shows a series of hypothetical choices between two interventions A and B that could be offered to the specified number of patients each, and responses to the choices. * Arrived at after having consequently voted for A in the preceding choices

The interactive nature of the survey allowed to program search algorithms in all PTO tasks that were different from the standard PTO design. Usually, people are asked to give a number *themselves* at which they are indifferent between two treatment options. As reported by others, the pre-test of this study showed that subjects had difficulties deciding on and matching precise numbers representing equal societal value to them [19]. Therefore, it was decided to *offer* numbers to respondents, similar to binary choices as in bidding game techniques. The elicitation procedure started with a scenario where both interventions could be offered to 100 patients each. Subjects then had to decide for which of the two treatments they would opt or whether they were indifferent. Had participants chosen one of the treatments, the response indicated the direction of inequality, i.e. which intervention was associated with higher value. Subsequently, a higher number of patients benefiting from the less favored treatment was presented. After each vote by which the participant indicated an unaffected direction of inequality, the number offered next was exponentially increased. The algorithm then identified the first choice where responders changed their preference for one of the alternatives, e.g. ticked the second treatment after they had chosen the first previously. The numbers in the range between this threshold and the preceding number (covering the point of indifference) were presented in a ping-pong style, alternating convergent numbers between high and low depending on the voting behavior. The task proceeded until subjects were indifferent, when the next pair-wise comparison was started. Examples of series of hypothetical choices and responses are given in table 3. The whole elicitation procedure was facilitated by various visual aids. For example, in the SG section, a balance was shown that changed its appearance depending on the risk the respondent was willing to take (Figure 2). Health states were explained by written text and small animations representing the information. In the subsequent sections, these animations were used to represent health states and to avoid redundant text and information overload. Written information on health states and the tasks was always accessible by hyperlinks that opened pop-up windows restating the information.

Subjects

A self-selected sample was recruited over the Internet by indexing the survey's web page in search engines, announcing it on survey meta-listings and by postings in German language Usenet groups, such as de.sci.*, de.soc.* and de.talk.*. Subjects were invited to visit the survey's webpage. As an incentive, it was announced that those taking the questionnaire are offered the chance to win a coupon from an online bookstore worth ~ 10€. By separate transmission of responses and email addresses anonymity was guaranteed. The risk that the same person could provide multiple responses was minimized by

placement of mutually dependent 'cookies'. Cookies are small files that are written to the users' hard disk by the Web server when accessing a web page and can be read and re-written alongside the visit by this server. In the study, information was written, checked, and tracked in multiple cookies, referring to each other, when initially accessing the survey, while proceeding the survey, before and after successful transmission of responses, as well as before and after submission of the email-address. The information was not stored in the usual temporary cookie folder on the users' hard disk. This programming inhibited repetitive restarts of the survey by the same user as well as moving freely through the survey's pages. Multiple submissions of responses were also suppressed. Still, it cannot be ruled out completely that a person experienced in web programming with some energy and time might have been able to manipulate cookies to provide multiple responses.

Data was analyzed with the statistical package STATA v7 [20].

Results

After 15 days, 129 individuals had completed the survey, of which 65% were male. The average age was 35 years, ranging between 19 and 63 years. 87% had tertiary education, predominantly in medicine (32%) or economics (19%), nursing sciences (11%) or psychology (8%). 24% were currently unemployed or retired and 9% were students. The average time needed to complete the survey was 21 minutes (median: 35). Data of two respondents had to be removed, as their answers to the PTO-Up choices were obviously not serious. Both subjects indicated that they were indifferent between a smaller health gain offered to a smaller number of patients and a larger benefit offered to more patients. 9% of respondents stated that they found it not difficult, little difficult (39%) or quite difficult (45%) to answer the questions, and 7% said they found it very difficult. Analyses of web statistics and access-logs show that only few individuals (7) started the questionnaire and dropped out during progress. All subjects provided the same, "correct" and consistent ranking of severity of the health states requested in the first task of the survey.

