Contingent Valuation of Eight New Treatments: What is the Clinician's and Politician's Willingness to Pay?[#]

Erkki J. Soini[*,1], Jarmo Kukkonen[2,3], Markku Myllykangas[2] and Olli-Pekka Ryynänen[2,4]

[1]*ESiOR Oy, Kuopio, Finland*

[2]*Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland*

[3]*Joensuu Health Center, Joensuu, Finland*

[4]*Primary Health Care, Kuopio University Hospital, Kuopio, Finland*

Abstract: *Objective*: To assess the willingness to pay (WTP) for eight new treatments from a life-long perspective. *Methods*: A contingent valuation with virtual examples and dichotomous choice questions is circulated to Finnish clinicians (N 146) and politicians (N 73). Costs and utilities (15D, EQ-5D) are obtained from Finnish sources, and the health care payer perspective is assumed. Health benefits are measured using life-years gained (LYG) and quality-adjusted life-years (QALY) gained, and 3% and 0% annual discounting is done. The results are presented as different WTP thresholds (incremental and aggregate cost-effectiveness ratios, and incremental investments, II). Heterogeneity is handled using conditional (Hurdle) modeling. *Results*: In 1,092 decisions, the mean discounted (undiscounted) incremental WTP/QALY gained is €102,616 (€78,686) and €94,770 (€77,856) measured with 15D and EQ-5D, respectively. The mean discounted (undiscounted) incremental WTP/LYG is €66,277 (€58,160). The highest incremental WTPs are reported for cancer (€205,994–250,509/QALY gained) and lowest for metabolic disease (€23,492–43,398/QALY gained) treatment. The discounted (undiscounted) IIs to health care are €83,886 (€85,398) Euros; metabolic presenting the highest (€199,499-213,808) and coronary heart disease treatment (€36,124-36,736) the lowest value for the lifetime of the patient. WTP is dependent upon disease/treatment, patient's age, time preference, health benefit type and discounting. Minor differences between clinicians and politicians are observed. *Conclusion*: WTP vary for different diseases and is not explained by incremental costs. Thus, a single WTP for all treatments/diseases hypothesis do not gain empirical support - WTP is better explained by treatment and patient/disease characteristics. Cost-effectiveness and II have a trade-off, which encourages studies including both efficiency and affordability.

Keywords: Budget impact, cost allocation, cost-utility, decision making, discrete choice, economic evaluation, quality of life, willingness to pay.

INTRODUCTION

The implementation limitations for new treatments in health care have centered on cost saving strategies. The reasons are obvious: costs are rising and resources are becoming more limited, resulting in decision situations where evidence of treatment efficiency and affordability is needed. However, a question remains: what is the willingness to pay (WTP) for an efficient or affordable treatment?

Treatment selection and targeting are addressed with cost-effectiveness (CEA, value for money, investment turn-over, "efficiency") and budget impact (BIA, money for value, investment impact on budget, "affordability") analysis. CEA and BIA should use the decision maker's perspective in the correct time window. Typically, BIA handles net budget changes in monetary terms through the cost and health impacts during a limited time horizon of 1–5 years.

When performing BIAs, there is a need to estimate the budget. Obviously, the settlement of the budget is also a question of justice and equity.

On the other hand, CEAs characterize the marginal value of treatments in the form of additional cost per additional unit of health benefit gained or saved during a with-in trial, restricted time-frame or life-long perspective. WTP means, for example, the acceptable cost for benefit that can be considered a decision maker's WTP for the health benefit gained (welfarism) or a decision maker's shadow budget (extra-welfarism) [1-3]. WTP has been the subject of wide theoretical discussion [1-7].

The current study focuses on acceptable CEA and BIA thresholds through the WTP for various health gains and on incremental investment (II, "willingness to budget") per patient for health care in a lifelong scenario. For CEA, health benefits are measured in different units including life-years gained (LYG) and quality-adjusted life-years (QALY) gained. The outcome presentation methods of CEA also include aggregate cost-effectiveness ratio (SCER) to sum up the average WTP and incremental cost-effectiveness ratio (ICER) for the comparison of mutually dependent options.

*Address correspondence to this author at the ESiOR Ltd., Tulliportinkatu 2 LT4, 70100 Kuopio, Finland; Tel: +358 40 053 3971;
E-mail: erkki.soini@esior.fi

The basic tool in the modern assessment of CEA is ICER:

$$ICER = \frac{\Delta\overline{C}}{\Delta\overline{E}} = \frac{\overline{C}_1 - \overline{C}_0}{\overline{E}_1 - \overline{E}_0},$$

where C stands for average costs and $\overline{E}$ for average health benefits (subscripts indicate treatment). A treatment in which ICER is at or falls below the incremental WTP (λ) for the health benefit gained indicated as a threshold or ceiling ratio [8-11] is cost-effective - i.e., for efficient treatment ICER$\leq\lambda$. In the case of perfect divisibility and constant returns of scale, ICER results in efficiency [12, 13] if WTP is the shadow price for budget constraint or the marginal opportunity cost of resources [14]. ICER is a ratio that has a problematic distribution. Thus, it should be handled in a stochastic or probabilistic form [15].

In addition, the difference in average cost-effectiveness ratios (CER) is not the same as ICER [16, 17]:

$$ICER \neq \frac{\overline{C}_1}{\overline{E}_1} - \frac{\overline{C}_0}{\overline{E}_0} = CER_1 - CER_0.$$

If CERs are estimated without budget constraints [18], they have only descriptive value. CER can be used for the comparison of mutually independent options, but not in the comparison of mutually dependent treatment options. That is, in comparative decision situations CER compares the treatments to unrealistic "do nothing" situations [15, 16, 19], which results in the lowest cost per health benefit (i.e., a form of cost-minimization analysis). In simple terms, the comparator of the CER analysis tends to be "instant death" with zero costs and effectiveness, i.e.:

$$ICER = \frac{\overline{C}_1 - \overline{C}_0}{\overline{E}_1 - \overline{E}_0} = CER$$

is true only when C_0=0 and E_0=0. Thus, there are probably only two special cases when CER=ICER: if the comparator has zero costs and effectiveness or CER for all compared treatments is equivalent.

The probability of cost-effectiveness [20] or highest expected net monetary benefit (NMB [21]) conditional to WTP can have an impact on the decision: if only two treatments are compared, typically a higher probability is expected for the given WTP (i.e., $\geq$50%, which also means that the treatment is potentially cost-effective); if more than two treatments are compared, a lower than 50% probability may be accepted [22]. In addition to the number of treatments compared, the goodness of the data and the value of information (VOI [20]) are likely to impact the decision making. Since ICER and the highest VOI correspond to the change of optimal treatment in the cost-effectiveness acceptability frontier (CEAF [20, 23]) in the comparison of two treatments, neither CEAF nor VOI curves were presented here.

The questions asked are:

1.   What is the level of WTP (in Euros) for different diseases and treatments measured using different health units and as an incremental investment?

2.   What has an impact on the level of WTP for different diseases and treatments (e.g. respondent, patient characteristics and different outcome presentation methods) - i.e., could a single WTP for all treatments hypothesis gain empirical support?

