

ORIGINAL ARTICLE

Evaluation of three methods assessing the relative value of a dental program

RISTO TUOMINEN

Department of Public Health, University of Turku, Turku, Finland

Abstract

Objective. The aim of this study was to evaluate three commonly used methods for preference measurement in medical and dental programs. **Material and Methods.** In a questionnaire filled in by 156 first-year medical ($n = 120$) and dental ($n = 36$) students, they were asked to evaluate five programs, each costing approximately 2 million euros annually. The programs were helicopter ambulance service, MPR vaccination (measles, parotitis and rubella), breast cancer screening, 250 hip replacement operations, and a dental check-up program for 7-year-olds. **Results.** The respondents were willing to make the highest donations for running the helicopter ambulance service (39 euros) and least for the dental check-up programme (14 euros) ($p < 0.01$). When using the visual analog scale (VAS) method to evaluate the importance of the five programs, MPR vaccination was considered the most important (score 88), and hip replacement operations and dental check-up the least important (scores 68) ($p < 0.01$). With the rank order (RO) method, MPR vaccinations were ranked the highest (score 1.8) and hip replacement operations (score 4.2) and dental check-up (score 4.0) the lowest ($p < 0.01$). The respondents considered VAS and RO methods significantly easier than the willingness to pay method ($p < 0.001$). **Conclusions.** It is concluded that depending on the nature of the information required, each of the three methods can be used for assessing preferences. The willingness to pay method seemed to be the most difficult to comprehend, but it provides more detailed and ready-to-use information for economic evaluations than the other two methods.

Key Words: *Contingent valuation, economic evaluation, preferences*

Introduction

With scarce resources in health-care systems compared to demand, it is necessary to establish priorities. In economic evaluation studies, several methods have been applied to systematically compare the usefulness of health-care actions. Accurate estimates of both costs and benefits are necessary if reasonable comparisons of the programs are to be made and preferences set between them. Although forming accurate estimates of the costs has seldom jeopardized reliability of the evaluations, estimation of the utility of the programs has been more problematic.

Methods of measuring preference among alternative programs have comprised rank order (RO), rating scale (RS) (also called visual analog scale (VAS)), time trade-off (TTO), standard gamble (SG), and willingness to pay (WTP). Quality-adjusted life (or tooth) years (QAL(T)Y) is an

outcome measure that can be applied in cases where there are trade-offs between quality of life and years of survival. All of these measures have their pros and cons, with one favored by some and another by others [1].

WTP has gained in popularity in medical studies in recent years, with the number of studies applying WTP methodology increasing rapidly. In 1996, Donaldson [2] counted 28 studies of WTP for health care published in the English language, and in 1998 Olsen and Donaldson counted 32 [3]. By the end of 2006, the number in the MEDLINE database alone had increased more than tenfold.

Although the WTP methodology is tried and tested [4], there have been few attempts to apply it in evaluating the usefulness of dental programs [5–8]. As resources in the health-care sector overall are scarce, systematic comparisons of dental and medical programs are justified. However, there are no studies making such comparisons.

Correspondence: Risto Tuominen, Department of Public Health, Lemminkäisenkatu 1, FIN-20014 University of Turku, Finland. Tel: +358 2 333 8434. Fax: +358 2 333 8439. E-mail: risto.tuominen@utu.fi

(Received 22 October 2007; accepted 1 February 2008)

ISSN 0001-6357 print/ISSN 1502-3850 online © 2008 Taylor & Francis
DOI: 10.1080/00016350801958274

The aims of this study were: 1) to evaluate three commonly used methods for preference measurement in medical and dental programs by comparing the outcomes of WTP, VAS and RO, and 2) to evaluate the relative values of the selected programs.

Material and methods

Preference measurements

The RO method is among the simplest to solicit people's preferences. In this method, subjects are provided with a list of alternatives and asked to rank them from the most preferred to the least preferred. With adequate description of the alternatives, this method provides information on the order of the alternatives, but gives no information on the relative importance of the programs, nor about the magnitude of the differences between them. The RO method is easy and fast to perform, which reduces the costs of data collection.