As shown in figure 3, the differences in mean utility/social value of health states between the three elicitation techniques are considerable. Values obtained with the "prevention"-PTO method (PTO-Down) are generally higher than those assessed by the standard "curative treatment" procedure (PTO-Up). The difference between PTO techniques and SG is substantial and increases with severity of the health state. It was investigated whether the fixed format of the successive bids may have provided incentives to shorten the length of the survey by voting strategically

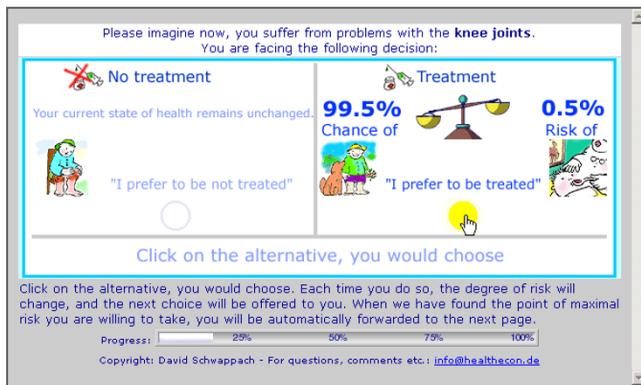


Figure 2
Screenshot of the standard gamble task (translated from the German original)

and avoiding the ping-pong procedure. If this was true, one would expect that the fraction of responders taking the easiest way to do so, namely "opting out" themselves by choosing the "indifference"-option, would increase with successive tasks. However, this was not the case. The number of responders ending a PTO task by clicking the "indifference"-option ranges between 35 (28%) and 46 (36%) per choice, and these figures are not correlated with the sequential placement of the respective PTO-choice among all 13 PTO-choices (*Spearman* $r = -0.09, p = 0.8$). Also, the number of successive bids per PTO-figure does not decrease with the sequential placement of the respective PTO-choice among the 4 resp. 5 PTO-choices of each PTO-task.

The vast majority of responders (69%) preferred the allocation of resources to already diseased patients over avoidance of the same health state in *all* five comparisons. The opposite was true for 2 subjects (2%). 8 responders (6%) gave equal priority to improvements in health and avoiding decline in all comparisons. The remaining subjects (23%) favored cure or avoiding decline depending on the respective health state. There was no significant association between sex, employment-status, education or professional background and preferences towards cure or avoiding decline. However, respondents that voted in favor of cure or avoiding decline dependent on the compared health effects were more likely to find the survey difficult ($\chi^2 = 21.79, p = 0.01$). The mean WAD is nearly constant for all except health state B, for which it is moderately higher than the WAD's for the other health states ($p = 0.000$ compared to WAD for health states C,D,E; $p = 0.05$ compared WAD for health state F). The size of the WAD values implies that to be of equal value to responders, a health decline has to be avoided in twice as many persons than persons have to experience a health gain of the same magnitude.

The mean intervals between health states calculated based on the values elicited by the different techniques and the substitution rate between avoided downward and gained upward movements (WAD) are presented in table 3. All differences between SG and the two PTO techniques are statistically significant at the 5% level or better. Notably, differences in size of intervals between health states inferred from the two PTO-tasks increase with severity of health states.

The number of reversals observed in the direct ranking task is substantial. Calculated over all subjects, 151 reversals occurred in favor of curative treatments, i.e., an improvement was ranked higher than an avoided decline of greater size. In 71 rankings, the treatment that would improve health from C to B was ranked higher than avoiding decline from health state B to D and 80 individuals attached higher priority to the change from E to C compared to preventing the decline from C to F. Both reversals were prevalent in 68 rankings.

9 reversals occurred in favor of prevention. The improvement $F \rightarrow B$ was attached less priority than avoided decline from $C \rightarrow F$ and $B \rightarrow D$ in 5 and 4 cases respectively. Occurrence of both reversals was observed in 3 rankings. Only 3 reversals occurred among curative treatments. In all three cases, the improvement $E \rightarrow C$ ranked better than $F \rightarrow B$.

A comparison between the direct ranking task and the rankings predicted by the four valuation techniques reveals considerable deviation in the predicted rankings on the individual level. Table 4 shows the mean number of