MATERIALS AND METHODOLOGY

Perfect competitive markets do not exist in health care and the real prices for benefits cannot be directly observed as consumer behavior. On the other hand, health markets have a so-called principal agent structure, which means that the agents (e.g. politicians making macro-level allocations and clinicians making micro-level care decisions; here henceforth agents) can be the "true" decision makers with a considerable impact on the principal (the patient).

Contingent Valuation and Dichotomous Choice Methods

In the present study, the missing markets are simulated using contingent valuation (CV), which has performed well [24]. Through hypothetical survey questions, the CV aims to establish an agent's WTP for health benefits that do not have a market price. Stated – not revealed – preferences [25] are used in the CV. In this way, the maximum average WTP corresponds to the maximum amount of money for indifference. In the CV, each treatment is a bundle of potential attributes, each responder has a set of unique relative utility weights for the attributes and attribute levels, and the process of combining the utilities for different attributes provides overall relative utility [26].

The study is influenced by Louviere *et al.* [27]. Thus the CV is started by building eight vignettes presenting new futuristic treatments with varied outcomes: stem cell treatment for dementia (dementia), a medication for the addiction of illegal drugs (addiction), a new cancer medication (cancer), a genetic therapy for coronary heart disease (CHD), a medication for serious neurological disease (neurologic), a treatment for a serious congenital metabolic disease (metabolic), the treatment of diabetes by implantation of insulin-producing cells (DM), and the treatment of traumatic paraplegia by an electronic advice giving artificial nerve impulses (paraplegia).

A discrete choice approach is used in the CV: the attribute levels were assumed to determine the value of alternative [28] and the agents either accepted or rejected the futuristic scenario contingent to the current treatment. Thus, every vignette contains two possible treatments: a new, futuristic treatment and a more conservative alternative, which is targeted to present the most common treatment currently in use. This approach is called as the dichotomous choice (DC) analysis, which mimics true consumer decisions. The DCs used here are probably not as prone to problems presented in the bounded rationality [29] and prospect [30] theories as the conventional CVs or multiple discrete choices. These problems include starting point bias [4, 31], a typical phenomenon with open-ended questions [32].

The vignettes contain information about the patients' age, gender, health before (fixed inputs) and after the treatment, and marginal costs (stochastic inputs). A cognitively simple illustration of excepted average effectiveness in terms of lifetime for both alternatives and marginal costs are included in the vignettes. Some assumptions (homogenous treatments

[28], perfect divisibility and deterministic behavior) are relaxed by using somewhat heterogeneous future treatments, by dealing with an infinite and mutually exclusive set of alternatives, and by using random utility [33,34]: the respondents are assumed to implicitly attach utilities for the outcomes and to have perfect discrimination capacity.

Analysis Sample and Perspective

A set of five vignettes is randomly chosen for the questionnaires and sent as a postal questionnaire with a cover letter to a random sample of 200 physicians of working age (derived from the registers of the Finnish Medical Association) and 218 politicians on the executive boards of 20 Finnish health districts (the health districts were asked to give their names). In Finland, health districts are owned by municipal governments and are responsible for the secondary and tertiary care. The districts have a politically elected executive board. Altogether, 418 questionnaires are sent (a total of 2,090 vignettes). The questions about the respondent's age, gender, working environment or socio-economic status are excluded in order to get high response rates.

In line with the sampling, a health care payer perspective excluding productivity losses and value added tax is assumed. Discounting is done using the officially recommended annual rates of 0% and 3%.

Analysis Inputs

For the analysis, the inputs for the current life-time treatment costs are estimated from the national Finnish unit cost list from the year 2006 [35]. These costs are case-mix adjusted, and present the average cost of various treatment units in Finland. The marginal costs for the futuristic treatments are taken as stated in the questionnaires (Table **1**).

The effectiveness inputs in terms of life expectancy are taken as stated in the questionnaires. Expected lifetime is converted to quality-adjusted life-years (QALY) by attaching the age-, gender- and disease adjusted quality of life (QoL) values to the individual vignettes and by extrapolating these to the remaining lifetime. Among the patient reported outcome (PRO) measures, QALYs gained are seen as the primary outcome in the assessments made by the National Centre for Health and Clinical Excellence (NICE) in the United Kingdom [36-38] as well as in Australia, China, Sweden and the United States, and in the estimation of WTP [37]. QoL values are obtained from the Health 2000 study during which both EQ-5D [39] and 15D [40] generic QoL values has been collected from the same patients ([41] Table **1**).

Analysis Outcomes

The primary outcome of the analysis is measured as WTP for incremental cost (IC) per incremental QALY gained (ICUR). The secondary outcomes are WTPs measured as IC per incremental LYG (ICER), aggregate outcomes measured as sum cost per sum QALY (SCUR) and sum cost per sum life-years (SCER), and incremental and average outcomes per patient. In addition, budget impacts as incremental investments (II, willingness to budget) per life or patient are derived based on the lifetime cost difference between the treatment options, *ceteris paribus*.

Incremental QALYs gained, LYGs and ICs are estimated only when the futuristic option is accepted by the agent. Meanwhile, the aggregate cost-effectiveness (based on regression-adjusted average per-patient QALYs, LYGs and costs) are defined for the whole population based on the choice the agent makes:

$$SCUR = \frac{\sum_{i=1}^{j} C}{\sum_{i=1}^{j} E},$$

where C_j is the cost of the jth decision and E_j is the health benefit of the jth decision. The per patient total costs and

Table 1.   Vignettes and Their Variations in Life Expectancy, Quality of Life and Lifetime Incremental Costs of Treatment

Disease (Vignettes)	New Treatment	Life Expectancy (LE) Variation	QoL in the Index Year EQ-5D (15D) new *vs.* old	Incremental Costs (€), not discounted
Dementia (148)	Stem cell transplantation	Add 6-12 months compared to present treatment	0.748 (0.801) *vs.* 0.748 (0.801)	65,000-200,000
Addiction (157)	Pharmaceutical treatment together with a rehabilitation period in an institution	5-15% of patients achieve normal LE, compared to 10 years in present treatment	0.843 (0.889) *vs.* 0.843 (0.889)	40,000-160,000
Cancer (165)	Pharmaceutical treatment for metastatic disease	10-12 months compared to 4 months in present treatment	0.815 (0.887) *vs.* 0.815 (0.887)	35,000-120,000
CHD (159)	Genetic therapy	6-8 years compared to 5.5-6 years in present treatment	0.831 (0.874) *vs.* 0.831 (0.874)	40,000-220,000
Neurologic (142)	Pharmaceutical treatment for severe muscular atrophy	7-10 years compared to 6 years in present treatment	0.802 (0.892) *vs.* 0.802 (0.892)	60,000-200,000
Metabolic (93)	Enzyme therapy for congenital metabolic disease in a new-born	30 years, compared to 20 years without treatment	0.556 (0.600) *vs.* 0.659 (0.700)	Annual cost: 40,000-300,000
DM (135)	Transplantation of Langerhans' islets	No effect on LE	0.844 (0.896) *vs.* 0.885 (0.917)	40,000-210,000
Paraplegia (93)	Electronic device giving artificial nerve impulses	No effect on LE	0.784 (0.889) *vs.* 0.885 (0.917)	30,000-200,000

CHD = coronary heart disease, DM = diabetes mellitus.

other per patient total outcomes for the regression model adjustment, SCURs and SCERs are estimated based on the choice the agent makes: the costs and other outcomes are estimated as for conventional scenario for the agents who did not approve the futuristic scenario for particular disease, and the costs and other outcomes related to futuristic scenario for the agents who approved the futuristic scenario for particular disease. These aggregate measures (SCUR, SCER) can show the average outcome of what the agents choose based on the whole population (in comparison, for example, the incremental outcomes apply only for those cases investing in the futuristic treatment).