The rating scale method is usually in the form of a VAS with clearly defined endpoints or anchors. It may be a line of a certain length or a 'feeling thermometer bar' with a scale drawn from 0 to 100 [9]. Subjects are asked to mark the position on the line or bar to indicate their opinion. Endpoints may be certain states of health (e.g. death and perfect health) or indications of importance (e.g. extremely unimportant and extremely important). The rating scale method has been widely used [10] because it is simple to apply and easily understood [11]. One commonly mentioned disadvantage, however, is the lack of a time component; for example, the length of time a certain state of health is expected to last. In addition, VAS scores do not give any definite monetary value, which is necessary if preferences are to be estimated in relation to scarce resources.

With the WTP method, respondents are asked how much money they would be willing to contribute in order to get a service, to reach a certain state of health, or to support a health-care program. In surveys, they are given descriptions of the service or program or told about the likelihood of reaching a certain state of health. They then consider their own health and economic status and evaluate how important the program would be to them and how much they could afford to pay. This method offers the exact monetary value one is willing to contribute to respective programs and these values can be directly compared with the actual or expected costs of each program. It is then possible to see which programs can be expected to produce the highest cost-effectiveness ratios. One disadvantage of the method is that in reality no monetary sacrifice needs to be made which might distort the estimates.

The survey

After ethics committee approval, this survey was conducted among all first-year medical ($n=121$) and dental ($n=36$) students on the second day of their studies. After giving their written informed consent they were asked to fill out a questionnaire during a lecture. There was no time limit for completing the questionnaire. One medical student did not want to participate, which gave a final sample size of 156 responses.

Among the first entries on the form were questions about age, sex, whether they were studying medicine or dentistry. One question was to estimate the total value of their monetary assets and personal property.

Five programs, each costing approximately 2 million euros annually, were chosen. The cost estimates were ascertained from the organizations responsible for running them. The programs were: running a helicopter ambulance in the metropolitan Helsinki area for one year, running a national MPR vaccination program for children for one year, running a breast cancer screening program for 50 to 59-year-old women for one year, conducting 250 hip replacement operations within one year, and running a dental check-up program for 7-year-old children for one year. The content and organization of each program were described, as well as the size and characteristics of the target populations. The expected program effects with time-perspectives were then presented based on the available evidence from documented research. The programs were described as not mutually exclusive, i.e. that choosing one program would not affect the opportunity to run the others.

The preference questions were ordered so that, first, the students were asked their view of the importance of a helicopter ambulance in the metropolitan Helsinki area, which was to be their place of residence for at least the next six years. This was to be done by marking an 'X' on a line exactly 10 cm in length. The left side endpoint of the line was labeled 'extremely unimportant' and the right side endpoint 'extremely important'. To ease the task, the mid-point of the line was labeled 'indifferent'. The VAS value was measured as distance in millimetres from the left side endpoint.

After this they were asked to say how much they would be willing to make as a one-off donation for the helicopter ambulance to be functioning for one calendar year, in the event this service was planned to be terminated due to lack of funds.

Following from the two helicopter ambulance questions, similar questions, in the same order and manner, were asked for the publicly funded and nationally run MPR vaccination program, the breast cancer screening program, the 250 hip replacement operations, and the dental check-up program.

The students were then asked to give their rank order for the 5 programs by giving each a value from 1 to 5 (1 = highest ranked, 5 = lowest ranked). Finally, they were asked to estimate the difficulty of answering with each of the three methods by scoring from 0 = extremely easy to 10 = extremely difficult.

Statistical evaluation of the data was based on independent and paired samples *t*-tests, the chi-square tests and correlation coefficients. When analysing the effects of age, sex, studying medicine or dentistry and economic status on perceptions, regression models were fitted. Linear regression models were used when perception was measured in continuous form (VAS, WTP), while logistic regression models were used for categorical data (RO) and when dichotomous data transformations were analysed.