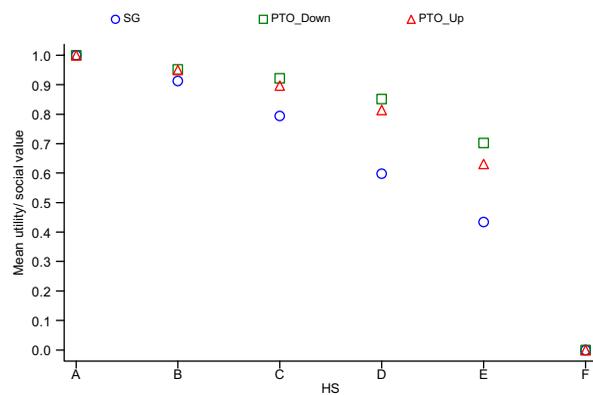


Figure 3
Mean utility / social value of health states by elicitation technique

Table 3: Mean intervals and WAD for movements between health states

Mean interval between health states				
Health states	SG [§] [95% CI]	PTO-Down [°] [95% CI]	PTO-Up [§] [95% CI]	WAD* [95% CI]
A-B	0.0875 [0.0637-0.1113]	0.0480 [0.0223-0.0736]	0.0502 [0.0304-0.0700]	0.6443 [0.5811-0.7075]
A-C	0.2056 [0.1836-0.2276]	0.0778 [0.0488-0.1069]	0.1032 [0.0716-0.1348]	0.5248 [0.4616-0.5873]
A-D	0.4015 [0.3738-0.4292]	0.1484 [0.1111-0.1857]	0.1857 [0.1389-0.2325]	0.4748 [0.4066-0.5431]
A-E	0.5658 [0.5354-0.5961]	0.2977 [0.2487-0.3467]	0.3694 [0.3112-0.4277]	0.4919 [0.3951-0.5888]
A-F	--	--	--	0.5167 [0.3542-0.6791]

N = 127 Results of paired t-tests: [§]Significant different to PTO-Up at 1% or better; [°] Significant different to SG at 1% or better; [§]Significant different to PTO-Down at 1% or better for A - E, at 10% for A - D and A - C; *WAD values were calculated as described in eq. 4.1 (section 5) of the study design

"errors" found in each predicted ranking compared to the direct ranking task. While the PTO-WAD ranking produces on average one reversal among the ranks of two scenarios, the unadjusted rankings implied by participants' answers to the common PTO-UP procedure show on average one additional scenario ranked incorrectly. The mean absolute deviation in ranks between the direct prioritization ranking and the predicted rankings illustrates that the PTO-WAD not only produces the smallest number of errors, but also that these errors are less serious (table 4). The average absolute deviation of two ranks suggests that the PTO-WAD produces one adjacent reversal of two scenarios compared to the direct ranking. The superiority of the PTO-WAD technique compared to the other valuation techniques on the aggregate level can be mainly attributed to the fact that it reproduces a considerable fraction of the reversals observed in the direct ranking task. 49% of the reversals between curative and preventive treatments reported about above were exactly replicated by the PTO-WAD technique.

A considerable fraction of participants provided detailed and serious qualitative comments on the content level. These comments were discussing mainly (1) reasons underlying their choices in favor of curative or preventive healthcare, (2) issues of translating their choices into practice, which was mainly expressed by participants with a professional healthcare background, and (3) general issues of priority setting, health care policy and the German health care system. Surprisingly, several healthcare professionals stated, that they appreciate the presentation of the resource allocation problems and the opportunity to transmit their views anonymously as "...these are the issues we have to face and we should aim to honestly and openly discuss them".

Discussion

This study used an interactive web-based questionnaire to examine preferences towards improvements in health and

avoidance of decline in a healthcare resource allocation context using a self-selected sample. All four preference elicitation techniques gave results of high internal consistency within and between subjects. This is especially surprising for the direct ranking and the PTO-tasks, since they were cognitively demanding. On the methodological level, this may be explained by the mode PTO choices were presented. Responders were confronted with numbers as in bidding games, which might be easier to handle than to match them "freehand". Another explanation is that the health scale used did not include the conventional "death"-state and therefore all interventions were equal with respect to the fact that ethical principles and preferences regarding life-saving treatments, such as the rule of rescue, were not involved. As the values obtained by the new PTO-Down method suggest, the same allocation preferences seem to underlie rationing choices regarding preventive interventions and curative treatments each evaluated among each other. Health effects of greater magnitude, gained upwards or avoided downwards, contributed to higher value. However, it is unclear for both types of health directions whether it is the size of health effect or the severity of initial, or avoided health state that was the main rationing criteria for responders. Unfortunately, the construction of the PTO choices does not allow decomposition of these factors.