Statistical Methods

The analysis is carried out using Stata 10 statistical software. The incremental outcomes are adjusted against heterogeneity using the conditional Hurdle-type regression [42, 43], which reflects a two-stage decision-making process with an appealing interpretation [44]:

1.   Logistic regression models (logit) are used to estimate the probability of agents dichotomously (yes or no) accepting the futuristic treatment. The models are adjusted for the potential differences in treatment groups between the clinicians and the politicians, and also reveal the potential differences between the acceptance in different treatment groups and between different agent groups.

2.   The ordinary least squares (OLS without or with ln-transformation (ln OLS, semi logarithmic OLS) for the dependent variable; conversion to unbiased means Duan smoothing [45]) or generalized linear models

(GLM with log link function and gamma distribution) are used for the outcomes. The models are adjusted for potential differences in outcomes between the clinicians and the politicians.

3.   The adjusted expected incremental outcome is obtained at the data level by weighting the outcome 2) with the probability of acceptance 1).

The fitness of the first and second stage models are assessed with coefficient of determination (R^2, the proportion of variability that is accounted for in OLS), log likelihood (LL), Akaike (AIC) and Bayesian (BIC) information criteria [45-49]. Of these, AIC and BIC are measures of relative goodness of fit: they penalize over fitting and help in determining between the bias and variance trade-off. For average per patient outcomes, only the second stage 2) is done.

RESULTS

Of the questionnaires sent, 219 (52.4%) are returned properly filled out and are included in the analysis. Of these, 146 are from physicians and 73 are from politicians, with the respective valid response rates of 73.0% and 33.5%. Altogether, 1,092 vignettes (52.2% of those sent) are analyzed.

Acceptance Probabilities

In the unadjusted total data, 43.5% of the physicians and 47.1% of the politicians prefer the new treatment over the present treatment. The difference between the groups is not statistically significant (acceptance difference 3.6%, odds ratio [OR] 0.87, p=0.271; Table **2**, Logit1). Among the diseases, only addiction (p=0.003) treatment reveals statistically

Table 2.   **Logistic Regression Models (N 1092) for Treatment Choice (in Favor of Futuristic Treatment); Logit2 is the First Stage in the Hurdle-Type Model**

Model	Logit1	Logit2			Logit3	Logit4	Logit5
Parameter[#]	*OR*	*OR*	*SE*	*p*	*OR*	*OR*	*OR*
Clinician	0.8674 p=0.271	0.7591	0.1102	0.058	0.7809 p=0.070	0.8692 p=0.279	0.8361 p=0.174
Dementia		0.5889	0.1511	0.039*			
Addiction		2.5315	0.5898	0.000***			
Cancer		0.5461	0.1379	0.017*			
Neurologic		1.4731	0.3517	0.105			
Metabolic		1.4344	0.3866	0.181			
DM		9.0378	2.5401	0.000***			
Paraplegia		9.6881	3.1296	0.000***			
Baseline age, patient					0.9743***		
Marginal costs						1.0000 p=0.227	
Time preference							1.4059***
Tests							
Chi^2	1.21 p=0.271	223.62		0.000***	114.55***	2.70 p=0.258	37.52***
Pseudo R^2	0.0008	0.1484			0.0760	0.0018	0.0249
LL	752.60	641.39			695.93	751.85	734.44
AIC	1,509.19	1,300.79			1,397.86	1,509.70	1,474.89
BIC	1,519.19	1,345.75			1,412.85	1,524.69	1,489.88

[#]Politician and CHD are the reference groups. Logit = logistic regression model. OR = odds ratio. SE = standard error. p = p value. N = number of observations. Chi^2 = Chi^2 test. R^2 = coefficient of determination. LL = Log likelihood. AIC = Akaike information criteria. BIC = Baysian (Swartz) information criteria. DM = Diabetes mellitus. CHD = Coronary heart disease. *** p value<0.001, * p value<0.050.

significant and CHD treatment suggestive (p=0.072) differences between physicians and politicians; the politicians favor the futuristic treatment more than the physicians (unadjusted data) in both.

The Table **2** presents the logit models for futuristic treatment acceptance (yes or no) in detail (in order to avoid dummy variable trap, CHD and politician are reference groups against which the ORs and p values of particular variable can be interpreted). The treatment and the patient's age (Logit3, 0.97 odds for accepting treatment with one year increase in age – i.e., treating of younger persons is preferred) at the baseline have an impact on the decision. Based on the adjusted results, the politicians favor futuristic treatments more than the physicians - however, the difference remains insignificant. When the results are adjusted for heterogeneity in treatments (Logit2) or patient's baseline age (Logit3) between the responder groups, the statistical significance becomes suggestive (OR=0.76 and 0.78; p=0.058 and 0.070 respectively) between the clinicians and the politicians.

Treatment, time preference and patient's age are correlated and could not be used in the same logit model - thus we present five models in the Table **2**. Time preference has a statistically significant impact (OR 1.41, p<0.0001) on the results: the agents are more likely (1.4 odds in comparison to high present ICER) to accept treatments that have low present ICER (i.e., ICER that are affected by discounting more than the average in the data). The incremental costs have insignificant and indifferent effect on the decision

which indicates that the decision is not driven by the explicitly stated incremental costs.

The model adjusting for heterogeneity in treatments (Logit2) has the best fit to the data and is (clinically and statistically) the most relevant basis for the Hurdle-type model for the ICUR, ICER and II outcomes. Based on the significant or suggestively significant variables in the Logit2 model, the clinicians are less likely to accept (odds 0.76 in comparison to the politicians) the futuristic treatment. The futuristic treatment for dementia (odds 0.59) and cancer (odds 0.55) are less accepted, and the futuristic treatment for addiction (odds 2.53), DM (odds 9.04) and paraplegia (odds 9.69) are more accepted; all the odds in comparison to the CHD.

Incremental and Average Per Patient Costs and Benefits

According to the cost models in Table **3** (politician deciding on CHD is the reference group, constant), there are statistically significant differences in the level of mean total per patient costs but not in the level of marginal costs the agents are willing to accept. The clinicians accept statistically significantly lower mean costs than the politicians. Accordingly, the politicians accept higher incremental costs than the clinicians. In all cost outcomes, ln OLS has the best fit to the data and is used in the Hurdle model.