Results

With the WTP method, both female and male students were ready to make the highest donations for running the helicopter ambulance. The average donations for the helicopter ambulance were significantly higher than for mammography ($p < 0.05$), hip replacement operations ($p < 0.001$) and dental check-up ($p < 0.005$) programs. The overall differences between MPR vaccination or mammography programs and hip replacement operations ($p < 0.01$) or dental check-up ($p < 0.01$) programs were statistically significant. The differences between the helicopter ambulance service and the MPR vaccination program, hip replacement and dental check-up programs were not significant (Table I).

With the VAS method, the MPR vaccination program was considered the most important by both male and female respondents ($p < 0.001$), followed by mammography screening and the helicopter ambulance service with almost equal values. Females considered dental check-ups least favorably, while males considered only hip replacement operations less favorably than the dental program. The importance of both dental check-ups ($p < 0.01$ among females and $p < 0.05$ among males) and hip replacement operations ($p < 0.01$ among females and $p < 0.001$ among males) differed from those of the other three programs (Table II).

With the RO method, the MPR vaccination program was most often ranked as having the highest

value, followed by the helicopter ambulance service and mammography screening. Dental check-ups and hip replacement operations were least favorably ranked ($p < 0.01$). Females ranked hip replacements higher ($p < 0.01$) than males did, and males ranked dental check-ups higher ($p < 0.01$) than females did (Table III).

When the respondents were asked to evaluate the difficulty of each method for assessing their preferences, VAS and RO methods were considered significantly easier than the WTP method ($p < 0.001$) among both males and females. Females perceived their perception task more difficult than males did, particularly the RO and the WTP methods (Table IV).

With the VAS method, younger respondents perceived the MPR vaccination program significantly ($p < 0.05$) more important than older respondents did. Otherwise age, economic status, and whether studying medicine or dentistry did not affect the perceptions of any of the programs with the three methods.

Preference measurements with the VAS and WTP methods correlated positively. However, the correlations were statistically significant only for the dental check-up program ($r = 0.215$, $p < 0.001$) and hip replacement operations ($r = 0.174$, $p < 0.05$). Subjects were willing to donate the highest amounts and scored highest perceptions in VAS to the programs they ranked the highest.

Discussion

The participation rate was considered good. Only one medical student refused to participate. Six respondents had not given any figure for the value of their monetary assets and personal property. Otherwise, all 156 forms were completed. The findings of this study can therefore be taken as a good representation of the views of first-year medical and dental students. However, they should be generalized only cautiously in other populations; this study is to be considered as a pilot study.

These five programs were chosen because they are run in society, each costing approximately 2 million euros per year. WTP comparisons between helicopter ambulance and hip replacement operations have been done in an earlier study [3].

Table I. Willingness-to-pay in single donation (in Euros) to maintain the five health care programs for one calendar year by sex.

	Females ($n = 96$)	Males ($n = 60$)	All ($n = 156$)
Program			
Helicopter ambulance	36.27	43.12	38.90
MPR vaccination	25.03	38.75	29.92
Mammography screening	21.41	25.61	23.02
Hip replacement operations	16.77	19.38	17.78
Dental check-up for 7-year-olds	13.60	15.51	14.34

Table II. Relative importance of the five health care programs by sex measured with visual analog scale (from 0 =not at all important, to 100 =extremely important).

	Females (<i>n</i> =96)	Males (<i>n</i> =60)	All (<i>n</i> =156)
Program			
Helicopter ambulance	78	79	79
MPR vaccination	88	89	88
Mammography screening	81	79	80
Hip replacement operations	70	65	68
Dental check-up for 7-year-olds	66	71	68

All the selected programs were such that subjects could be expected to be able to make value evaluations and comparisons between them without having personal experience of them. A common problem with preference measurements has been that the alternatives offered may not have been adequately understood by the respondents. Asking a lay person to give a monetary value for risk reductions of dying in traffic accidents [12] could serve as an example.