Contrary to the study by Ubel et al., which reports moderate preferences for preventive over curative treatments [12], responders to the present study strongly favored improving hypothetical patients' health compared to avoiding decline. The prioritization of curative treatments was present both in the person trade-off procedure as well as in the direct ranking task (represented by the high number of reversals in favor of cure). This is surprising since the onset of disease was defined as a certain event in the near future. It is likely that the value people place on preventive as compared to curative treatments diminishes as the extent of uncertainty and time to disease onset increase. One

Table 4: Mean and median number of errors and absolute rank deviation in the predicted rankings by elicitation technique

	SG	PTO-Down	PTO-Up	PTO-WAD
Mean number of errors	3.0	2.7	2.7	1.9
Median	4	4	4	2
Mean absolute deviation in ranks	3.9	3.0	3.0	2.1
Median	4	4	4	2

N = 127

explanation for the apparent discrepancy in results may be the different framing of questions. Ubel et al. presented the task as choices among different "new projects" of which a fixed number of nursing home residents could benefit from, i.e., on the program, rather than on the individual patient level. Contrary, in this study, participants were confronted with choices among treatments that could be offered to a varying number of patients. In addition, the SG tasks in which participants were asked to imagine themselves suffering from the health states preceded the rationing choices. The presence of suffering or need in terms of actual manifestation of the disease could have caused subjects to focus patients' situation and to place more value on curative treatment, even if the distance between health states is considerably smaller. This interpretation is strengthened by the qualitative comments provided by participants of the present study. A reasonable fraction argued "...it is cruel to deny treatment to those suffering with reference to the fact that resources are reserved to save others from entering this situation. What are patients left to expect from health care then?". Other qualitative comments suggest, that preferences for improving patients' health rather than avoiding decline might in part be caused by fairness concerns. Several responders stated considerations such as: "It is some form of luxury to prevent a disease as long as there are people already suffering the same disease, and as long as it can be treated". Spending resources on already suffering persons first and accepting others getting seriously ill, which might then be prioritized against others again, equalizes each individuals' chances to obtain appropriate treatment. Preferring avoidance of decline over improving patients' health in resource allocation, on the other hand, would increase inequality between those remaining ill and those fortunate to profit from treatment before suffering.

However, participants' preferences towards curative treatments were not absolute. As both, results from the combined PTO-WAD technique as well as the direct prioritization ranking, show, a vast majority was willing to trade their preference towards cure against a larger number of patients that could be saved from decline. The

basic idea of quantifying this relation between equally sized effects of different directions and applying this value as a weight for avoided health decline in comparisons between curative and preventive treatments proved feasible and has led to a substantial improvement in predicting individuals' rationing decisions concerning curative and preventive treatments. However, using the individual-specific mean as a weight may be a too rough estimate since a reasonable fraction of responders made their choices for or against curative treatments dependent of the health states involved in the PTO-WAD comparisons. Other, more complex weighting functions may produce even better results. Optimally, one could use (traditional) PTO-Up values to evaluate the benefit of curative treatments, PTO-Down values for the assessment of interventions avoiding decline, and the weighting technique for choices involving both types of interventions.

The inferior performance of the SG in predicting individuals' rationing choices as compared to the PTO procedures, at least at the ordinal level, is not surprising since Pinto and others report similar results [18,21,22]. One could argue that it is not reasonable to expect that SG utilities would accurately predict societal rationing choices among preventive and curative treatments, since this is not what they are intended to measure. While this is of course true, it is exactly what is done in health economic evaluation and it is the purpose of the study to question this practice by presenting the disagreement between SG, PTO and the direct prioritization ranking. For example, recently Zaric and Brandeau developed a resource allocation model for HIV prevention programs [23]. The calculation of potential gain in QALYs of preventing HIV infection was based on utilities attached to HIV associated health states by HIV-infected patients. In addition to the problem of personal versus social valuation of health states, this approach assumes that society would attach the same value to avoiding and leaving HIV-associated health states.

The study has several limitations and weaknesses. A main limitation is the small sample, which is highly biased to-

