Eur J Health Econom 2003 · 4:222–231
DOI 10.1007/s10198-003-0182-5
Published online: 16. July 2003
© Springer-Verlag 2003

Wolfgang Greiner[1] · Tom Weijnen[2] · Martin Nieuwenhuizen[2] · Siem Oppe[2] · Xavier Badia[3] Jan Busschbach[4] · Martin Buxton[5] · Paul Dolan[6] · Paul Kind[7] · Paul Krabbe[8] · Arto Ohinmaa[9] David Parkin[10] · Montserat Roset[3] · Harri Sintonen[11] · Aki Tsuchiya[12] · Frank de Charro[2]
[1] Institut für Versicherungsbetriebslehre, University of Hanover, Germany · [2] Centre for Health Policy and Law, Erasmus University, Rotterdam, The Netherlands · [3] Health Outcomes Research Group, Barcelona, Spain · [4] Department of Medical Psychology and Psychotherapy, Erasmus MC, The Netherlands · [5] Health Economics Research Group, Brunel University, Uxbridge, UK · [6] ScHARR, University of Sheffield, UK · [7] Outcomes Research Group, Centre for Health Economics, York University, UK · [8] Department of Medical Technology Assessment, University Medical Centre, Nijmegen, The Netherlands · [9] Department of Public Health, University of Alberta, Edmonton, Canada [10] Department of Economics, City University, London, UK · [11] Department of Public Health, University of Helsinki, Helsinki, Finland · [12] SHEG, ScHARR, University of Sheffield, Sheffield, UK

A single European currency for EQ-5D health states

Results from a six-country study

The EQ-5D questionnaire is a standardised, generic instrument for describing and valuing health that was designed by the EuroQol Group (an international research network established in 1987) [1, 2]. The EQ-5D self-report questionnaire (commonly referred to as EQ-5D) consists of a five-dimensional descriptive system and a visual analogue scale (EQ VAS) together with brief demographic questions. The descriptive system defines health in terms of five dimensions: mobility, self-care, usual activities, pain or discomfort, and anxiety or depression. Each dimension is divided into three levels, indicating no problem, some or moderate problems or extreme problems. Respondents record their problem in each of the five dimensions. Combinations of these levels defines a total of 243 health states. A health state defined by the descriptive system of EQ-5D can be described by a five-digit number. For example, 12113 would imply no problems with mobility, usual activities and pain/discomfort, some problems with self-care, and extreme problems with anxiety/depression. According to this system, full health is indicated by 11111 and the poorest health state by 33333. While the EQ-5D descriptive system asks respondents to *describe* their health status, the EQ VAS asks respondents to *rate* their perception of their overall health on a vertical visual analogue scale with 'best imaginable health state' set at 100 and 'worst imaginable health state' set at 0.

Translated into more than 50 languages, EQ-5D is now in widespread use in many countries and has been applied in many different settings. EQ-5D is an integral feature of many clinical trials and economic studies and is increasingly used in population health surveys.

The EQ-5D valuation questionnaire

EQ-5D can be used in a variety of ways one of which is to represent health-related quality of life as a single summary index. This is achieved by applying scores from a standard set of general population preference weights to health states defined by the EQ-5D descriptive system. One of the primary objectives of the EuroQol Group is the investigation of values of health states in the general population in different countries. For this purpose the EuroQol Group designed an valuation questionnaire. Respondents are first requested to fill in the EQ-5D and then to perform a valuation task consisting of two pages where they are asked to value a number of hypothetical health states defined by the EQ-5D descriptive system on a vertical scale. This scale, known as the EQ-5D VAS, is the same standard visual analogue as the EQ VAS with the same endpoints of 'best imaginable' and 'worst imaginable' health. Finally respondents are asked a standard set of socio-demographic questions including age, sex, education, social status, present health state and smoking (EQ SDQ).

Over the years members of the EuroQol Group have conducted general population surveys in different countries using the EQ-5D VAS to collect values for health states. This contribution reports on an analysis of the data sets collected by members of the EuroQol Group in Finland, Germany, The Netherlands, Spain, Sweden and the United Kingdom. The data have been brought together and harmonised within the framework of the EQ-net project funded by the European Union under the Biomed II scheme. The objectives of the project were to investigate differences in the valuations of EQ-5D health states across European countries and to explore the option of describing the values for EQ-5D health states by a common model. Earlier studies had indicated that Western European industrialised countries share similar values [2]. A more recent paper comparing the Finnish and US EQ-5D VAS valuations concluded that differences in valuations of EQ-5D health states are generated in Finland and the United States [3]. However, these country-specific differences were not large.