Programs offered for evaluation need to be described in such a manner that respondents are able to make rational comparisons between them. Methods like TTO, SG and QAL(T)Y require that subjects are able to estimate the value of a certain time before they die, remain in a certain state of health or lose a tooth, which undoubtedly is extremely difficult, even impossible. Furthermore, in the SG method, subjects are asked to give values for increasing or decreasing probabilities of encountering a certain state of health. Such tasks can be expected to be overwhelming for most lay persons, which raises doubts about the reliability of the obtained estimates.

The starting point for preference measurements is to obtain population level estimates of the values of planned or running actions or programs, usually in connection with CEA or CBA evaluations. Both the VAS and RO methods are simple and easy to perform, but they do not provide information on monetary value, which limits their use. The WTP method is considered relatively easy to comprehend [13,14] and can provide the directly comparable monetary values necessary for economic evaluation techniques to be applied.

In some surveys dealing with public programs, the subjects have been asked to give their preference

measure as willingness to contribute in the form of increased taxes [3,6]. It can be difficult to estimate one's marginal tax burden and thus the taxation alternative was not used in this survey.

In this survey, the MPR vaccination, helicopter ambulance, and mammography screening programs formed a higher valued group with all three methods. Hip replacement operations and dental check-ups were valued significantly less favorably. In a previous study [3], both helicopter ambulance and hip replacement programs were also estimated to cost approximately 2 million euros. In this earlier work, the subjects were willing to donate significantly more to the helicopter ambulance than to a higher number of hip replacements.

Each of the three methods produces values that can be used to differentiate between the more and less favored programs, particularly when the value differences are large enough. If there is no need for more detailed monetary values, and information is needed only on relative importance, then less complicated and easy to perform methods such as VAS and RO could be used.

In VAS and WTP measurements, the dental program was offered as the last of the five programs to be evaluated. Thus, the respondents had already made their estimates of the importance and relative value of the other four programs. When an interview technique is used for soliciting values, the order of presentation of the alternatives for evaluation can produce bias. When the respondent gives his/her value statement for the first alternative, it can set limits for the further evaluations. In this study, order bias obviously did not affect the value statements, because the respondents had the opportunity to change all values before returning the questionnaire

Table III. Rank-order scores for the five health care programs by sex (from 1 =most preferable, to 5 =least preferable).

	Females (<i>n</i> =96)	Males (<i>n</i> =60)	All (<i>n</i> =156)
Program			
Helicopter ambulance	2.3	2.3	2.3
MPR vaccination	1.9	1.7	1.8
Mammography screening	2.7	2.8	2.7
Hip replacement operations	4.0	4.4	4.2
Dental check-up for 7-year-olds	4.2	3.7	4.0

Table IV. Average perceived difficulty in estimating the values of five health care programs by three methods and sex (from 0 = extremely easy, to 10 = extremely difficult).

	Females (<i>n</i> = 96)	Males (<i>n</i> = 60)	All (<i>n</i> = 156)
Willingness to pay	6.9	5.8	6.4**
Visual analog scale	4.3	4.2	4.3
Rank order	4.9	3.9	4.5*

Statistical evaluation between sexes by Student's *t*-test: **p* < 0.05, ***p* < 0.01.

form. With a larger sample it would be possible to offer programs for evaluation in varying order and any possible effect of order bias could be evaluated. Dental check-up starting from 7-year-olds instead of one year later has a similar value to this student population as 250 annual hip replacement operations. The other three health-care programs seemed to have significantly higher value.

Among the five programs, the helicopter ambulance service is the only one using various donation campaigns and other forms of active marketing efforts to raise funds for running the program. These marketing activities are well known to the public. The other four programs are funded solely from public funds; there has been no active marketing or fund raising for them. The subjects in this study, their relatives or friends may already have contributed to the helicopter ambulance program, which may have influenced their reported willingness to pay for it. Olsen & Donaldson suggested [3] that one major reason for valuing the helicopter ambulance program is 'selfishness'; people may think that they or their close ones might one day urgently need this particular service.