wards the well-educated and those using the Internet. In addition to the traditional self-selection bias common to online-research, the current study is also characterized by a strong overrepresentation of participants with a medical background. This bias is likely to result in part from the chosen newsgroups, in which the survey was announced and might be reduced by a more wide-spread invitation strategy. However, the main aim of the study was to introduce a technique for the joint valuation of health effects experienced in different directions and to examine its internal consistency with a direct prioritization task, rather than to yield representative figures. The unconventional scaling of health states, the presentation of the PTO questions as bidding games and the heavy demand on cognitive capabilities during the ranking task highlight the experimental character of the study. In addition, the PTO tasks did not provide information on the preceding duration of illness of hypothetical patients. Some of the health states used might have been recognized as congenital disabilities, which is likely to introduce particular bias. Also, preferences towards curative treatments might in part result from uncertainty or discounting future outcomes. Though both curative and protective actions were described as having certain and immediate effects, subjects might have intuitively perceived the latter as less certain or subject to a time shift, or even questioned the credibility of this description, since preventive programs with "certain effects in the near future" are scarce. Contrary, curative treatments might have been perceived as more harmful or associated with side effects, which would have introduced a bias towards prioritization of avoiding deterioration. A serious concern is the additive definition of three health states. Whereas they may fit an additive scheme of cure of each condition, an additive scheme of prevention may have been seriously doubted by participants. Realism would have introduced a strong bias in favor of cure. One would then expect individuals with a healthcare professional background being most sensitive to this issue. However, no "knowledge effect", in terms of association between professional background and preferences towards cure or avoidance of decline was observed. Finally, the observed strong preferences towards cure might also be explained by the fact, that – traditionally – the healthcare budget is perceived to produce above all treatment and cure for diseased patients. Again, this potential bias might be even more serious taking the high fraction of medical professionals into account, that can be expected to be for the most part practicing and experiencing curative care. On the other hand however, targeting and moving resources to preventive care is often claimed by public opinion and the medical profession. Preventive programs often rank very high in priority setting lists [24]. As Ubel et al. point out, such attitudes might be based on the belief, that it only needs "an ounce of prevention to bring large benefits" [12]. As in the present study effects of

avoiding decline and opportunity costs in terms of foregone curative effects were quantified and directly comparable, however, pro-preventive attitudes might have been altered.

There are also a number of methodological concerns. The fact that all PTO-choices started with the compared treatments offered to 100 patients each may have introduced bias from anchoring effects (starting-point bias) [25,26]. Though anchoring responses at 100/100 might have led to a systematic underestimation of participants' "true" PTO-values, this should not have affected their relative size or their relation to the direct ranking task. More generally, the PTO may be susceptible to order effects which may in part explain the differences between the PTO-Up and PTO-Down results [19]. A major concern is that the PTO-WAD technique, which was used to directly evaluate the value of health improvements compared to avoided decline of the same magnitude might have introduced and mixed action, status-quo and omission biases [27]. To explain results, action bias must be strong enough to overcome status-quo and loss omission bias, both of which suggest that more participants opt for avoiding decline, as this keeps the current situation unchanged for all patients and the decision maker cannot be held responsible for health deterioration. Patt and Zeckhauser observed such strong action bias in some individuals in experimental studies involving environmental decisions though [28]. Finally, one could be concerned with the level of engagement and reflection of responders in the presented numerical exercises. Given that participants are not forced to deliberation, they might present quick and unreflective answers, not necessarily consistent with their deeper values. The high number of thoughtful and serious qualitative comments, the observed internal consistency in the PTO-tasks, and between PTO-WAD and direct prioritization ranking indicate that subjects were not responding haphazardly and made considered choices. Also the small number of reversals among curative treatments observed in the ranking task compared to those found among curative and preventive treatments indicate that these may reflect "true values" rather than resulting from ill judgment. Conversely, there is no objective and direct evidence on the involvement of participants and the rationales underlying their choices.

Despite these limitations, the results of this study, however, suggest that the utility of being cured of a given health state might not be a good approximation for the societal value of avoiding this health state, especially in cases of competition between preventive and curative interventions.

Conclusions

Participants strongly prioritized improving patients' health rather than avoiding (further) decline. Weighting PTO values according to the direction in which changes in health are experienced improved their accuracy in predicting a direct prioritization ranking. This study also adds to the existing evidence that health state utilities obtained by the standard gamble method do not reflect social values in resource allocation contexts.

Competing interests

None declared.

Acknowledgements

I am particularly indebted to Thomas J. Strasmann for his support in graphical design and web programming. Special thanks are due to Erik Nord for his valuable comments on an earlier draft. I also thank Christian Koeck and Nicoletta von Laue for facilitating the survey. Any errors or omissions are the responsibility of the author alone.