The EQ-net project compared European EQ-5D VAS valuations. It was concluded that there is considerable agreement between health state valuations using the EQ-5D VAS across European countries. This implies that it is a priori possible to determine a European EQ-5D VAS-based set of values. This paper explores whether a standard set of general population preference weights can be derived from existing European data. This set of general population preference weights can then potentially be applied to calculate a European VAS-based health status index for all the potential 243 EQ-5D health states defined by the EQ-5D descriptive system. The pooled data set was derived from 11 studies, and the analysis was performed on 82,910 VAS valuations of 44 EQ-5D health states elicited from 6,870 respondents. A multi-level regression model was applied to estimate the European set of EQ-5D preference weights.

Material and Methods

The EuroQol valuation questionnaire was used to collect data on EQ-5D health states, usually via postal surveys. The standard set of health states collected consists of 13 'core' EQ-5D health states, selected by the EuroQol Group in the 1980s to cover as wide a spectrum of health status as possible and includes the best and worst states defined by EQ-5D, i.e. 11111 and 33333, respectively. These two states are repeated on both pages of the valuation task to act as common reference points. Respondents are also requested to value unconscious and to value dead twice in comparison with the other states. Research initially focused on eliciting valuations for this standard set of health states, but over the years EQ Group members have gone further, eliciting valuations for additional states generated by the EQ-5D descriptive system.

Of the studies reported here six were postal and five interview-based. Of the interview-based studies three elicited VAS valuations in combination with the time tradeoff (TTO) method. In these studies health states were ranked before the VAS valuation task was performed. The analysis reported here was carried out on the pooled data set, which contains data from the 11 studies from six European countries. Details of the 11 data sets are as follows.

The Netherlands

The Dutch data (NL) were collected via a postal survey in Rotterdam in January 1991. The sample of 1,400 households was randomly selected from the population living on the right bank of the River Maas in Rotterdam. Districts with over 20% immigrants were excluded because of expected language problems. The sample is therefore not representative of the population of The Netherlands. After two reminders, 993 questionnaires were returned, yielding a response rate of 70.9%. In the Dutch study a total of 28 health states (including unconscious) were valued. 'Dead' was not included. The 13 core states (including 11111 and 33333) valued in the Dutch study are included in the current analysis [4, 5].

United Kingdom

The Frome IV study (UK-1), was a postal survey carried out in a general practice population in Frome, Somerset, in June–August 1991. Questionnaires were posted to 340 persons aged 16 years and over, chosen randomly from the computerised register of a Frome general practice. The response rate was 35.9% (122 questionnaires). The respondents valued the 13 core states (including 11111 and 33333) as well as unconscious and dead [6].

A second UK data set resulted from the MVH study (UK-2) [7, 8] carried out at York University, UK, and was based on a representative sample of the non-institutionalised adult population of England, Scotland and Wales. From the postcode address file 6,080 addresses were selected [9]. This resulted in a total of 5,324 in-scope addresses and 3,395 (63.7%) respondents who agreed to be interviewed.

Abstract

Eur J Health Econom 2003 · 4:222–231
DOI 10.1007/s10198-003-0182-5
© Springer-Verlag 2003

**Wolfgang Greiner · Tom Weijnen
Martin Nieuwenhuizen · Siem Oppe
Xavier Badia · Jan Busschbach
Martin Buxton · Paul Dolan · Paul Kind
Paul Krabbe · Arto Ohinmaa · David Parkin
Montserat Roset · Harri Sintonen
Aki Tsuchiya · Frank de Charro**

A single European currency for EQ-5D health states. Results from a six-country study

Abstract

The EQ-5D questionnaire is a widely used generic instrument for describing and valuing health that was developed by the EuroQol Group. A primary objective of the EuroQol Group is the investigation of values for health states in the general population in different countries. As part of the EuroQol enterprise 11 population surveys were carried out in six Western European countries (Finland, Germany, The Netherlands, Spain, Sweden and the UK) to value health states as defined by the EQ-5D using a standardised visual analogue scale (EQ-5D VAS). This contribution reports how a European set of general population preference weights was derived from the data collected in the 11 valuation studies. The scores from this set of preference weights can be applied to generate a VAS-based weighted health status index for all the potential 243 EQ-5D health states for use in multi-national studies. To estimate the preference weights a multi-level regression analysis was performed on 82,910 valuations of 44 EQ-5D health states elicited from 6,870 respondents. Stable and plausible solutions were found for the model parameters. The R^2 value was 75%. The analysis showed that the major source of variance, apart from 'random error', was variance between individuals (28.3% of the total residual variance). These results suggest that VAS values for EQ-5D health states in six Western European countries can be described by a common model.

