VALUE IN HEALTH 15 (2012) 475–484

Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/locate/jval

QALY Weights for Diabetic Retinopathy—A Comparison of Health State Valuations with HUI-3, EQ-5D, EQ-VAS, and TTO

Emelie Heintz, MSc[1,], Ann-Britt Wiréhn, PhD[2], Beatrice Bourghardt Peebo, MD[3], Ulf Rosenqvist, MD[4], Lars-Åke Levin, PhD[1]*

[1]*Center for Medical Technology Assessment, Department of Medical and Health Sciences, Linköping University, Linköping, Sweden;* [2]*Local Health Care Research and Development Unit, Östergötland County Council, Linköping University, Linköping, Sweden;* [3]*Department of Ophthalmology, Linköping University Hospital, Linköping, Sweden;* [4]*Department of Internal Medicine and Endocrinology, Motala Hospital, Motala, Sweden*

A B S T R A C T

Objective: To estimate quality-adjusted life-year weights for patients with diabetic retinopathy by using various methods and to investigate the empirical validity of the different measures. **Methods:** The study population comprised 152 patients with diabetes in Östergötland County, Sweden. Participants were interviewed by telephone by using the time trade-off (TTO) method and a visual analogue scale (EQ-VAS) (direct valuations) as well as the EuroQol five-dimensional questionnaire (EQ-5D) and the health utilities index mark 3 (HUI-3) (indirect valuations). The quality-adjusted life-year weights were adjusted for potential confounders by using analysis of covariance. The empirical validity of the measures was examined by testing their ability to detect hypothetical differences between severity levels of diabetic retinopathy and by investigating the correlation between the measures and the 25-item National Eye Institute Visual Function Questionnaire (NEI VFQ-25). **Results:** All measures detected significant differences in scores between patient groups classified according to visual impairment in the better eye (analysis of covariance, $P < 0.05$), but only HUI-3 and EQ-VAS detected significant differences between patient groups classified according to visual impairment or pathological progression in the worse eye. HUI-3 recorded a difference of 0.43 in values between normal vision and blindness in the better eye, which was more than twice the differences captured by the other measures (0.15–0.20). In addition, HUI-3 showed the highest correlation with NEI VFQ-25 ($r = 0.54$; $P < 0.001$). **Conclusions:** In cost-utility analyses, the choice of quality-adjusted life-year measure may affect whether an intervention is considered cost-effective. Furthermore, if decisions are to be based on values from the general public, HUI-3 can be recommended for cost-utility analyses of interventions directed at diabetic retinopathy.

Keywords: diabetes mellitus, diabetic retinopathy, EQ-5D, health utility index mark 3, NEI-VFQ-25, QALY, time trade-off, TTO, VAS.

Introduction

In cost-effectiveness analyses used to inform decision makers about the cost-effectiveness of alternative interventions, quality-adjusted life-years (QALYs) have been recommended as the main effect outcome measure [1,2]. By combining the effect on survival and health-related quality of life (HRQOL) into one measure and enabling comparisons across different disease areas, QALYs have become a common measure of health outcome and treatment effect. To calculate QALYs, the time in a specific health state is multiplied by a score representing the value of that specific health state (death is represented by 0 and full health by 1).

The value of health states can be estimated by using various measures. These measures include direct valuation methods such as standard gamble (SG) [3], time trade-off (TTO) [4], and the visual analogue scale (VAS) [5] as well as indirect questionnaire-based measures such as the health utilities index mark 3 (HUI-3) [6], the EuroQol five-dimensional questionnaire (EQ-5D) [7], the 15-D measure [8], and the six-dimensional health state short form (SF-6D, derived from short form 36 health survey) [9]. There are various methodological differences between these measures. For example, when used on patient populations, the values estimated with the direct and indirect measures represent different respondent perspectives. The direct methods are used to capture values that patients assign to their own health state, whereas the indirect measures use published tariffs to assign values of the general public to patients' descriptions of their health. In addition, the indirect measures differ in what dimensions their questionnaires include, how many response levels each question has, and the direct valuation method used to create the tariff. Concerning the direct methods, the most important difference is that the values of the TTO and SG methods are elicited by asking respondents to choose between different scenarios, while the VAS is based on a rating exercise. However, there are also methodological differences between TTO and SG.

Despite the dissimilarities between the measures, in practice they are often used interchangeably. For instance, in Sweden

Beatrice Bourghardt Peebo has previously worked as a medical advisor for Pfizer and has given lectures for Novartis and Alcon. Ulf Rosenqvist has lectured and worked as a consultant at pharmaceutical companies such as MSD, Sanofi-Aventis, Lilly, Novo Nordisk, Pfizer, AstraZeneca, and Bristol-Myers Squibb. None of the authors has shares in or is employed by these companies.

* *Address correspondence to:* Emelie Heintz, Center for Medical Technology Assessment, Department of Medical and Health Sciences Linköping University, SE 581 83 Linköping, Sweden.

E-mail: emelie.heintz@liu.se.

1098-3015/$36.00 – see front matter Copyright © 2012, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

doi:10.1016/j.jval.2011.11.031

pharmacoeconomic guidelines state that patient values elicited by using direct choice-based methods are preferred [10]; however, in the absence of these, decision makers accept valuations from the general public and from any of the above-mentioned direct or indirect methods. If the measures give different results, the choice of measure could, when measuring the effectiveness of an intervention, play an important role, which could potentially have an impact on whether an intervention is considered cost-effective. Various studies have suggested that the measures do give different results (e.g., [11–15]). Though, because the performance of the measures may depend on the characteristics of the health state being valued, there is a need to investigate how these measures behave in specific disease areas.

In the previous literature concerning QALY weights for the diabetic eye complication diabetic retinopathy (DR), patient valuations with TTO have been the most commonly used method [16]. Most of these studies, however, used vision-specific questions in which patients were asked to trade life years for full vision rather than full health. This vision-specific version operates on another scale and has previously been argued to overestimate the decrements in HRQOL related to visual impairment compared with conventional methods [17,18]. Relatively few studies have used indirect measures to estimate QALY weights for DR [16]. Thus, to be able to populate decision models in line with guidelines in different countries, there is still a need for QALY weights for DR based on patient as well as general public values. These weights should be classified according to the pathological progression of DR, because this is the dominating primary end point in clinical studies [19–21]. Few studies have compared the performance of measures among patients with DR, and results have varied [22,23]. The aim of the current article, therefore, was to elicit QALY weights for DR and to investigate the empirical validity of QALY measures in patients with different severity levels of DR.

Methods

Study population

The study population consisted of 152 patients with diabetes (types 1 and 2) living in the county of Östergötland, Sweden. In this county, all patients with diabetes are included in a screening program for DR. The study included patients attending these screening examinations at the eye clinic of Linköping University Hospital in 2008 to 2009. Because patients with the most severe vision loss are normally excluded from the screening program, we recruited these patients through two low vision rehabilitation centers in Östergötland. To be eligible for inclusion in the study, participants had to have had diabetes for at least 1 year and be at least 18 years old. After giving informed consent, including access to their medical records, patients with cataract, glaucoma, or other eye disease compromising visual acuity (VA) were excluded from the study. Other exclusions were those with dietary treatment only, patients with Alzheimer's disease, dementia, or difficulty understanding instructions because of very old age, patients with severe concomitant disease that could affect the quality of life (as judged by the physician who included patients in the study), and patients who did not have a sufficient level of understanding Swedish to complete the interview. In the inclusion process, we strived for an equal distribution of types 1 and 2 diabetes. The study was approved by the Regional Ethical Review Board of the county of Östergötland in Linköping (DNR: M115-08).