Table **4** (politician deciding on CHD is the reference group, constant) reveals that there are no statistically significant differences in the benefits the agents accept. The

Table 3. Different Types of Linear Models for the Accepted Average and Marginal Per Patient Lifetime Costs (€, 2008 value). Semi Logarithmic (OLS ln) Models Provide the Best Fit to the Average and Marginal per Patient Cost Data

Costs	Average costs 3% (N 1092)			Marginal costs 3% (N 501)			Average costs 0% (N 1092)			Marginal costs 0% (N 501)		
Parameter[#]	*OLS*	*OLS ln*	*GLM*	*OLS*	*OLS ln*	*GLM*	*OLS*	*OLS ln*	*GLM*	*OLS*	*OLS ln*	*GLM*
Constant	36,940**	9.8959***	10.3234***	70,763***	10.4982***	10.8812***	42,326*	9.9698***	10.3715***	81,146	10.4947***	10.8962***
Clinician	-23,442	-0.1114*	-0.1088	-79,582	0.01332	-0.0074	-3,5978	-0.1086*	-0.1095	-122,975	0.0107	-0.0127
Dement	57,416***	1.4037***	1.0980***	56,991***	1.0078***	0.7455***	65,298***	1.4423***	1.1599***	65,419***	1.0801***	0.8185***
Addiction	24,901***	0.4281***	0.5879***	31,097**	0.6664***	0.4126***	26,887***	0.4470***	0.5892***	34,804**	0.6978***	0.4268***
Cancer	6,135	0.3063***	0.1754	30,765*	0.5691***	0.3463*	5,079	0.2314***	0.1276	34,845*	0.5735***	0.3329*
Neurol.	397,606***	2.4655***	2.6798***	759,321***	3.0608***	2.7172***	518,154***	2.5920***	2.8830***	1,015,616***	3.3409***	2.9768***
Metabolic	2,197,436***	4.3636***	4.3325***	3,539,235***	4.5034***	4.2112***	3,094,782***	4.6027***	4.6233***	5,092,382***	4.8558***	4.5575***
DM	93,660***	1.6078***	1.4322***	77,775***	1.1087***	0.8560***	105,564***	1.7495***	1.4882***	61,180***	0.8763***	0.6763***
Paraplegia	92,735***	1.6402***	1.4238***	64,889***	0.7241***	0.6507***	127,515***	1.9090***	1.6310***	74,388***	0.7249***	0.6508***
Tests												
Chi^2	72.13***	269.15***		103.18***	545.30***		88.99***	318.94***		106.75***	635.42***	
R^2	0.4864	0.6673		0.8123	0.8816		0.4738	0.7155		0.7960	0.8215	
LL	16,108.84	1,392.15	13,718.60	7,233.20	452.12	6,501.26	16,512.07	1,341.42	13,855.06	7,442.44	485.25	6,520.66
AIC	32,235.68	2,802.31	27,455.20	14,484.40	922.23	13,020.53	33,042.15	2,700.83	27,728.13	14,902.87	988.51	13,059.32
BIC	32,280.64	2,847.27	27,500.16	14,522.35	960.18	13,058.48	33,087.11	2,745.79	27,773.09	14,940.82	1,026.46	13,097.27

OLS = ordinary least squares regression model. OLS ln = OLS model with logarithmic dependent variable. GLM = generalized linear model (gamma distribution and logistic link). SE = standard error. N = number of observations. Chi^2 = Chi^2 test. R^2 = coefficient of determination. LL = Log likelihood. AIC = Akaike information criteria. BIC = Baysian (Swartz) information criteria. DM = Diabetes mellitus. CHD = Coronary heart disease. 3% = discounting with 3% per annum. 0% = not discounted. [#] Politician deciding on CHD is the reference group. *** p value<0.001, ** p value<0.010, * p value<0.050. OLS ln models were used to adjust the lifetime costs in the Hurdle model.

Table 4. OLS Models for the Accepted Average and Marginal per Patient Benefits during Lifetime

Model	15D QALYs				EQ-5D QALYs				LYGs			
Parameter[#]	Av 3%	Incr 3%	Av 0%	Incr 0%	Av 3%	Incr 3%	Av 0%	Incr 0%	Av 3%	Incr 3%	Av 0%	Incr 0%
Constant	4.7972***	0.8011***	5.3299***	0.9876***	4.5606***	0.7617***	5.0667***	0.9392***	5.4898***	0.9166***	6.0988***	1.1301
Clinician	-0.0378	0.0251	-0.0460	0.0348	-0.0339	0.0235	-0.0406	0.0326	-0.0466	0.0284	-0.0542	0.0395
Dement	-0.8434***	-0.2407**	-0.9839***	-0.3195***	-0.8687***	-0.2384***	-1.0088***	-0.3152***	-0.5539***	-0.2164*	-0.6738***	-0.2952*
Addiction	4.5520	0.6152***	6.0915***	1.1092***	4.3042***	0.5813***	5.7626***	1.0493***	5.0282***	0.6764***	6.7496***	1.2284***
Cancer	-4.4360***	-0.5536***	-4.9659***	-0.7435***	-4.2253***	-0.5319***	-4.7292***	-0.7123***	-5.0768***	-0.6344***	-5.6831***	-0.8516***
Neurologic	0.9295***	1.2436***	1.1621***	1.6191***	0.5883***	1.0763***	0.7702***	1.4041***	0.9316***	1.3755***	1.1802***	1.7920***
Metabolic	6.1542***	3.9825***	10.4621***	7.9981***	5.6723***	3.8734***	9.6964***	7.6975***	11.3863***	3.7947***	18.1157***	8.8537***
DM	21.8117***	-0.1969***	58.6127***	0.4720***	21.0117***	0.4253***	56.3854***	1.9211***	23.6477***	-0.9251***	63.9177***	-1.1419***
Paraplegia	18.7283***	-0.3230***	40.3754***	-0.0526	18.0382***	0.3088***	38.8386***	1.1469***	20.2590***	-0.9288***	43.9233***	-1.1470***
Tests												
N	1,092	501	1,092	501	1 092	501	1,092	501	1,092	501	1,092	501
Chi[2]	12,566***	557***	26,186***	1,013***	12,422***	512***	24,945***	1,027***	14,150***	564***	25,649***	1,218***
R[2]	0.9893	0.9006	0.9949	0.9428	0.9892	0.8928	0.9946	0.9435	0.9905	0.9016	0.9947	0.9519
LL	1,432.83	236.93	2,004.23	388.22	1,399.93	195.24	1,989.94	347.42	1,475.12	297.17	2,111.66	448.19
AIC	2,883.65	491.86	4,026.47	794.44	2,817.87	408.48	3,997.88	712.84	2,968.23	612.35	4,241.31	914.38
BIC	2,928.61	529.80	4,071.43	832.39	2,862.83	446.43	4,042.85	750.79	3,013.19	650.30	4,286.27	952.33

OLS = ordinary least squares regression model. QALY = quality-adjusted life-year. LYG = life-year gained. Av = average total per patient. Incr = incremental per patient. 3% = discounting with 3% per annum. 0% = not discounted. N = number of observations. ns = not statistically significant. Chi[2] = Chi[2] test. R[2] = coefficient of determination. LL = Log likelihood. AIC = Akaike information criteria. BIC = Baysian (Swartz) information criteria. DM = Diabetes mellitus. CHD = Coronary heart disease. 3% = discounting of costs and effects with 3%. 0% = costs and effects were not discounted. [#] Politician deciding on CHD is the reference group. *** p value<0.001, ** p value<0.010, * p value<0.050.

clinicians accept lower mean total per patient benefits and higher incremental benefits than the politicians.