Keywords

Quality of life · EQ-5D · Health status index
Visual analogue scale

Table 1

Response rates, exclusions, inconsistencies and number of health states per study

	Finland	Ger-1	Ger-2	Ger-3	Neth	Sp-1	Sp-2	Sp-3	Sweden	UK-1	UK-2	Total
Sample	2,530	1,000	1,000	4,000	1,400	682	–	–	1,000	340	5,324	
Response rate	64.4%	37.0%	12.4%	8.5%	70.9%	88.0%	–	–	54.2%	35.9%	63.7%	
n	1,630	370	124	339	993	600	300	294	542	122	3,395	8,709
Exclusions												
Due to incomplete data	207 (12.7%)	161 (43.5%)	19 (15.3%)	–	330 (33.2%)	–	–	–	114 (21.0%)	2 (1.5%)	40 (1.2%)	873 (10.0%)
Due to more than three inconsistencies	485 (29.8%)	25 (6.8%)	11 (8.9%)	6 (1.8%)	132 (13.3%)	3 (0.5%)	34 (11.3%)	7 (2.4%)	59 (10.9%)	24 (19.7%)	180 (5.3%)	966 (11.1%)
Number of respondents not excluded	938 (57.5%)	184 (49.7%)	94 (75.8%)	333 (98.2%)	531 (53.5%)	597 (99.5%)	266 (88.7%)	287 (97.6%)	369 (68.1%)	96 (78.7%)	3,175 (93.5%)	6,870 (78.9%)
Number of health states in the analysis	32	13	13	36	13	11	13	43	13	13	43	44
Number of valuations in the analysis	10,341	2,213	1,206	4,411	4,028	6,558	3,455	3,729	4,664	1,211	41,094	82,910

Trained interviewers conducted the interviews between August and November 1993. At each interview the respondent classified his/her own health according to the EQ-5D descriptive system and rated his/her health on the EQ VAS. Next, respondents ranked the health states and valued the states on the EQ-5D VAS. Afterwards the respondents continued by valuing the same states using TTO (reported elsewhere). Each respondent valued a selection of health states drawn from a subset of 43 EQ-5D health states. All respondents valued two very mild states, three mild states, three moderate states and three extreme states. They also valued states 11111 and 33333. Unconscious and dead were also valued.

Finland

The Finnish valuation questionnaires were sent by post in November 1992 to a sample of 2,530 randomly chosen persons over 16 years of age. After two reminders the response rate was 64.4% (1,630 questionnaires). The sample was divided into 17 sub-samples who received different versions of the questionnaire. Some health states were common to every questionnaire, while some were unique to various sub-samples. The Finnish study elicited valuations for 43 EQ-5D health states (including 11111 and 33333). Unconscious and dead were also valued [3, 10]. This analysis included 32 states. These were states common to the UK-2 sample

Spain

Three data sets were collected in Spain. The first was collected in November-December 1992 in a primary health care centre in l'Hospitalet del Llobregat (Sp-1) in Catalonia. Physicians invited patients to join the study. If someone agreed to participate, he or she was sent to an adjacent room and asked to fill in the EQ-5D questionnaire under the supervision of a trained nurse. A total of 682 patients were asked to participate. The sample size was fixed at 600 valid questionnaires, a response rate of 88.0%. A quota sampling procedure was used with gender, age, occupational class, and patient/non-patient status as control variables. Twelve of the 13 core health states (including 11111 and 33333) were valued as well as an additional state (22322). Respondents also valued unconscious and dead [11]. This study included 11 states, 10 of which were core states.

The second data set was collected in Navarra (Sp-2) drawn from the general population using quota sampling by age and sex. In September 1995, 300 respondents were interviewed at home by a qualified interviewer. The interviewers assisted only when the respondents requested clarification. Twelve of the 13 core states were valued (including 11111 and 33333) as well as an additional state (22322). Respondents also rated unconscious and dead [12].