Patient classification

Patients were, based on the worse eye, recruited to one of five severity groups representing the following health states: no DR (Early Treatment Diabetic Retinopathy Study level 10), background retinopathy (BR; levels 20–53), proliferative diabetic retinopathy (PDR; level 60+ or treated earlier with panretinal photocoagulation), diabetic macular edema (DME; defined in accordance with Early Treatment Diabetic Retinopathy Study guidelines for clinically significant macular edema or treated earlier for clinically significant macular edema), and legal blindness (Snellen VA $\leq$ 20/200 in the better eye). These health states have previously been used in decision models [24–27]. Legal blindness represented the end point of the disease. A sample size calculation (α = 5% and 1 – β = 80%) via the *F* distribution was performed on the basis of previously published data on mean values (no DR 0.94, BR 0.87, PDR 0.83, legal blindness 0.81) and standard deviation (SD) (0.14) [28]. The calculation gave a sample size of at least 25 participants in each severity level. Grading of the patients into the severity levels was performed by an ophthalmologist on the basis of photographs taken during the screening appointment. Dilated pupil, red-free fundus photography (Topcon Retina Camera T3C 50 IX, Topcon Corporation, Tokyo, Japan) was performed in both eyes in the following fields: optic nerve, central fundus, and temporal, superial, and inferior retina. If both PDR and DME were present in the same eye, the patient was classified as having DME, because this was assumed to affect patients' VA more than PDR. Besides the classification based on pathological progression of the disease, patients were also divided into four groups on the basis of their Snellen VA in the better eye (using *International Classification of Diseases, Ninth Revision, Clinical Modification*/World Health Organization criteria): normal vision (defined as VA 20/10–20/25), mild visual impairment (VA 20/32–20/63), moderate visual impairment (VA 20/80–20/160), and severe visual impairment (VA $\leq$ 20/200). The VA was checked during the screening appointment. In line with standard procedure, a trained ophthalmic nurse started each screening visit by assessing best-corrected VA starting with the right eye, followed by the left eye, using a 5-m Snellen chart.

Data collection

Included patients were contacted and interviewed over the telephone. Each interview lasted approximately 45 minutes and included two direct valuation methods, TTO and EQ-VAS, and two indirect generic multiattribute instruments, HUI-3 and EQ-5D. In addition, the 25-item National Eye Institute Visual Function Questionnaire (NEI-VFQ-25) was used to collect vision-specific information. The instruments were administered in the same order for all patients. The order was the following: HUI-3, TTO, EQ-5D, EQ-VAS, and NEI-VFQ-25. All interviews were performed by the same trained researcher (E.H.). Background information and data on the duration of diabetes, type of diabetes, and other diabetes complications were collected from patient records. In the few cases where information was not available in the patient records (hemoglobin A_{1c} value for one patient and diabetes duration for another), data were complemented by asking the patients.

Direct valuation

In the TTO method [4], the preferences of each respondent for a specific health state are elicited by asking the respondent to choose between two different health states, each assigned a specific number of years followed by death. In the present study, respondents were asked to choose between living in their current health state for their remaining life expectancy (*t*), followed by death, and a state of full health for a shorter period (*x*), also followed by death. The time *x* was varied until the number of years was found for which the respondent was indifferent between the two alternatives. The TTO value was calculated by dividing *x* by the time of the first alternative *t*. The remaining life expectancy used in the questions was based on the Swedish population's average remaining survival (rounded to the nearest decade) for the age and sex group of the respondent [29]. For the EQ-VAS rating

Table 1 – Comparison of the measures.

Measure	Possible range	Dimensions	Response levels	Valuation technique	Valuation sample
HUI-3	−0.36 to 1.00	Vision, hearing, speech, ambulation, dexterity, emotion, cognition, pain	5–6	SG and VAS	504 representatives of the general population in Canada
EQ-5D	−0.59 to 1.00	Mobility, self-care, usual activities, pain/discomfort, anxiety/depression	3	TTO	3 395 representatives from the general population in UK
EQ-VAS	0–100	N/A	N/A	VAS	Study participants
TTO	0–1.00	N/A	N/A	TTO	Study participants
NEI-VFQ-25	0–100	General health, general vision, ocular pain, near activities, distance activities, social functioning, mental health, role difficulties, dependency, driving, color vision and peripheral vision	6	Average of the scores in the dimensions	N/A

EQ-5D, EuroQol five-dimensional questionnaire; EQ-VAS, EuroQol Visual Analogue Scale; HUI-3, health utilities index mark 3; N/A, not applicable; NEI-VFQ-25, 25-item National Eye Institute Visual Functioning Questionnaire; SG, standard gamble; TTO, time trade-off; VAS, visual analogue scale.

scale [7], which is a separate part of the EQ-5D, respondents are asked to estimate their own health on a scale from 0 to 100, with 0 representing the worst imaginable health state and 100 the best.

Multiattribute instruments

With both HUI-3 [6] and EQ-5D [7], information on the respondent's health status is collected by asking the respondent to complete a specific health questionnaire or by conducting an interview following a detailed interview manual. The responses in the questionnaires can then be scored by using published tariffs that have been derived by using one of the direct valuation methods. In this study, we have for both these instruments used specific versions that have been developed for telephone interviews [30,31]. The HUI-3 interview manual consists of 41 questions, and the included questions concern how respondents perceive their own HRQOL in the dimensions of vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain (Table 1). The health description resulting from the questionnaire was scored by using a multiattribute scoring function with values elicited from a random sample of the Canadian general population (n = 504) by using VAS in combination with the SG method. The EQ-5D interview manual consists of five questions, each representing one HRQOL dimension: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. In each dimension, respondents can classify themselves into one of three levels of severity: no problems, some problems, and extreme problems. The scoring function used in the present study was derived from a random sample of the general population in the United Kingdom (n = 3395) [32] by using the TTO method.

NEI-VFQ-25

NEI-VFQ-25 [33] assesses how eye disease affects a patient's ability to live without pain, to work productively, and to interact with family and friends. The results can be presented as an overall score or as a health profile using 12 dimensions (Table 1). The NEI-VFQ-25 composite score is presented as an index between 0 and 100, with 0 representing the worse possible score and 100 the best. It assumes equal intervals between the response choices within each dimension, and the composite score of the measure is an unweighted average of all attributes except for general health.