Willingness to Pay

According to the adjusted WTP results presented in Table **5** for all respondents, the discounted (undiscounted) incremental WTPs per QALY gained are €102,616 (€78,686) and €94,770 (€77,856) valued with 15D and EQ-5D respectively. The respective aggregate WTPs per QALY are €16,062 (€15,843) and €17,277 (€17,045). The discounted (undiscounted) incremental WTP per LYG is €66,277 (€58,160) and the aggregate WTP per LY is €13,847 (€13,679). On an aggregate level, this means that the agents are willing to invest some €13,700 for a year of treatment (value of life-year from health care payer perspective) or at least €15,900 for treatment resulting in one QALY.

Incremental outcomes (ICUR, ICER) are more sensitive to discounting compared to aggregate outcomes (SCUR, SCER). On average, the undiscounted WTPs are lower than the discounted WTPs due to the fact that a significant part of the disease/treatment related costs occur (peak) early meanwhile effectiveness occurs over time (for most scenarios). This explains why incremental outcomes are more sensitive to the discounting. The discounting of costs and outcomes is rarely accounted for in a WTP study. However, this phenomenon of lower undiscounted ICURs or ICERs in comparison to the discounted ICURs or ICERs is commonly observed in cost-effectiveness analyses.

The highest incremental WTPs per QALY gained (ICUR) and per LYG (ICER) are reported for cancer (€205,994–250,509 per QALY gained; €180,666–204,095 per LYG) and lowest for metabolic (€23,492–43,398 per QALY gained; €21,188–42,656 per LYG) treatment (Table **5**). When measuring aggregate outcomes, the highest aggregate WTP per QALY (SCUR) and per LY (SCER) are also reported for cancer (€77,329–79,516 per QALY; €62,781–65,306 per LY), but the lowest values are reported for treatment of diabetes (€825–2,024 per QALY; €753–1,776 per LY). On an aggregate level, this means that the agents are willing to invest some €1,200–64,000 for a year of treatment depending on the disease or treatment, or €1,400–78,000 for treatment resulting in a year of perfect QoL or one QALY depending on the disease or treatment.

Generally, the adjusted WTP results are comparable between clinicians and politicians (Table **5**). The politicians are willing to pay more for treatments of CHD, dement, cancer, addiction, DM and paraplegia, and less for neurologic and metabolic treatments than the clinicians. However, there seemed to be some differences in outcomes when different QoL tools are used. The most evident differences are observed for treatments of paraplegia and DM - approximately two times higher WTP per QALY gained is estimated for 15D than for EQ-5D. This indicates that higher ICUR values can be accepted for 15D-based results compared to EQ-5D-based results in the cost-effectiveness analyses of DM or paraplegia treatments, and lower ICUR values can be accepted for 15D-based results compared to EQ-5D-based results

Table 5. **Hurdle-Type Model Result Summary: Average Maximum Willingness to Pay for Different Outcomes (€, Annual Discounting Rate is Given)**

Outcome	ICUR (15D)		ICUR (EQ-5D)		SCUR (15D)		SCUR (EQ-5D)		ICER		SCER		II	
Clinician	*3%*	*0%*	*3%*	*0%*	*3%*	*0%*	*3%*	*0%*	*3%*	*0%*	*3%*	*0%*	*3%*	*0%*
CHD	44,464	35,712	46,786	37,572	3,731	3,628	3,923	3,814	38,874	31,218	3,262	3,172	36,736	36,512
Dement	125,976	108,051	134,892	115,670	10,899	10,889	11,668	11,655	101,234	86,858	8,729	8,718	73,759	75,949
Addiction	42,471	29,082	44,799	30,672	2,723	2,439	2,871	2,571	37,756	25,851	2,422	2,168	61,217	61,990
Cancer	211,456	205,994	227,568	221,394	71,723	74,240	76,958	79,516	185,586	180,666	63,306	65,306	57,643	57,452
Neurologic	72,074	60,002	80,139	66,710	10,817	10,683	12,030	11,881	64,287	53,517	9,653	9,532	149,179	158,496
Metabolic	42,043	23,702	43,398	24,663	8,727	6,822	9,338	7,296	42,656	21,331	5,659	4,446	202,174	213,808
DM	123,098	45,843	63,995	23,681	1,743	825	1,813	858	na	na	1,592	753	77,466	68,508
Paraplegia	125,868	64,941	57,895	29,726	1,996	1,221	2,078	1,271	na	na	1,824	1,116	63,337	62,980
Politician	*3%*	*0%*	*3%*	*0%*	*3%*	*0%*	*3%*	*0%*	*3%*	*0%*	*3%*	*0%*	*3%*	*0%*
CHD	45,251	36,577	47,591	38,462	4,138	4,010	4,352	4,218	39,549	31,965	3,616	3,504	36,250	36,124
Dement	129,877	112,469	139,085	120,418	12,067	12,010	12,923	12,863	103,946	89,999	9,666	9,621	72,783	75,141
Addiction	42,652	29,250	44,979	30,843	3,032	2,708	3,198	2,856	37,920	26,004	2,695	2,407	60,407	61,331
Cancer	229,819	232,857	247,520	250,509	71,784	72,298	77,329	77,975	201,560	20,095	62,781	63,306	56,880	56,840
Neurologic	71,993	60,156	80,090	66,918	12,011	11,825	13,359	13,152	64,223	53,663	10,712	10,546	147,205	156,809
Metabolic	41,705	23,541	43,041	24,492	9,721	7,582	10,404	8,111	42,345	21,188	6,308	4,945	199,499	211,532
DM	126,516	46,436	64,398	23,696	1,945	919	2,024	956	na	na	1,776	839	76,441	67,779
Paraplegia	130,724	66,641	58,383	29,869	2,228	1,360	2,319	1,416	na	na	2,035	1,243	62,499	62,309
All	*3%*	*0%*	*3%*	*0%*	*3%*	*0%*	*3%*	*0%*	*3%*	*0%*	*3%*	*0%*	*3%*	*0%*
Total	102,616	78,686	94,770	77,856	16,062	15,843	17,277	17,045	66,277	58,160	13,847	13,679	83,886	85,398
Clinicians	103,627	78,544	95,774	77,589	16,276	16,239	17,483	17,427	66,864	57,857	14,087	14,054	83,327	84,942
Politicians	104,736	81,202	96,660	80,467	16,340	15,938	17,600	17,190	67,711	60,212	14,033	13,728	87,076	88,531

SCER = aggregate cost-effectiveness ratio. ICER = incremental cost-effectiveness ratio. SCUR = aggregate cost-utility ratio. ICUR = incremental cost-utility ratio. 3% = discounting of costs and effects with 3%. 0% = costs and effects were not discounted. II = incremental investment. DM = Diabetes mellitus. CHD = Coronary heart disease.

in the cost-effectiveness analyses of CHD, dementia, addiction, cancer, neurologic or metabolic disease treatments.