WTP has been criticized for giving more weight to the wealthy, who can contribute more than the poor [15]. However, if the values are collected in order to form a comprehensive picture in a certain population, then the distribution of wealth will be implied in the average values. In this study, socio-economic and other background factors affected the preference measurements very little. One reason for this could have been the homogeneity of the respondents. With a more heterogeneous sample greater variation might have been found [16].

Medical and dental students seemed to have equal values. A priori, one might have expected that dental students would have valued the dental check-up program more than their medical colleagues did, and the opposite to occur with some medical programs.

Although considered more difficult than the other two methods, the WTP method gives more information. Large numbers of missing values or zero values have been thought to reduce the reliability of the WTP study findings [17]. In this study, however, all

156 subjects gave their preferences with all three methods and no missing values were found in preference questions, indicating that the WTP method is indeed not too complicated, but subjects' preferences can be measured by estimating their willingness to pay.

Conclusion

The order of importance of health care programs to the population could be obtained by any of the three methods. Although the study subjects considered the WTP method more difficult than the other two, its use is preferable because of the more detailed information it provides.

Acknowledgment

This study has been supported by a grant from the Finnish Dental Society Apollonia.

References

- [1] Krabbe P, Essink-Bot M-L, Bonsel G. The comparability and reliability of five health-state valuation methods. *Soc Sci Med* 1997;45:1641–52.
- [2] Donaldson C. Willingness to pay for publicly provided health care. PhD Thesis, University of Aberdeen, Aberdeen, Scotland, 1996.
- [3] Olsen J, Donaldson C. Helicopters, hearts and hips: Using willingness to pay to set priorities for public sector health care programmes. *Soc Sci Med* 1998;46:1–12.
- [4] Matthews D, Gafni A, Birch S. Preference based measurements in dentistry: a review of the literature and recommendations for research. *Community Dental Health* 1999;16:5–11.
- [5] Matthews D, Birch S, Gafni A. Willingness to pay for periodontal therapy: Development and testing of an instrument. *J Publ Health Dent* 1999;59:44–51.
- [6] Dixon S, Shackley P. Estimating the benefits of community water fluoridation using willingness-to-pay technique: results of a pilot study. *Community Dent Oral Epidemiol* 1999;27:124–9.
- [7] Cunningham S, Hunt N. Relationship between utility values and willingness to pay in patients undergoing orthognathic treatment. *Community Dental Health* 2000;17:92–6.
- [8] Matthews D, Rocchi A, Gafni A. Putting your money where your mouth is: willingness to pay for dental gel. *Pharmacoeconomics* 2002;20:245–55.
- [9] Torrance P, Feeny D. Utilities and quality-adjusted life years. *Int J Technol Assess in Health Care* 1989;5:559–75.
- [10] Froberg D, Kane R. Methodology for measuring health-state preferences-II: Scaling methods. *J Clin Epidemiol* 1989;42:459–71.
- [11] Matthews D, McCulloch C. Evaluating patient perceptions as short-term outcomes of periodontal treatment: a comparison of surgical and nonsurgical therapy. *J Periodontol* 1993;64:990–7.
- [12] Norinder A, Hjalte K, Persson U. Scope and scale insensitivities in a contingent valuation study of risk reductions. *Health Policy* 2001;57:141–53.
- [13] Robinson R. Cost-benefit analysis. *Br Med J* 1993;307:924–6.

- [14] Cross M, March L, Lapsley H, Tribe K, Brnabic A, Courtenay B, Brooks P. Determinants of willingness to pay for hip and knee joint replacement surgery for osteoarthritis. *Rheumatology* 2000;39:1242–8.
- [15] Donaldson C. Eliciting patients' values by use of 'willingness to pay': letting the theory drive the method. *Health Expect* 2001;4:180–8.
- [16] Frew E, Wolstenholme JL, Whynes DK. Willingness-to-pay for colorectal cancer screening. *Eur J Cancer* 2001;37:1746–51.
- [17] Gyldmark M, Morrison GC. Demand for health care in Denmark: results of a national sample survey using contingent valuation. *Soc Sci Med* 2001;53:1023–36.