Statistical analysis

All patients who answered all questions in the four generic measures were included in the analysis. The bakground characteristics are presented as proportions of patients or mean values with SD. For each generic measure, adjusted mean scores (QALY weights) and 95% confidence intervals were calculated for the severity levels by using analysis of covariance models. The models' *F* statistics were used to obtain a measure of the methods' ability to detect differences between the severity levels. The variables included in the adjustment were age, sex, diabetes duration, hemoglobin A_{1c} level (Mono S), cardiovascular disease (known peripheral vascular disease, transient ischemic attack/stroke, angina pectoris/myocardial infarction, or previous vascular intervention), neuropathy (clinical signs of any form of neuropathy mentioned in the patient's clinical record), and nephropathy (proteinuria and/or an increase in serum creatinine level). The adjusted scores were calculated on the basis of the distribution of these variables in the no-DR group. Sex and age were included in the model because these variables previously have been seen to have an effect on HRQOL [34]. In addition, all variables presented in Table 2 except for visual impairment and the NEI VFQ-25 scores were considered for inclusion in the model because they a priori were believed to affect the HRQOL of the patients. However, they were included in the model only if they differed between the severity levels of DR. The residuals were tested for non-normality and heteroscedasticity by plotting the residuals toward the fitted values and using the Breusch-Pagan test. Because the residuals were found to be heteroscedastic, this was corrected for by using White's robust standard errors [35]. This approach has been recommended when the purpose is to inform economic evaluations [36]. The model was also tested for multicolinearity, but none of the variables had a variance inflation factor (VIF) of more than 2. When relevant, pairwise comparisons were made via analysis of variance corrected with the Tukey-Kramer procedure [37]. The statistical software used for the basic analyses was STATA/IC 10.1 (Stata Corp LP, Texas, USA). The adjustment with analysis of covariance was performed in SAS 9.1.3 (SAS Institute Inc, North Carolina, USA).

Comparison of valuation methods

In evaluating the performance of QALY measures, comparisons can be made in terms of practicality, reliability, and validity [38]. The last of these, validity, concerns to what extent the measures

Table 2 – Clinical and sociodemographic characteristics by severity level of DR in the worse eye.

	No retinopathy (n = 35)		BR (n = 37)		PDR (n = 26)		DME (n = 31)		Legal blindness (n = 23)		P
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
Age (y)	51.06	17.31	57.62	13.40	51.69	13.05	58.19	12.09	67.00	11.07	<0.001
Hb A_{1c} (Mono S)	6.48	1.05	7.08	0.97	7.60	1.53	7.35	1.26	6.95	1.08	0.004
Diabetes duration	11.03	8.08	21.54	13.12	27.50	11.29	19.58	9.55	33.00	19.92	<0.001
NEI-VFQ-25 score	96.37	4.94	95.29	7.18	88.68	13.22	90.86	10.67	49.63	21.68	<0.001
	n	%	n	%	n	%	n	%	n	%	P
Sex (women)	12	34.3	14	37.8	12	46.2	7	22.6	7	30.4	0.432
Type 1 diabetes	17	48.6	18	48.7	17	65.4	15	48.4	12	52.2	0.670
Neuropathy	0	0.0	5	13.5	6	23.1	4	12.9	12	52.2	<0.001
Nephropathy	0	0.0	5	13.5	7	26.9	4	12.9	11	47.8	<0.001
Cardiovascular disease	4	11.4	7	18.9	3	11.5	5	16.1	11	47.8	0.015
Visual impairment (best eye)											
Normal vision	34	97.1	36	97.3	21	80.8	21	67.7	2	8.7	<0.001
Mild visual impairment	1	2.9	1	2.7	5	19.2	9	29.0	1	4.4	
Moderate visual impairment	0	0.0	0	0.0	0	0.0	1	3.2	5	21.7	
Severe visual impairment	0	0.0	0	0.0	0	0.0	0	0.0	15	65.2	
Visual impairment (worse eye)											
Normal vision	34	97.1	35	94.6	13	50.0	17	54.8	0	0.0	<0.001
Mild visual impairment	1	2.9	2	5.4	10	38.5	9	29.0	0	0.0	
Moderate visual impairment	0	0.0	0	0.0	3	11.5	5	16.1	0	0.0	
Severe visual impairment	0	0.0	0	0.0	0	0.0	0	0.0	23	100.0	
Smoker	4	11.4	4	10.8	1	3.9	3	9.7	2	8.7	0.903
Former smoker (of nonsmokers)	15	48.4	18	54.6	11	44.0	14	50.0	13	61.9	0.791
Education*											
Nine-year compulsory school and/or upper secondary school	21	60.0	25	67.6	18	69.2	23	76.7	17	73.9	0.680
Vocational training or university	14	40.0	12	32.4	8	30.8	7	23.3	6	26.1	
Civil status											
Married or cohabiting	25	71.4	31	83.8	19	73.1	22	71.0	18	78.3	0.680
Single, divorced, or widower	10	28.6	6	16.2	7	26.9	9	29.0	5	21.7	

BR, background retinopathy; DME, diabetic macular edema; DR, diabetic retinopathy; Hb A_{1c}, hemoglobin A_{1c}; NEI-VFQ-25, 25-item National Eye Institute Visual Functioning Questionnaire; PDR, proliferative diabetic retinopathy; VA, visual acuity.

* Data missing on one patient. Normal vision was defined as VA 20/10-20/25, mild visual impairment as VA 20/32-20/63, moderate visual impairment as VA 20/80-20/160, and severe visual impairment as VA of 20/200 or less. The means were compared by one-way analysis of variance using the *F* test, and the proportions were compared by using Fisher's exact test.

measure what they are intended to do. Ideally, this would be tested by comparing the results of the measure to a gold standard. However, because no such gold standard exists for QALY weights, other more indirect methods have been recommended. Empirical validity measures to what degree the measure is able to identify or measure what it is supposed to in practice. One way of doing this is to study whether the measure generates QALY weights that are in line with expected differences between groups of patients, so-called hypothetical preferences. This method has previously been used to investigate the validity of QALY measures in various disease areas [15,39–43].

The empirical validity of the measures in this study was examined by testing a hypothetical preference rule that the scores should differ significantly between groups with different severity levels of DR and that the scores should decrease with disease severity. The ability of the measures to detect differences in HRQOL between patient groups was compared by using relative efficiency statistics [44]. The relative efficiency of the measures was calculated by taking the square of the *F* statistic of the comparator measure divided by the square of the *F* statistic of the reference measure (the measure with the lowest *F* statistic). The adjusted scores from the generic measures were also tested toward the patients' VA and their results from the NEI-VFQ-25 by calculating the Spearman's rank correlation coefficient. Our assumption was that health states with higher VA or a higher score on the NEI-VFQ-25 scale would be preferable to health states with lower scores. Spearman's rank correlation coefficient was chosen as the main measure of correlation, because we were primarily interested in the correlation between the measures' ranking of individuals and not the correlation between the absolute values. A correlation was considered moderate if it exceeded 0.30 and strong if it exceeded 0.5 [45]. To investigate whether any of the measures had a higher correlation with any of the external indicators (NEI-VFQ-25 and VA) than did the other measures, the correlations were tested against each other by using Wolfe's test for comparing dependent correlation coefficients [46].