Affordability results given as II to health care are given in Table **5**. The adjusted discounted (undiscounted) incremental investments (II) in health care are €83,886 (€85,398) per life. The highest IIs per life are reported for metabolic (€199,499–213,808) and lowest for CHD (€36,124–36,736) treatment. Table **5** also indicates that there is a trade-off between affordability (II) and efficiency (ICER or ICUR).

DISCUSSION

This CV with DC questions on eight new treatments carried out for Finnish clinicians and politicians reveals that the agents attach different levels of WTP to different treatments. The differences are considerable, ranging from the metabolic treatment at €24,000–43,000 per QALY gained to the cancer treatment with €206,000–251,000 per QALY gained depending, for example, on the discounting, agent and QoL tool. The incremental WTP based on the whole data ranges from €78,000 to 103,000 per QALY gained depending on the discounting and QoL tool.

Quite a few monetary threshold values have been set for QALY and sometimes it is hard to comprehend from the guidelines whether this threshold is, for example, defined for ICER or CER, and whether this threshold is absolute limit or maximum mean WTP - a general €50,000 per QALY based on the dialysis argument, a £20,000-30,000 per QALY gained suggested by the National Institute for Health and Clinical Excellence (NICE), €80,000 per QALY proposed in the Netherlands and two times gross domestic product (GDP) per person per year. The total result of this study is in line with the threshold the Netherlands has set and close to 2xGDP (some €70,000 in Finland). However, based on this study a fixed setting of cost per QALY threshold is questionable, at least as a single WTP for all diseases and treatments.

Setting an exact limit in a monopoly industrial environment would also lead to a situation where every new treatment would be priced to be just on the threshold level. There is no reason why dialysis has been set as a comparator of other treatments and, as is demonstrated here, different patient groups can justify different cost per QALY gained limits. For example, if the patient is a child, the fair innings

principle would make higher cost per QALY gained values acceptable, which is supported by this study. The acceptance of expensive treatments may also be easier if the disease is rare and only a few patients need the treatment. The questionnaire used in this study lacked information about the prevalence of disease and we could not directly assess the impact of disease rarity. The WTPs in different patient groups, however, could indicate that such a dependency is not implicit or a strong one. However, the present study supports the general idea that the severity of the disease can have an impact on the decision. Accordingly, it would be important to evaluate these potential dependencies in future research.

However, dependencies between WTP and the patient's age at baseline, time preference, QoL tool and the manner in which the results are estimated are found. The level of WTP changes when patients have different ages or when discounting, different health units, different QoL tools or different outcome presentation methods are used. Generally, incremental outcomes presented higher WTPs than aggregate outcomes. Also, the ICURs are higher than the correspondent ICERs and, generally, the undiscounted WTPs are lower compared to the discounted WTPs. This is reasonable as the benefit side in ICUR includes qualitative dimensions, whereas the benefit side in ICER presents mere quantities, and a significant part of the treatment related costs occur early meanwhile effectiveness for the most scenarios occurs over time.

Generally, an ICER approach should be preferred over CER. Actually, a CER type of analysis should result in a cost-consequence analysis in comparison with mutually independent treatments or in an aggregate description of health economic data, as was done in the present study in terms of SCURs and SCER. Incremental outcomes (ICUR, ICER) can be considerably higher than aggregate outcomes (SCUR, SCER) or CERs, especially if the disease and its treatment have a long duration.

In this study, the incremental costs have an insignificant and indifferent effect on the decision the agents make. The WTPs attached to benefits by clinicians and politicians are nearly identical: only the acceptance rate for additional futuristic treatment is significantly higher among the politicians. The QoL tool has a much more significant impact on the results: in paraplegia and DM approximately two times higher accepted maximum WTP per QALY gained is estimated by the 15D than by the EQ-5D. The utility data used for QALY estimation with EQ-5D and 15D come from the same nationwide population, which makes the comparison between EQ-5D- and 15D-based WTPs relatively valid. More effort should be allocated to studying the differences and validity of QoL measurement tools in different patient populations.

Accepted IIs per life (affordability) range from €36,000 to 202,000 depends on the discounting, treatment and agent. On average, the agents are willing to invest an additional €83,886 per patient in health care for the treatment of a disease, which, naturally, may be an over-exaggeration due to our somewhat rare and severe disease mix. For example, for CHD, lifetime II is some €36,000. This corresponds to recent cost estimations among the Finnish CHD population [21]. However, there seems to be a trade-off between affordability and efficiency - if ICER was low, II was high, and *vice versa*. Thus future research should study this potential trade-off in more detail. The time preference and its relationship to WTP and affordability would also make an excellent study topic.

Previous CV and, especially, DC studies are limited in number. Typically, just one patient group has been included in the CVs. Bradford *et al.* [50] has compared the WTP estimates for a new telemedicine technology in two populations: patients who are treated for chronic heart failure (CHF) and patients who are treated for hypertension. The analysis indicates a negative relationship between price and the likelihood of purchase, and patients with CHF are less responsive to price changes than those with hypertension [50]. In this study, the incremental cost and acceptance have a very weak non-significant inverse association. As is also observed in the present CV, this former result supports the severity of disease argument related to WTP.

Only a few health care WTP assessments have taken place in Finland. The present study is the first to sample the WTP of Finnish clinicians and politicians, and the present study also includes a variety of treatments and outcomes. An unpublished study directed at 1,000 Finnish clinicians has resulted in a WTP of €66,000 per QALY, but the result is based on a low 22% response rate to the WTP question [Jormanainen V *et al.*: Finnish Health Economist Convention 2008, Lecture]. The response rate could be negatively influenced by the direct question of WTP for QALY (some respondents may not answer when open ended questions are used). However, the undiscounted ICURs obtained in this study (some 78,000 per QALY gained) are close to those obtained by Jormanainen *et al.*

In an earlier Finnish study [51] concerning the attitudes to cuts in expenditure and increased fees in health care, all groups that have been included indicated the greatest willingness to cut expenditure on health education, occupational health services, hygiene inspection, substance abuse care, rehabilitation services for war veterans, and family planning. All of the groups are least willing to make cuts in home care for disabled and elderly people, maternity services and clinics for under-fives. Primary health care and prevention of diseases for small children, mothers, the elderly and disabled people, are prioritized by all of the groups [51]. The present study is in line with these earlier results.