The third Spanish study was conducted in Cornellà de Llobregat (Sp-3). Two trained interviewers recruited a random sample of 294 individuals, both patients and non-patients, attending a primary care centre during October–December 1996. Individuals were recruited according to age and sex quota sampling so as to be representative of the Catalan general population. The UK-2 study was replicated with each respondent valuing a selection of health states drawn from a subset of 43 EQ-5D health states (including 11111 and 33333). Unconscious and dead were also valued [13].

Sweden

The Swedish data were elicited from a nationwide sample of 1,000 citizens drawn from the address register. The ages ranged

Table 2

The dummy variables and their values in the model

Dummy for	EQ-5D dimension rated on		
	Level 1	Level 2	Level 3
Study methodology			
Ranked, interview-based studies with UK-2 protocol design	RID=0 or 1	RID=0 or 1	RID=0 or 1
Ratings on at least one EQ-5D dimension			
At least one dimension=2 or 3	N2=0	N2=1	N2=1
At least one dimension=3	N3=0	N3=0	N3=1
Interaction			
RID studies ×N2	RN2=0	RN2=0 or 1	RN2=0 or 1
RID studies ×N3	RN3=0	RN3=0	RN2=0 or 1
Ratings on the EQ-5D descriptive system			
Mobility	Mo2=0 Mo3=0	Mo2=1 Mo3=0	Mo2=0 Mo3=1
Self-care	Se2=0 Se3=0	Se2=1 Se3=0	Se2=0 Se3=1
Usual activities	Ua2=0 Ua3=0	Ua2=1 Ua3=0	Ua2=0 Ua3=1
Pain/discomfort	Pd2=0 Pd3=0	Pd2=1 Pd3=0	Pd2=0 Pd3=1
Anxiety/depression	Ad2=0 Ad3=0	Ad2=1 Ad3=0	Ad2=0 Ad3=1

from 18 to 78 years. The questionnaires were posted in April 1994. A total of 542 questionnaires were returned (54.2%). The respondents valued the 13 core states (including 11111 and 33333) as well as unconscious and dead [14].

Germany

Of the three German datasets two collected values for health states via postal surveys. In the first study (Ger-1) 1,000 households across Germany were randomly sampled by the German National Telephone Company. The questionnaires were sent out by mail in June 1994, and 370 (37.0%) were returned. Respondents valued 12 of the 13 core health states (including 11111 and 33333) as well as an additional state (12121). Respondents also rated unconscious and dead [15].

Another German study (Ger-2) following the protocol from the earlier 1994 study was carried out in April 1997. Again, 1000 households were randomly selected from the register of the German National Telephone Company. In this study 124 questionnaires (12.4%) were returned. Respondents valued the 13 core states plus unconscious and dead [16].

A third German study (Ger-3) collected VAS data in a multi-method framework in which TTO values were also collected. A random sample of 4,000 addresses was selected from the telephone directory of the German National Telephone Company. In order to represent urban and rural regions equally, the selection was based on postal code areas. It was ascertained that one-half of the selected persons were women. In this study 339 persons (8.5%) agreed to be interviewed. The interviews took place between October 1997 and March 1998 in northern Germany. Respondents were asked to value up to 15 different health states from a sample of 43 states. The participants were first requested to give brief information about their socio-economic background. They were then given selected cards with a description of the health states. They were then asked to rank these cards before carrying out the VAS valuation task. A total of 36 health states (including 11111 and 33333) were valued. Dead and unconscious were not included [17]. Seven states valued in the UK-2 sample were not included in this German sample; otherwise the two studies valued the same health states.

Pooling these 11 data sets resulted in the EQ-net VAS database of 124,077 VAS valuations from 8,709 respondents. Most of the studies valued the health states 11111, 33333 and dead twice. The mean value was calculated for each respondent from the two valuations of the same health state. Only these mean values were used. As a result the analyses were performed on 114,220 valuations. Unconscious was not included in the data set.

Exclusions

Data from some respondents were considered to be unreliable and were therefore excluded from the analyses. Every effort was made, however, to ensure that the number of exclusions was as low as possible. Data from respondents who gave all states the same value or who valued fewer than three states were excluded. In addition, responses were excluded if they contained more than three inconsistencies. Ratings of two health states which were not in accordance with their rank order were counted as one inconsistency; for example, if the rating for 11113 was higher than the rating for 11112.