References

1. Menzel P, Gold MR, Nord E, Pinto-Prades J-L, Richardson J, Ubel P: **Toward a broader view of values in cost-effectiveness analysis of health.** *Hastings Center Reports* 1999, **29**:7-15
2. Ubel PA, Nord E, Gold M, Menzel P, Prades JL, Richardson J: **Improving value measurement in cost-effectiveness analysis.** *Med Care* 2000, **38**:892-901
3. Nord E: **An alternative to QALYs: the saved young life equivalent (SAVE).** *BMJ* 1992, **305**:875-877
4. Nord E, Richardson J, Macarounas KK: **Social evaluation of health care versus personal evaluation of health states. Evidence on the validity of four health-state scaling instruments using Norwegian and Australian surveys.** *Int J Technol Assess Health Care* 1993, **9**:463-478
5. P Dolan: **The measurement of individual utility and social welfare.** *J Health Econ* 1998, **17**:39-52
6. Richardson J, Nord E: **The importance of perspective in the measurement of quality-adjusted life years.** *Med Decis Making* 1997, **17**:33-41
7. Nord E, Pinto-Prades J-L, Richardson J, Menzel P, Ubel P: **Incorporating societal concerns for fairness in numerical valuations of health programmes.** *Health Econ* 1999, **8**:25-39
8. Nord E: **The trade-off between severity of illness and treatment effect in cost-value analysis of health care.** *Health Policy* 1993, **24**:227-238
9. Abellan-Perpiñan J-M, Pinto-Prades J-L: **Health state after treatment: A reason for discrimination?** *Health Econ* 1999, **8**:701-707
10. Ubel PA, Richardson J, Prades JL: **Life-saving treatments and disabilities. Are all QALYs created equal?** *Int J Technol Assess Health Care* 1999, **15**:738-748
11. Chapman RH, Stone PW, Sandberg EA, Bell C, Neumann PJ: **A comprehensive league table of cost-utility ratios and a sub-table of "panel-worthy" studies.** *Med Decis Making* 2000, **20**:451-467
12. Ubel PA, Spranca MD, DeKay ML, Hershey JC, Asch DA: **Public preferences for prevention versus cure: what if an ounce of prevention is worth only an ounce of cure?** *Med Decis Making* 1998, **18**:141-148
13. Johannesson M, Johannesson PO: **A note on prevention versus cure.** *Health Policy* 1997, **41**:181-187
14. Polsky D, Willke RJ, Scott K, Schulman KA, Glick HA: **A comparison of scoring weights for the EuroQol[®] derived from patients and the general public.** *Health Econ* 2001, **10**:27-37
15. Salkeld G, Ryan M, Short L: **The veil of experience: Do consumers prefer what they know best?** *Health Econ* 2000, **9**:267-270
16. Tversky A, Kahnemann D: **Loss aversion in riskless choice: A reference-dependent model.** *Quart J Economics* 1991, **106**:1039-1062
17. Nord E: **The person-trade-off approach to valuing health care programs.** *Med Decis Making* 1995, **15**:201-208
18. Pinto Prades JL: **Is the person trade-off a valid method for allocating health care resources?** *Health Econ* 1997, **6**:71-81
19. Green C: **On the societal value of health care: what do we know about the person trade-off technique?** *Health Econ* 2001, **10**:233-243
20. StataCorp: **Stata Statistical Software: Release 7.0.** College Station, TX, Stata Corporation 2001
21. Dolan P, Green C: **Using the person trade-off approach to examine differences between individual and social values.** *Health Econ* 1998, **7**:307-312
22. Ubel PA, Loewenstein G, Scanlon D, Kamlet M: **Individual utilities are inconsistent with rationing choices: A partial explanation of why Oregon's cost-effectiveness list failed.** *Med Decis Making* 1996, **16**:108-116
23. Zaric GS, Brandeau ML: **Optimal investment in a portfolio of HIV prevention programs.** *Med Decis Making* 2001, **21**:391-408
24. Bowling A: **Health care rationing: the public's debate.** *BMJ* 1996, **312**:670-674
25. Green D, Jacowitz KE, Kahneman D, McFadden D: **Referendum contingent valuation, anchoring, and willingness to pay for public goods.** *Resource Engery Economics* 1998, **20**:85-116
26. Stålhammar NO: **An empirical note on willingness to pay and starting-point bias.** *Med Decis Making* 1996, **16**:242-247
27. Baron J: **Thinking and deciding.** Cambridge 2000
28. Patt A, Zeckhauser R: **Action bias and environmental decisions.** *J Risk Uncertainty* 2000, **21**:45-72

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMedcentral will be the most significant development for disseminating the results of biomedical research in our lifetime."

Paul Nurse, Director-General, Imperial Cancer Research Fund

Publish with **BMC** and your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours - you keep the copyright



Submit your manuscript here:

<http://www.biomedcentral.com/manuscript/>

editorial@biomedcentral.com