Result Usability and Study Limitations

Preferences are stated in our study. The main reasons for using stated preferences is the assessment of hypothetical future treatment (i.e., revealed preferences are impossible to derive), the principal agent relationship, the asymmetric information and respective uncertainty between the principal and agent, and the design, which is targeted to enable direct estimation of the outcomes. DCs also enable the collection of high-quality data at low cost, allow characterization of the value of different attributes, define and describe the limits of the utility function, and the utility function values different characteristics of treatment differently and derive the WTP indirectly.

Even after the Hurdle-type regression adjustment, some over or under-sensitivity to the questions may have remained

in the data because the data do not include information about the respondent's age, gender, working environment or socio-economic status. These background questions have been excluded in order to get moderate to good response rates, which are accomplished. However, the coefficients of determination are high, the RESET test do not indicate missing values, and the link tests indicate proper function specifications for the regression models used in the estimation, all of which indicate the sufficiency of the data for the modeling purpose. The stated marginal costs do not have a significant impact on the results, and the mean costs but not the marginal costs that the agents were willing to accept are different. Thus it is likely that factors other than mere cost have an impact on the WTP decisions.

The main theoretical problems with the WTP approach used in this study are related to the fact that it may be impossible to predetermine the maximum WTP per health benefit gained, because the information related to ICERs can be incomplete [52, 53] and imperfect. Also, a dynamic WTP means that when a new treatment with a lower ICER compared to previous treatments is taken to use, the WTP is also likely to change to a lower level [54] and *vice versa*. Professor Allan Williams [55] has stated the problem: must the decision maker make an internal assessment of the WTP per health benefit gained and adjust the budget accordingly?

Considering the usability of the results, Brouwer *et al.* [56] have provide an application of benefits transfer, which focuses on the transfer of CV WTP estimates and associated value functions for reducing the health risks associated with solar ultraviolet exposure. They have showed that when transferring between similar contexts, simple mean-value transfers outperform more complex value function transfers [56]. In addition to the use of hypothetical future scenarios (i.e., the marginal costs are not local), this would suggest that the incremental results of this analysis may be directly transferable to similar contexts. The transferability of results based on a WTP or cost-effectiveness study can be further improved with the purchase power adjustment of results using, for example, the official purchasing power parities.

However, based on recent research, the cost-effectiveness of, for example, newer pharmaceuticals can be significantly distorted by the computational pharmaceutical pricing scheme (PPS) which determines retail prices based on wholesale prices [57]. Various PPSs may decrease the potential for direct extrapolation of the WTP results from one setting to another. In addition, the differences between median outcomes (used for, for example, censored data analysis in some trials) and mean or average outcomes should be noted when interpreting the WTP results presented here. All values presented in this analysis are means or averages and should not be mixed with medians. In the cost-effectiveness analyses of, for example, cancer [58], the estimation and use of mean (the expected value of distribution) can make difference compared to the simple use of median.

When interpreting these results, it must be noted that the clinicians and the politicians are not facing any explicit budget constraints in the study – the budget constraints are assumed to be their own shadow prices – and the obtained WTPs may be somewhat exaggerated. In contrast, these results are of real importance in revealing that different diseases or treatments are likely to have very different social values that agents relate to them, and a single WTP threshold may not have empiric support. In the future, a larger scale analysis with dozens of diseases/treatments would be needed to assess the social value of the diseases/treatments.

CONCLUSION

Two approaches are needed when budgets are adjusted. A cost-effectiveness analysis allows evaluation of "efficiency" and a budget impact analysis evaluates "affordability". According to this study, there are only some minor differences in WTP between clinicians and politicians, but efficiency and affordability can have a trade-off.

WTP is significantly dependent on patient characteristics (disease and age), treatment, time preference, benefits measure and discounting, but is not significantly related to incremental costs. Thus, based on this study, the hypothesis of a single WTP for all treatments or diseases does not gain empirical support.

CONFLICT OF INTEREST

E.J.S. is an employee of and a shareholder in ESiOR Oy. ESiOR carries out studies and health-economic analyses for several pharmaceutical companies, food industry companies, device companies, research institutions and hospitals. The other authors declare no potential conflict of interests. All authorship decisions were made on the basis of scientific consideration.

ACKNOWLEDGEMENT

E.J.S. contributed to the concept and design, and data collection and management, modelling, analysis and interpretation of data, and to drafting and revising the article critically for important intellectual content. J.K. and O.P.R. contributed to the concept and design, data collection and management, and to drafting the article. E.J.S. and J.K. shared an equal contribution. M.M. contributed to the concept and design, and to revising the article critically for important intellectual content. All authors have approved the final version for publication.

REFERENCES

[1]   Karlsson G, Johannesson M. The decision rules of cost-effectiveness analysis. Pharmacoeconomics 1996; 9: 113-20.

[2]   Johannesson M, Weinstein M. On the decision rules of cost-effectiveness analysis. J Health Econ 1993; 12: 459-67.

[3]   Weinstein M, Zeckhauser R. Critical ratios and efficient allocation. J Public Econ 1973; 2: 147-57.

[4]   Hanley N, Ryan M, Wright R. Estimating monetary value of health care: lessons from environmental economics. Health Econ 2003; 12: 3-16.

[5]   Johnson FR, Desvousges WJ, Ruby MC, *et al.* Eliciting stated health preferences: an application to willingness to pay for longevity. Med Decis Making 1998; 18: S57-67.

[6]   O'Brien BJ, Gafni A. When do the dollars make sense? Toward a conceptual framework for contingent valuation studies in health care. Med Decis Making 1996; 16: 288-99.

[7]   Pauly MV. Valuing health care benefits in money terms. In: Sloan FA, Ed. Valuing health care. Cambridge: Cambridge University Press 1995.

[8]   Willan AR, Briggs AH. Statistical analysis of cost-effectiveness data. Chichester: Wiley 2006.

[9]   Fenwick E, Marshall DA, Levy AR, Nichol G. Using and interpreting cost-effectiveness acceptability curves: an example using data from a trial of management strategies for atrial fibrillation. BMC Health Serv Res 2006; 6: 52.

[10]   Fenwick E, Palmer S, Claxton K, *et al.* An iterative bayesian approach to health technology assessment: application to a policy of preoperative optimization for patients undergoing major elective surgery. Med Decis Making 2006; 26: 480-96.

[11]   Eichler H, Kong S, Gerth W, *et al.* Use of cost-effectiveness analysis in health-care resource allocation decision-making: how are cost-effectiveness thresholds expected to emerge? Value Health 2004; 7: 518-28.

[12]   Birch S, Gafni A. Cost effectiveness/utility analysis: do current decision rules lead us to where we want to be? J Health Econ 1992; 11: 279-96.

[13]   Weinstein MC, Stason WB. Foundation of cost effectiveness analysis for health and medical practices. New Engl J Med 1977; 296: 716-21.

[14]   Birch S, Gafni A. Changing the problem to fit the solution: Johannesson's and Weinstein's (mis)application of economics to real world problems. J Health Econ 1993; 12: 469-76.

[15]   Briggs A, Fenn P. Trying to do better than average: a commentary on 'statistical inference for cost-effectiveness ratios. Health Econ 1997; 6: 491-5.