The rules for excluding inconsistent responses were derived on the basis of the conclusions by Ohinmaa and Sintonen [18]. They investigated the effect of the number of inconsistencies per respondent on the values resulting from the model and concluded that inconsistencies did not affect the modelling results in the Finnish data as long as the number of inconsistencies was three or lower. The analysis of the number of inconsistencies on modelling results was also tested for the values in the EQ-net VAS database. The Finnish findings were confirmed.

Table 1 shows the number of exclusions and inconsistencies that result from the application of the criteria. The exclusions due to incomplete data refer to respondents who gave all states the same value or who valued fewer than three states.

The EQ-5D database consists of data from 8,709 (100%) respondents from 11 different studies. A total of 873 (10.0%) respondents were excluded because they

Table 3

Results from the multilevel regression on the VAS values

	All coefficients	Aggregated coefficients
Constant	95.71*	97.66
RID studies	1.96*	
At least one 2 or 3 (N2)	−19.52*	−11.21
At least one 3 (N3)	−8.58*	−20.06
Interaction N2 and RID studies	8.30*	
Interaction N3 and RID studies	−11.48*	
Mobility=2	−5.78*	−5.78
Mobility=3	−16.03*	−16.03
Self care=2	−10.28*	−10.28
Self care=3	−13.67*	−13.67
Usual activities=2	−2.31*	−2.31
Usual activities=3	−7.54*	−7.54
Pain/discomfort=2	−8.15*	−8.15
Pain/discomfort=3	−14.35*	−14.35
Anxiety/depression=2	−7.81*	−7.81
Anxiety/depression=3	−11.31*	−11.31
R-square	0.745	–
AdjustedR^2	0.745	–
Multi-level variance components in shares of total variance (%)		–
Study	1.6	–
Respondent	28.3	–
Residual	70.1	–

*$P<0.001$

gave all states the same value or valued fewer than three states. We excluded 966 (11.1%) respondents because they returned questionnaires that contained more than three inconsistencies. Therefore a total of 1,839 (21.1%) respondents were excluded before the analysis was performed. Consequently 6,870 (78.9%) respondents are included in the analysis.

The number of exclusions was high in the Finnish, 1994 German (Ger-1), Dutch and Swedish samples. The number of exclusions due to more than three inconsistencies was especially high in the Finnish and Dutch data but low in two Catalonian samples (Sp-1, Sp-3), the UK-2 data and the 1998 German (Ger-3) sample, which are all interview-based studies.

After exclusions there were a total of 100,068 valuations from 6,870 respondents. The values for dead and unconscious could not be included in the regression model because these states are not defined in terms of the EQ-5D descriptive system. Also, valuations of health states which were rated in only a single study were not used in the analysis. This was to avoid a study-related bias in the overall results. In the Sp-1 study there was evidence of a translation problem in the wording for level 3 on the mobility dimension. Therefore the valuations of two health states with level three on mobility from this sample were excluded.

As a result, the regression analysis presented here was performed on 82,910 valuations. The 'Appendix' presents an overview of these valuations per health state and per sample.

Modelling

A regression analysis was performed on the individual data of all non-excluded respondents in the 11 studies. A multi-level (random effects) model was used to allow for differences in scale use by respondents [19]. A multi-level analysis assumes that the data are hierarchically structured. The EQ-net data set distinguished three levels. The first level concerned the evaluations of the health states. The second concerned respondents. Therefore evaluations of health states were nested within respondents. The third level concerned the studies. Respondents were nested within the studies. In a multi-level regression model the error variance is comprised three components: random error, error due to individual differences and error due to differences in study design. If systematic differences exist at the level of the respondents and/or the study, the multi-level regression analysis estimates the parameters in a more precise manner. The multilevel regression analysis was performed using MlwiN 1.02 software.

We used the same main effects model that was used in the UK-2 study [20]. The model describes how the valuation of a health state depends on the scores of the EQ-5D five-dimensional classification system for that particular health state. Background variables, i.e. socio-economic, individual, regional and country characteristics, are not included. A dummy variable (RID) was included for ranked, interview-based study design using a protocol originally developed for the UK-2 study in 1994. In this protocol the duration of the health states under consideration was 10 years [8]. RID was used to capture the implications of differences in study design on the valuations and indicates whether the data derived from one of the three interview-based studies in which the respondents ranked the health states prior to the VAS valuation task. Another dummy variable (N2) accounted for any move away from full health. This variable had the value '1' if one of the dimensions was not '1'. In addition, another dummy variable (N3) had the value '1' if one of the dimensions was at level 3. The N2 and N3 dummy variables were added to measure overall aspects not expressed by the scores on the individual dimensions. The model allowed for interaction between the RID variable and the N2 and N3 variables. This interaction was represented by the RN2 and RN3 dummy variables. The scores on the 5 EQ-5D dimensions were represented by ten dummy variables. For each of the five dimensions of EQ-5D there was a dummy variable which had the value '1' if the score on the dimension was at level 2 and another dummy variable with the value '1' if the score on the dimension was at level 3. **Table 2** summarises the model.