Results

Of the 251 patients who were asked to participate, 152 were included in the analysis (their clinical and sociodemographic characteristics are described in Table 2). Of those who were excluded, 24 chose not to participate and 61 were excluded either on exclusion criteria (after assessment of full patient records) or because they were not reachable at the time of the telephone interview. In

Table 3 – QALY weights (95% CI) for different severity levels of DR and the relative efficiency (RE) of the measures in detecting differences in HRQoL between the severity levels.

Valuation method	Classification	n	Measure outcome	*F*	*P*	RE
	Severity level in the worse eye		**Unadjusted mean value (95% CI)**			
HUI-3	No DR	35	0.88 (0.83–0.93)			
	BR	37	0.78 (0.69–0.86)			
	PDR	26	0.81 (0.74–0.88)	20.71	<0.001	8.59
	DME	31	0.79 (0.73–0.86)			
	Blindness	23	0.39 (0.26–0.51)			
EQ-5D	No DR	35	0.81 (0.75–0.86)			
	BR	37	0.72 (0.61–0.83)			
	PDR	26	0.77 (0.68–0.87)	3.26	0.01	1.35
	DME	31	0.81 (0.74–0.88)			
	Blindness	23	0.59 (0.45–0.74)			
TTO	No DR	35	0.85 (0.79–0.92)			
	BR	37	0.78 (0.71–0.86)			
	PDR	26	0.78 (0.70–0.86)	2.41	0.05	1.00
	DME	31	0.82 (0.74–0.89)			
	Blindness	23	0.68 (0.54–0.81)			
EQ-VAS	No DR	35	80.51 (77.33–83.70)			
	BR	37	72.16 (66.14–78.18)			
	PDR	26	75.46 (68.59–82.33)	5.74	<0.001	2.38
	DME	31	74.48 (68.25–80.72)			
	Blindness	23	59.96 (51.55–68.36)			
	Severity level in the worse eye		**Adjusted mean*value (95% CI)**			
HUI-3	No DR	35	0.88 (0.81–0.95)			
	BR	37	0.85 (0.77–0.93)			
	PDR	26	0.89 (0.79–0.99)	9.22	<0.001	98.29
	DME	31	0.86 (0.78–0.95)			
	Blindness	23	0.53 (0.41–0.66)			
EQ-5D	No DR	35	0.81 (0.73–0.89)			
	BR	37	0.77 (0.68–0.86)			
	PDR	26	0.83 (0.71–0.94)	1.87	0.12	4.04
	DME	31	0.86 (0.76–0.96)			
	Blindness	23	0.67 (0.52–0.81)			
TTO	No DR	35	0.85 (0.78–0.93)			
	BR	37	0.81 (0.73–0.89)			
	PDR	26	0.83 (0.73–0.94)	0.93	0.45	1.00
	DME	31	0.83 (0.74–0.92)			
	Blindness	23	0.72 (0.59–0.85)			
EQ-VAS	No DR	35	80.51 (75.36–85.67)			
	BR	37	73.84 (68.04–79.63)			
	PDR	26	77.74 (70.36–85.12)	3.12	0.02	11.25
	DME	31	77.03 (70.71–83.36)			
	Blindness	23	63.26 (53.93–72.59)			
	Severity level in the better eye		**Adjusted mean*value (95% CI)**			
HUI-3	No DR	39	0.88 (0.82–0.95)			
	BR	48	0.86 (0.79–0.93)			
	PDR	32	0.87 (0.78–0.97)	9.81	<0.001	75.37
	DME	18	0.83 (0.72–0.93)			
	Blindness	15	0.48 (0.34–0.61)			
EQ-5D	No DR	39	0.82 (0.74–0.89)			
	BR	48	0.78 (0.69–0.86)			
	PDR	32	0.84 (0.73–0.95)	1.13	0.34	1.00
	DME	18	0.86 (0.73–1.00)			
	Blindness	15	0.69 (0.52–0.85)			
TTO	No DR	39	0.87 (0.80–0.93)			
	BR	48	0.78 (0.71–0.86)			
	PDR	32	0.81 (0.72–0.91)	3.51	0.01	9.65
	DME	18	0.89 (0.78–1.00)			
	Blindness	15	0.62 (0.48–0.77)			

(*continued on next page*)

Table 3 (*continued*)

Valuation method	Classification	n	Measure outcome	*F*	*P*	RE
EQVAS	No DR	39	80.59 (75.74–85.44)			
	BR	48	74.67 (69.40–79.94)			
	PDR	32	76.70 (69.54–83.86)	3.66	0.01	10.49
	DME	18	76.94 (68.75–85.13)			
	Blindness	15	60.08 (49.79–70.37)			

* Adjusted for age, sex, diabetes duration, HbA1c (Mono S), cardiovascular disease, neuropathy and nephropathy.
BR, background retinopathy; CI, confidence interval; DME, diabetic macular edema; DR, diabetic retinopathy; EQ-5D, EuroQol five-dimensional questionnaire; EQ-VAS, EuroQol Visual Analogue Scale; HUI-3, health utilities index mark 3; PDR, proliferative diabetic retinopathy; RE, relative efficiency; TTO, time trade-off.

addition, 14 individuals were excluded from analysis because of missing data.

The percentages of patients reporting a score equivalent to full health, that is, a score of 1.0, was 8%, 24%, 26%, and 0% for HUI-3, TTO, EQ-5D, and EQ-VAS, respectively. HUI-3 resulted in scores below zero for two patients (1%); for EQ-5D, the corresponding number was six patients (4%). As expected, the mean NEI-VFQ-25 composite scores decreased with progression of the disease (Table 2). All the measures were able to identify significant differences between the different groups of visual impairment in the better eye (**Table 4**), but neither TTO nor EQ-5D identified any significant differences between the severity levels classified according to the pathological progression or visual impairment in the worse eye (Tables 3 and 4). In addition, EQ-5D did not identify significant differences between the severity levels classified according to the better eye. For all measures, the significant differences were mainly due to blindness (pair-wise comparisons corrected with Tukey-Kramer) and for the less severe states the scores did not always follow the hypothesized ordering (Tables 3 and 4). DME and PDR had in some cases higher scores than BR and in a few cases even higher than no DR. This was particularly evident in the results from EQ-5D and TTO, independent of classification according to the better eye or the worse eye. Even when patients were classified according to visual impairment in the better eye, TTO and EQ-5D resulted in a score ordering that contradicted our hypothesized expectations.

When the patients were classified according to the worse eye, HUI-3 resulted in scores higher than those from other measures for all severity levels except blindness, for which HUI-3 in contrast recorded the lowest score. Of the four measures, EQ-VAS gave the lowest score for the levels before blindness (the less severe health states) and TTO had the highest score for blindness (the most severe health state). **Figure 1** and **Table 3** show that the difference in HRQOL between PDR or less severe DR and blindness is larger if measured with HUI-3 (a difference of approximately 0.36) than with any of the other measures. The difference is smallest with the EQ-5D score (approximately 0.16). Compared with the other measures, HUI-3 had the highest relative efficiency (i.e., was the most efficient), independent of whether the patients were classified on the basis of pathological progression of DR or their visual impairment (Tables 3 and 4). HUI-3 also had a significantly higher correlation with VA (in the better and worse eye) and with NEI-VFQ-25 (Table 5) than did the other measures. The Spearman's rank correlation between HUI-3 and NEI-VFQ-25 was 0.54 ($P < 0.001$), meaning that the ranking of the patients according to HUI-3 correlated strongly with the ranking according to vision-specific measures. The other measures' ranking showed only low to moderate correlation with the ranking according to VA and NEI-VFQ-25.