[16]   Drummond M, Sculpher M. Common methodological flaws in economic evaluations. Med Care 2005; 43: S5-14.

[17]   Stinnett A, Paltiel A. Estimating CE ratios under second-order uncertainty: the mean ratio versus the ratio of means. Med Decis Making 1997; 17: 483-9.

[18]   Laska EM, Meisner M, Siegel C. The usefulness of average cost-effective ratios. Health Econ 1997; 6: 497-504.

[19]   Soini E. Incremental or average cost-utility of routine cataract surgery? Health Qual Life Outcomes 2006; 4: 74. Available from: http://www.hqlo.com/content/4/1/74/comments

[20]   Barton GR, Briggs AH, Fenwick EAL. Optimal cost-effectiveness decisions: the role of the cost-effectiveness acceptability curve (CEAC), the cost-effectiveness acceptability frontier (CEAF), and the expected value of perfection information (EVPI). Value Health 2008; 11: 886-97.

[21]   Soini EJ, Davies G, Martikainen JA, *et al.* Population-based health-economic evaluation of the secondary prevention of coronary heart disease in Finland. Curr Med Res Opin 2010; 26: 25-36.

[22]   Hallinen T, Soini EJ, Eklund K, Puolakka K. Cost-utility of different treatment strategies after the failure of tumor necrosis factor inhibitor in rheumatoid arthritis in the Finnish setting. Rheumatology (Oxford) 2010; 49: 767-77.

[23]   Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. Health Econ 2001; 10: 779-87.

[24]   Phillips KA, Johnson FR, Maddala T. Measuring what people value: a comparison of "attitude" and "preference" surveys. Health Serv Res 2002; 37: 1659-79.

[25]   Samuelson PA. Foundations of economic analysis. Cambridge: McCraw-Hill 1948.

[26]   Singh J, Cuttler L, Shin M, *et al.* Medical decision-making and the patient: understanding preference patterns for growth hormone therapy using conjoint analysis. Med Care 1998; 36: AS31-45.

[27]   Louviere JJ, Hensher D, Swait JD. Stated choice methods: Analysis and application. Cambridge: Cambridge University Press 2000.

[28]   Lancaster K. A new approach to consumer theory. J Polit Econ 1966; 74: 132-57.

[29]   Simon HA. A behavioral model of rational choice. Q J Econ 1955; 69: 99-118.

[30]   Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. Econometrica 1979; 47: 263-91.

[31]   Hanley N, Mourato S, Wright R. Choice modeling approaches: a superior alternative for environmental evaluation. J Econ Surv 2001; 15: 453-557.

[32]   Smith RD. Sensitivity to scale in contingent valuation: the importance of the budget constraint. J Health Econ 2005; 24: 515-29.

[33]   Thurstone LL. A law of comparative judgment. Psychol Rev 1927; 34: 273-86.

[34]   McFadden D. Conditional logit analysis of qualitative choice behavior. In: Zarembka P, Ed. Frontiers in econometrics. New York: Academic Press 1974.

[35]   Hujanen T, Kapiainen S, Tuominen U, Pekurinen M. Terveydenhuollon yksikkökustannukset Suomessa vuonna 2006. Helsinki: Stakes Työpapereita 2008.

[36]   Rawlings MD. Five NICE years. Lancet 2005; 9462: 904-8.

[37]   Buxton MJ. How much are health-care systems prepared to pay to produce a QALY? Eur J Health Econ 2005; 6: 285-7.

[38]   Rawlins MD, Culyer AJ. National Institute for Clinical Excellence and its value judgments. BMJ 2004; 329: 224-7.

[39]   Brooks R. EuroQol: the current state of play. Health Policy 1996; 37: 53-72.

[40]   Sintonen H. The 15D Instrument of health-related quality of life: properties and applications. Ann Med 2001; 33: 328-36.

[41]   Saarni S, Suvisaari J, Lönnqvist J, *et al.* The impact of 29 chronic conditions on health-related quality of life: a general population survey in finland using 15D and EQ-5D. Qual Life Res 2006; 15: 1403-14.

[42]   Heilbron D. Zero-altered and other regression models for count data with added zeros. Biometrical J 1994; 36: 531-47.

[43]   Mullahy J. Specification and testing of some modified count data models. J Econom 1986; 33: 341-65.

[44]   Soini EJ, Hallinen TA. Cost-utility of agomelatine, venlafaxine and placebo in the treatment of major depressive disorder (MDD) in Finland - Economic modeling study using representative population data. Value Health 2009; 12: A359.

[45]   Hallinen T, Martikainen JA, Soini EJ, *et al.* Direct costs of warfarin treatment among patients with atrial fibrillation in a Finnish health care setting. Curr Med Res Opin 2006; 22: 683-92.

[46]   Keränen J, Soini E, Ryynänen OP, Keränen U. Economic evaluation comparing From Home To Operation same day admission and preoperative admission one day prior to surgery process: a randomised, controlled trial of laparoscopic cholecystectomy. Curr Med Res Opin 2007; 23: 2775-84.

[47]   Burnham KP, Anderson DR. Multimodel inference - Understanding AIC and BIC in model selection. Social Methods Res 2004; 33: 261-304.

[48]   Cherkassky V, Ma Y. Comparison of model selection for regression. Neural Comput 2003; 15: 1691-714.

[49]   Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. J R Stat Soc Series B Stat Methodol 2002; 64: 583-639.

[50]   Bradford WD, Kleit A, Krousel-Wood MA, Re RM. Comparing willingness to pay for telemedicine across a chronic heart failure and hypertension population. Telemed J E Health 2005; 11: 430-8.

[51]   Myllykangas M, Ryynänen OP, Kinnunen J, Takala J. Attitudes to cuts in expenditure and increased fees in health care. Public Health 1997; 111: 71-5.

[52]   Gafni A, Birch S. NICE methodological guidelines and decision making in the National Health Service in England and Wales. Pharmacoeconomics 2003; 21: 149-57.

[53]   Devlin N. An introduction to the use of cost-effectiveness thresholds in decision-making: what are the issues? In: Towse A, Pritchard C, Devlin N, Eds. Cost effectiveness thresholds: economics and ethical issues. London: Kings Fund and Office of Health Economics 2002.

[54]   Sendi P, Gafni A, Birch S. Opportunity costs and uncertainty in the economic evaluation of health care interventions. Health Econ 2002; 11: 23-31.

[55]   Williams A. What Could Be Nicer than NICE? London: Office of Health Economics 2004.

[56]   Brouwer R, Bateman IJ. Benefits transfer of willingness to pay estimates and functions for health-risk reductions: a cross-country study. J Health Econ 2005; 24: 591-611.

[57]   Hallinen T, Soini EJ. The impact of the pharmaceutical pricing system on cost-effectiveness results: Finnish analysis. The Open Pharmacoeconomics Health Econ J 2011; 3: 6-10.

[58]   Soini EJ. Cost-utility and expected value of perfect information related to trabectedin in the treatment of metastatic soft-tissue

sarcoma: the publicly funded comments explored. Ann Oncol   2011; 22: 1465-6.