Discussion

Main findings

In this study, we estimated QALY weights for DR with four HRQOL measures and investigated their performance in patients with DR by examining whether they can identify hypothesized differences between severity levels and how well their results correlate with external indicators of the severity of DR. The results showed that all four measures identified significant differences between the groups of visual impairment in the better eye but only HUI-3 and EQ-VAS detected significant differences between the patient groups classified according to the visual impairment or pathological progression in the worse eye. Contrary to our expectations, the scores did not always follow the progression of the disease. Of the four measures, HUI-3 was the most sensitive at detecting differences between severity levels, with an HRQOL decrement for blindness that was at least two times greater than that of the other measures. The HUI-3 ranking was also found to have a stronger correlation with the ranking according to visual function. This large difference between HUI-3 and the other measures implies that their estimates could, if used in cost-effectiveness analyses, lead to different conclusions concerning the cost-effectiveness of an intervention.

Our results correspond well with those of Espallargues et al. [14], who showed that HUI-3 is more sensitive than EQ-5D, TTO, and VAS in detecting differences in HRQOL due to macular degeneration (which is a condition with a effect similar to that of DR on VA). Furthermore, in a comparison of EQ-5D, SF-6D, and the 15-D measure, Kontodimopoulos et al. [23] showed that the 15-D measure, which, like HUI-3, has a specific dimension for vision, is more sensitive to DR than are the other measures. In contrast, Lloyd et al. [22] found that HUI-3 and EQ-5D yield similar decrements in HRQOL as DR progresses. It is however unclear whether the values presented in the latter were adjusted for comorbidities. Regarding the results from our TTO question, the decrement from normal vision to blindness was smaller than in previous TTO studies [16]. However, the vision-specific TTO versions that have been used in most previous studies have been argued to overestimate the effects of visual impairment [17,18].

Inconsistent ordering of the scores has previously been observed by Lloyd et al. [22], who suggested that this finding may be related to the loss of independence in the early stages of visual impairment. In our study, however, the inconsistency was mostly related to relatively low scores for patients with BR, which is an asymptomatic stage of the disease and is not expected to be associated with loss of independence (when the patients were classified according to VA, this was reflected in relatively low scores in the group with normal vision). Patients with BR had a higher proportion of cardiovascular disease, but this was adjusted for in the analysis of covariance. It is possible, however, that we failed to adjust for some other important di-

Table 4 – QALY weights (95% CI) for different levels of vision impairment and the relative efficiency (RE) of the measures in detecting differences in HRQoL between impairment levels.

Valuation method	Classification	n	Measure outcome	*F*	*P*	RE
	Visual impairment in the worse eye		**Unadjusted mean value (95% CI)**			
HUI-3	Normal vision	99	0.82 (0.78–0.86)	25.94	<0.001	9.98
	Mild visual impairment	22	0.81 (0.76–0.87)			
	Moderate visual impairment	8	0.77 (0.59–0.96)			
	Severe visual impairment	23	0.39 (0.26–0.51)			
EQ-5D	Normal vision	99	0.77 (0.72–0.82)	3.92	0.01	1.51
	Mild visual impairment	22	0.83 (0.77–0.89)			
	Moderate visual impairment	8	0.77 (0.71–0.83)			
	Severe visual impairment	23	0.59 (0.45–0.74)			
TTO	Normal vision	99	0.82 (0.78–0.86)	2.60	0.05	1.00
	Mild visual impairment	22	0.78 (0.69–0.88)			
	Moderate visual impairment	8	0.79 (0.58–1.00)			
	Severe visual impairment	23	0.68 (0.54–0.81)			
EQ-VAS	Normal vision	99	75.95 (72.73–79.17)	6.79	<0.001	2.61
	Mild visual impairment	22	77.82 (72.48–83.15)			
	Moderate visual impairment	8	67.50 (50.00–85.00)			
	Severe visual impairment	23	59.96 (51.55–68.36)			
	Visual impairment in the worse eye		**Adjusted mean* value (95% CI)**			
HUI-3	Normal vision	99	0.82 (0.78–0.86)	13.16	<0.001	12.78
	Mild visual impairment	22	0.90 (0.81–1.00)			
	Moderate visual impairment	8	0.81 (0.66–0.95)			
	Severe visual impairment	23	0.51 (0.40–0.62)			
EQ-5D	Normal vision	99	0.77 (0.72–0.82)	2.52	0.06	2.45
	Mild visual impairment	22	0.87 (0.76–0.98)			
	Moderate visual impairment	8	0.78 (0.61–0.96)			
	Severe visual impairment	23	0.66 (0.53–0.79)			
TTO	Normal vision	99	0.82 (0.78–0.86)	1.03	0.38	1.00
	Mild visual impairment	22	0.81 (0.71–0.91)			
	Moderate visual impairment	8	0.81 (0.65–0.96)			
	Severe visual impairment	23	0.71 (0.59–0.83)			
EQ-VAS	Normal vision	99	75.95 (72.90–79.00)	4.17	0.01	4.05
	Mild visual impairment	22	80.43 (73.26–87.59)			
	Moderate visual impairment	8	69.20 (58.14–80.25)			
	Severe visual impairment	23	63.89 (55.66–72.13)			
	Visual impairment in the better eye		**Adjusted mean* value (95% CI)**			
HUI-3	Normal vision	114	0.83 (0.79–0.86)	17.49	<0.001	5.59
	Mild visual impairment	17	0.81 (0.71–0.91)			
	Moderate visual impairment	6	0.53 (0.37–0.70)			
	Severe visual impairment	15	0.40 (0.28–0.52)			
EQ-5D	Normal vision	114	0.78 (0.73–0.82)	3.28	0.03	1.05
	Mild visual impairment	17	0.84 (0.72–0.97)			
	Moderate visual impairment	6	0.54 (0.34–0.75)			
	Severe visual impairment	15	0.63 (0.48–0.77)			
TTO	Normal vision	114	0.82 (0.78–0.85)	3.13	0.03	1.00
	Mild visual impairment	17	0.79 (0.68–0.90)			
	Moderate visual impairment	6	0.85 (0.67–1.03)			
	Severe visual impairment	15	0.62 (0.49–0.75)			
EQ-VAS	Normal vision	114	76.18 (73.37–78.98)	5.62	0.001	1.80
	Mild visual impairment	17	75.59 (67.81–83.36)			
	Moderate visual impairment	6	61.70 (48.71–74.69)			
	Severe visual impairment	15	58.10 (49.06–67.14)			

* Adjusted for age, sex, diabetes duration, HbA1c (Mono S), cardiovascular disease, neuropathy and nephropathy. CI, confidence interval; EQ-5D, EuroQol five-dimensional questionnaire; EQ-VAS, EuroQol Visual Analogue Scale; HUI-3, health utilities index mark 3; RE, relative efficiency; TTO, time trade-off.

abetes-related variable. In a more detailed analysis of the dimensions included in HUI-3 and EQ-5D, the group of patients with BR was found to have a higher proportion of problems with dexterity. Given this unexpected finding and the fact that the differences in scores between the early stages are not significant, we would advise users of our results to set the QALY weight for BR at equal to no DR until we know more about the reason behind the decrement in this group.

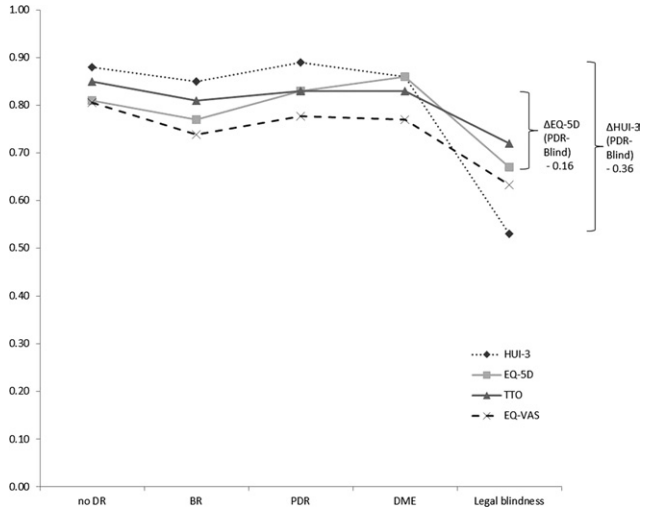

Fig. 1 – Adjusted QALY weights for different severity levels of DR based on the worse eye. The weights are adjusted for age, sex, diabetes duration, hemoglobin A_{1c} level, cardiovascular disease, neuropathy, and nephropathy. The differences are calculated by subtracting the score for blindness from that of PDR. BR, background retinopathy; DME, diabetic macular edema; DR, diabetic retinopathy; EQ-5D, EuroQol five-dimensional questionnaire; EQ-VAS, EuroQol Visual Analogue Scale; HUI-3, health utilities index mark 3; PDR, proliferative diabetic retinopathy; TTO, time trade-off.

Concerning the differences between the direct and indirect measures, patients have previously been found to report higher scores than the general public [47–49], especially for severe health states. A possible reason is that patients adapt to their own health state. This corresponds well with our results for HUI-3, where HUI-3 gave lower scores for blindness than did the direct measures. Adaption could perhaps also explain why patients with moderate visual impairment reported higher TTO scores than did those with no or mild visual impairment. In addition, the questionnaires of the indirect methods, such as EQ-5D and HUI-3, limit patients' descriptions of their health states to specific dimensions. It has been argued that potentially positive aspects that would be captured with the direct methods might be missed by the indirect estimations [50]. On the other hand, it could be argued that patients do not take their vision into account when responding to the TTO questions. When patients had finished the TTO exercise, however, they were asked whether they had considered full health as also including full vision, and 60% stated that they had. Of those who had not, 80% (vs. 73% among those who had done so) had normal vision. Further research on this subject is required to learn more about the role of vision in TTO exercises.

Of the indirect measures, HUI-3 has an advantage over EQ-5D when used for patients with diseases leading to visual impairment, because it includes a specific dimension for vision and focuses more on functional impairment. For example, while the dimensions Mobility in EQ-5D and Ambulation in HUI-3 may seem very similar, the framing of the questions differ. In EQ-5D, the respondents are asked about their ability to walk with or without difficulties, but in HUI-3, the focus is on whether they can walk around the neighborhood without help from anyone and/or without walking equipment. Because a patient with visual impairment may not have difficulties with moving in general but may need some sort of guidance or support when walking from one place to another, these questions may give completely different results. Furthermore, the measures differ because their tariffs have been elicited by using different valuation methods and modeling techniques. While the tariff for EQ-5D is based on TTO values, the HUI-3 tariff is derived from a combination of SG and VAS. TTO and SG have previously been found to give different results [51].

When comparing the two direct methods, EQ-VAS gave lower scores than did TTO, in line with previous literature [51]. A possible explanation could be that no patient reported full health with EQ-VAS but it could also be due to the different methodologies. Because EQ-VAS is not a choice-based method, many do not consider it an appropriate method for estimating QALY weights.

Limitations

This article has a number of limitations that need to be discussed. First, the questionnaires were administered in the same order to all patients. This could potentially have biased our comparison of measures because of carryover effects or fatigue. Administration order, however, has been found in various previous studies to have no or only a marginal effect on HRQOL measurements [52–54]. Furthermore, we believe that the administration of the study by telephone required patients to answer actively, reducing the risk of fatigue, and this was supported by the patients' tone of voice. Nevertheless, caution is warranted when comparing the results of the measures, especially because HUI-3, which was the first measure in our interviews, was found to be the most sensitive.

Second, we did not use any visual support for the TTO and EQ-VAS exercises. Because of the severe visual impairment of some of the patients in the study, using any form of visual aid could have resulted in a severe bias, because those patients would not have been able to use any such aid. In addition, TTO and EQ-VAS exercises administered by telephone have previously been found to give similar results as do face-to-face interviews [55–57]. Our TTO question could also be considered to have an advantage over those of other studies, because the years patients were willing to trade were elicited by using "ping-pong" questions rather than an open-ended question.

Third, our sample size was relatively small. This could explain why there were no large differences between the stages prior to

Table 5 – Spearman's rank correlation between NEI-VFQ-25, Snellen visual acuity, and the generic HRQOL measures.

	r			
	HUI-3	EQ-5D score	TTO	EQ-VAS
NEI-VFQ-25	0.54	0.27	0.33	0.32
Visual acuity (worse eye)	0.59	0.31	0.35	0.34
Visual acuity (better eye)	0.50	0.24	0.28	0.27
HUI-3		0.70	0.58	0.73
EQ-5D score			0.52	0.97
TTO				0.46

All correlations were significantly higher than 0 ($P < 0.01$). The correlations between HUI-3 and the external indicators of severity of DR (NEI-VFQ-25 and visual acuity) were significantly higher than the correlations between the other measures and the external indicators (tested with Wolfe's test for comparing dependent correlation coefficients). The scores from the generic measures were adjusted for age, sex, diabetes duration, Hb A_{1c} level, cardiovascular disease, neuropathy, and nephropathy.
EQ-5D, EuroQol five-dimensional questionnaire; EQ-VAS, EuroQol Visual Analogue Scale; Hb A_{1c}, hemoglobin A_{1c}; HRQOL, health-related quality of life; HUI-3, health utilities index mark 3; NEI-VFQ-25, 25-item National Eye Institute Visual Functioning Questionnaire; TTO, time trade-off.

blindness. However, because the more severe levels for some measures had even higher scores than did the less severe levels, it is not likely that these would have been associated with significant decrements, even if we have had a larger sample. In addition, the lack of significant decrements could also be explained by the fact that patients with more severe levels of DR had better vision than expected.

Fourth, our classification of the patients into severity levels according to the pathological progression of DR could be questioned. A combination of VA and contrast sensitivity has been recommended for classifying patients into different severity levels relevant for their HRQOL [58]. In this study, however, we were looking for QALY weights that could be used to populate decision models, and many of the models [24–27] that have been developed in the DR field have been based on the pathological progression in the worse eye. Even though HRQOL may mostly be affected by the patients' ability to see, worries about the disease may also affect their HRQOL. In addition, treatment effect and costs are related to the pathological state of the eye. However, because there were such small differences in values between the stages prior to blindness, we also decided to present values for different degrees of VA.

Implications

The fact that there are differences between the sensitivity of the measures and that they give different decrements for blindness may have important policy implications. In cost-utility analyses of interventions directed at DR, the choice of measure could have a significant impact on the incremental cost per QALY. Imagine that a new treatment for DR, compared with the standard treatment, prevents one additional patient from progressing from PDR to blindness (on the worse eye) at the age of 60 years and that the patient, independent of his or her eye condition, would live until the age of 80 years. Using the estimates from HUI-3 (Fig. 1), this would result in an incremental effect of 7.2 QALYs (0.36×20), while the estimates from EQ-5D would give an incremental effect of 3.2 QALYs (0.16×20). If we also assumed that the new treatment costs GBP 150,000 (1.601 GBP/USD July 7, 2011) more per case of blindness prevented, this would result in a cost per QALY of approximately GBP 21,000 and GBP 47,000 for HUI-3 and EQ-5D, respectively. As many decision-making bodies, such as the National Institute of Clinical Excellence in the United Kingdom, use a threshold of GBP 20,000 to 30,000 [59] to determine whether a treatment is cost-effective, the choice of measure could have an important influence on whether a treatment is implemented.

Where it is likely that the results of different measures diverge, there are various aspects that should be considered when choosing between measures. On a policy level, it should be determined whether the decisions concerning the allocation of resources are to be based on patient values or on values of the general public. In our study, the choice would be between direct patient valuation with TTO or EQ-VAS and indirect valuation based on public values with HUI-3 or EQ-5D. If the chosen policy is compromised because of the lack of relevant measures, decision makers should be aware that this may give other results. On a more practical level, assuming that the measures are practical and reliable, it is important to consider whether they are able to capture the effect of the specific disease on HRQOL [39]. This can be achieved by studying the construct of the measures in terms of face and content validity and by investigating whether the instruments capture hypothesized differences (empirical validity). Our findings concerning the sensitivity of HUI-3 and its correlation with visual function suggest that HUI-3 captures DR-related differences in HRQOL that are not detected by EQ-5D. This could be explained by differences in the measures' construct. Thus, if choosing between measures based on public values, we would recommend HUI-3 for cost-effectiveness calculations related to interventions directed at DR. Regarding the direct measures, none of the measures correlated strongly with visual function but EQ-VAS showed values that were more in line with the hypothesized ordering than TTO. We would not recommend EQ-VAS, however, for use in cost-effectiveness analyses because it is not a choice-based method. Other studies have solved the insensitivity of TTO by using vision-specific TTO questions. The use of these vision-specific questions in cost-utility analysis is questionable, however [17,18]. Therefore, future research is necessary on the role of vision in direct valuation exercises.

Conclusions

In conclusion, our study has shown that the choice of measure for estimation of QALY weights for cost-utility analyses matters, because it may have an impact on whether an intervention is considered cost-effective. This is of interest not only to those conducting future cost-utility analyses of interventions directed at DR but also to decision makers evaluating cost-effectiveness evidence for interventions directed at patients with comprised vision. Furthermore, the study suggests that HUI-3 is the most sensitive measure for detecting differences in HRQOL due to DR and that it can, if decisions are to be made on the basis of values from the general public, be recommended for use in cost-utility analyses of interventions directed at DR.

Acknowledgments

The authors thank the staff of the eye clinic at Linköping University Hospital and the patient counselors at the low vision rehabilitation centers at Linköping University Hospital and Vrinnevi Hospital for their help in the recruitment of patients. We are also grateful to David Andersson, statistician at the Center for Medical Technology Assessment, for his useful comments on the statistical analysis, and to Colin Green at the Peninsula Medical School, Exeter, for his helpful comments on the manuscript.

Source of financial support: The study was sponsored by AstraZeneca AB. However, the research group had full control over the results and the decision to submit them for publication.

REFERENCES

[1] National Institute for Health and Clinical Excellence. Guide to the Methods of Technology Appraisal. London: National Institute for Health and Clinical Excellence, 2008.
[2] International Society for Pharmacoeconomics and Outcomes Research. Pharmacoeconomic guidelines around the world. 2008. Available from: http://www.ispor.org/peguidelines/index.asp. [Accessed September 28, 2011].
[3] Von Neumann J, Morgenstern O. Theory of Games and Economic Behaviour. Princeton University Press. Princeton, New Jersey. 1944.
[4] Torrance GW, Thomas WH, Sackett DL. A utility maximization model for evaluation of health care programs. Health Serv Res 1972;7:118–33.
[5] Patrick DL, Bush JW, Chen MM. Methods for measuring levels of well-being for a health status index. Health Serv Res 1973;8:228–45.
[6] Feeny D, Furlong W, Torrance GW, et al. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Med Care 2002;40:113–28.
[7] The EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199–208.
[8] Sintonen H. The 15D instrument of health-related quality of life: properties and applications. Ann Med 2001;33:328–36.
[9] Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ 2002;21:271–92.
[10] The Pharmaceutical Benefits Board. General Guidelines for Economic Evaluations from the Pharmaceutical Benefits Board (LFNAR 2003:2). The Pharmaceutical Benefits Board, Stockholm, 2003.
[11] McDonough CM, Tosteson AN. Measuring preferences for cost-utility analysis: how choice of method may influence decision-making. Pharmacoeconomics 2007;25:93–106.

[12] Seymour J, McNamee P, Scott A, et al. Shedding new light onto the ceiling and floor? A quantile regression approach to compare EQ-5D and SF-6D responses. Health Econ 2010;19:683–96.
[13] Kopec JA, Willison KD. A comparative review of four preference-weighted measures of health-related quality of life. J Clin Epidemiol 2003;56:317–25.
[14] Espallargues M, Czoski-Murray CJ, Bansback NJ, et al. The impact of age-related macular degeneration on health status utility values. Invest Ophthalmol Vis Sci 2005;46:4016–23.
[15] Barton GR, Sach TH, Doherty M, et al. An assessment of the discriminative ability of the EQ-5D index, SF-6D, and EQ VAS, using sociodemographic factors and clinical conditions. Eur J Health Econ 2008;9:237–49.
[16] Fenwick EK, Pesudovs K, Rees G, et al. The impact of diabetic retinopathy: understanding the patient's perspective. Br J Ophthalmol 2010;95:774–82.
[17] Lee BS, Kymes SM, Nease RF Jr, et al. The impact of anchor point on utilities for 5 common ophthalmic diseases. Ophthalmology 2008;115: 898–903, e4.
[18] Kymes SM, Lee BS. Preference-based quality of life measures in people with visual impairment. Optom Vis Sci 2007;84:809–16.
[19] Sjolie AK, Klein R, Porta M, et al. Effect of candesartan on progression and regression of retinopathy in type 2 diabetes (DIRECT-Protect 2): a randomised placebo-controlled trial. Lancet 2008;372:1385–93.
[20] Chaturvedi N, Porta M, Klein R, et al. Effect of candesartan on prevention (DIRECT-Prevent 1) and progression (DIRECT-Protect 1) of retinopathy in type 1 diabetes: randomised, placebo-controlled trials. Lancet 2008;372:1394–402.
[21] Mauer M, Zinman B, Gardiner R, et al. Renal and retinal effects of enalapril and losartan in type 1 diabetes. N Engl J Med 2009;361:40–51.
[22] Lloyd A, Nafees B, Gavriel S, et al. Health utility values associated with diabetic retinopathy. Diabet Med 2008;25:618–24.
[23] Kontodimopoulos N, Pappa E, Chadjiapostolou Z, et al. Comparing the sensitivity of EQ-5D, SF-6D and 15D utilities to the specific effect of diabetic complications. Eur J Health Econ 2012; 13:111–20. Epub 2010 Dec 5.
[24] Aoki N, Dunn K, Fukui T, et al. Cost-effectiveness analysis of telemedicine to evaluate diabetic retinopathy in a prison population. Diabetes Care 2004;27:1095–101.
[25] Vijan S, Hofer TP, Hayward RA. Cost-utility analysis of screening intervals for diabetic retinopathy in patients with type 2 diabetes mellitus. JAMA 2000;283:889–96.
[26] Zhou H, Isaman DJ, Messinger S, et al. A computer simulation model of diabetes progression, quality of life, and cost. Diabetes Care 2005;28: 2856–63.
[27] Brandle M, Herman WH. The CORE diabetes model. Curr Med Res Opin 2004(20, Suppl. 1):S1–3.
[28] Tung TH, Chen SJ, Lee FL, et al. A community-based study for the utility values associated with diabetic retinopathy among type 2 diabetics in Kinmen, Taiwan. Diabetes Res Clin Pract 2005;68:265–73.
[29] Statistics Sweden. Återstående livslängd efter kön och ålder. Available from: http://www.ssd.scb.se/databaser/makro/MainTable.asp?yp=dddbbf&xu=96487001&omradekod=BE&omradetext=Befolkning&lang=1&langdb=1. [Accessed November 12, 2008].
[30] The EuroQol Group. EQ-5D products. September 28, 2011. Available from: http://www.euroqol.org/eq-5d/eq-5d-products.html. [Accessed September 28, 2011].
[31] HUInc. HEALTH UTILITIES INC An Overview 2011. Available from: http://www.healthutilities.com/. [Accessed September 28, 2011].
[32] Dolan P, Gudex C, Kind P, et al. A Social Tariff for EuroQol: Results from a UK General Population Survey. Discussion Paper No. 138. York: Centre for Health Economics, University of York, 1995.
[33] Mangione CM, Lee PP, Gutierrez PR, et al. Development of the 25-item National Eye Institute Visual Function Questionnaire. Arch Ophthalmol 2001;119:1050–8.
[34] Dolan P, Gudex C, Kind P, et al. The time trade-off method: results from a general population study. Health Econ 1996;5:141–54.
[35] White H. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 1980;48:817–38.
[36] Pullenayegum EM, Tarride JE, Xie F, et al. Analysis of health utility data when some subjects attain the upper bound of 1: are Tobit and CLAD models appropriate? Value Health 2010;13:487–94.
[37] Neter J. Applied Linear Statistical Models (4th ed). Chicago, IL: Irwin, 1996: 1408 pp.
[38] Brazier J, Deverill M. A checklist for judging preference-based measures of health related quality of life: learning from psychometrics. Health Econ 1999;8:41–51.
[39] Petrou S, Morrell J, Spiby H. Assessing the empirical validity of alternative multi-attribute utility measures in the maternity context. Health Qual Life Outcomes 2009;7:40.
[40] Grutters JP, Joore MA, van der Horst F, et al. Choosing between measures: comparison of EQ-5D, HUI2 and HUI3 in persons with hearing complaints. Qual Life Res 2007;16:1439–49.
[41] Luo N, Johnson JA, Shaw JW, et al. Relative efficiency of the EQ-5D, HUI2, and HUI3 index scores in measuring health burden of chronic medical conditions in a population health survey in the United States. Med Care 2009;47:53–60.
[42] Petrou S, Hockley C. An investigation into the empirical validity of the EQ-5D and SF-6D based on hypothetical preferences in a general population. Health Econ 2005;14:1169–89.
[43] Marra CA, Woolcott JC, Kopec JA, et al. A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. Soc Sci Med 2005;60:1571–82.
[44] Fayers PM, Machin D, Quality of Life: Assessment, Analysis and Interpretation. Chichester, UK: Wiley, 2000.
[45] Cohen J. A power primer. Psychol Bull 1992;112:155–9.
[46] Rosner B. Fundamentals of Biostatistics (6th ed.). Belmont, CA: Thomson-Brooks/Cole, 2006.
[47] Burstrom K, Johannesson M, Diderichsen F. A comparison of individual and social time trade-off values for health states in the general population. Health Policy 2006;76:359–70.
[48] Zethraeus N, Johannesson M. A comparison of patient and social tariff values derived from the time trade-off method. Health Econ 1999;8: 541–5.
[49] Polsky D, Willke RJ, Scott K, et al. A comparison of scoring weights for the EuroQol derived from patients and the general public. Health Econ 2001;10:27–37.
[50] Brazier J, Ratcliffe J, Salomon J, et al. Measuring and Valuing Health Benefits for Economic Evaluation. Oxford: Oxford University Press, 2007.
[51] Drummond MF, Sculpher MJ, Torrance GW, et al. Methods for the Economic Evaluation of Health Care Programmes (3rd ed.). Oxford: Oxford University Press, 2005.
[52] Childs AL. Effect of order of administration of health-related quality of life interview instruments on responses. Qual Life Res 2005;14:493–500.
[53] McColl E, Eccles MP, Rousseau NS, et al. From the generic to the condition-specific? Instrument order effects in quality of life assessment. Med Care 2003;41:777–90.
[54] Rat AC, Baumann C, Klein S, et al. Effect of order of presentation of a generic and a specific health-related quality of life instrument in knee and hip osteoarthritis: a randomized study. Osteoarthritis Cartilage 2008;16:429–35.
[55] McPhail S, Lane P, Russell T, et al. Telephone reliability of the Frenchay Activity Index and EQ-5D amongst older adults. Health Qual Life Outcomes 2009;7:48.
[56] van Wijck EE, Bosch JL, Hunink MG. Time-tradeoff values and standard-gamble utilities assessed during telephone interviews versus face-to-face interviews. Med Decis Making 1998;18:400–5.
[57] Wu AW, Jacobson DL, Berzon RA, et al. The effect of mode of administration on medical outcomes study health ratings and EuroQol scores in AIDS. Qual Life Res 1997;6:3–10.
[58] Szabo SM, Beusterien KM, Pleil AM, et al. Patient preferences for diabetic retinopathy health states. Invest Ophthalmol Vis Sci 2010;51: 3387–94.
[59] McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: what it is and what that means. Pharmacoeconomics 2008;26:733–44.