The full model can be expressed by the following equation:

$$\begin{aligned} VAS_{\text{health state x}} = {} & C + c_1RID + c_2N2 + c_3RN3 \\ & + c_4RN2 + c_5RN3 \\ & + c_6Mo2 + c_7Mo3 + c_8Se2 \\ & + c_9Se3 + c_{10}Ua2 + c_{11}Ua3 \\ & + c_{12}Pd2 + c_{13}Pd3 \\ & + c_{14}Ad2 + c_{15}Ad3 \end{aligned}$$

where $VAS_{\text{health state x}}$ is the VAS value for health state X predicted by the model, C is the constant, and c_1–c_{15} are the 15 dummy variables.

Rescaling

For particular applications, for example, in cost-effectiveness analysis, where quality-adjusted life years (QALYs) are to be computed the quality index must be anchored on full health and dead with corresponding values of '1' and '0'. The standard VAS scale in EQ-5D has endpoints that are labelled 'best imaginable' and 'worst imaginable'. Raw and estimated scores generated using this VAS do not conform to the 0–1 scale requirement for QALY computations and need to be rescaled. This was carried out using average values for dead applied to the following formula 3 (both mean and median values for dead were used, these values being 0.1 and 0.02, respectively):

$$X_{res} = \frac{100.(X - dead_{\text{mean or median}})}{(11111 - dead_{\text{mean or median}})}$$

where X_{res} is the rescaled VAS value for health state X, X is the VAS value for health state X, $dead_{mean/med}$ is the mean or median value for dead, and 11111 is the value for full health.

Results

Multi-level regression analysis

The results of the multi-level regression analysis are given in **Table 3.** The full model is presented in the first column. In the last column of **Table 3** the values are recalculated for the cases in which the RID=1. This column presents the model coefficients from the point of view that the study design includes ranking and is interview-based. The 'Appendix' gives an overview of means and medians of observed VAS values of the estimated VAS values and of the differences between observed and estimated values.

Of the total variance 28.3% was due to individual differences and 1.6% to study differences. The goodness-of-fit (R^2) was 0.745. All regression weights were significant at $P<0.001$. For all the dimensions of the EQ-5D descriptive system the coefficients indicating the change in the index because of a change from the first level (no problems) to the second level (some problems) or vice versa were smaller than the coefficients associated with a change on an EQ-5D dimension from level 2 (some problems) to level 3 (extreme problems). All the coefficients had the expected signs. **Table 3** shows the results from the multilevel regression on the VAS values.

The valuations of mild states in the RID studies tended to be higher than in the non-RID studies; the RID valuations of severe states tended to be lower. In the RID studies the N3 variable, which indicated at least one 3 on the descriptive system, together with the variable which indicated interaction of N3 with RID, showed the greatest effect (–20.06). When the first 3 appeared on the descriptive system the predicted VAS value declined by 33 points, in addition to the relevant score on the dimension on which the 3 occurred.

Figure 1 shows the estimated VAS values for the observed health states compared to the mean and the median values of these health states. The health states are ordered according to the mean. For mild health states the estimated values are close to the median values, but higher in general. For health states with lower values the estimated values were lower than the mean values. Figure 1 also shows the difference between the observed mean and median values and the estimated values. The estimation procedure was based on the minimisation of differences between observations and the mean and not the median values. **Figure 1** indicates the health states which showed differences smaller than –6 or greater than 6.

Rescaling

The predicted values using the mean for death were rescaled for the model results where RID=1. This procedure generated rescaled values for all EQ-5D health states. Of the 243 health states generated by the EQ-5D descriptive system rescaled values for five health states were negative. These states were: 33333, 32333, 33233, 33323 and 33332. Because there were no estimated values lower than 6.5 (for 33333), rescaling of the predicted values using the median of the observations for dead (0.02 or 2%) did not result in any negative rescaled values.

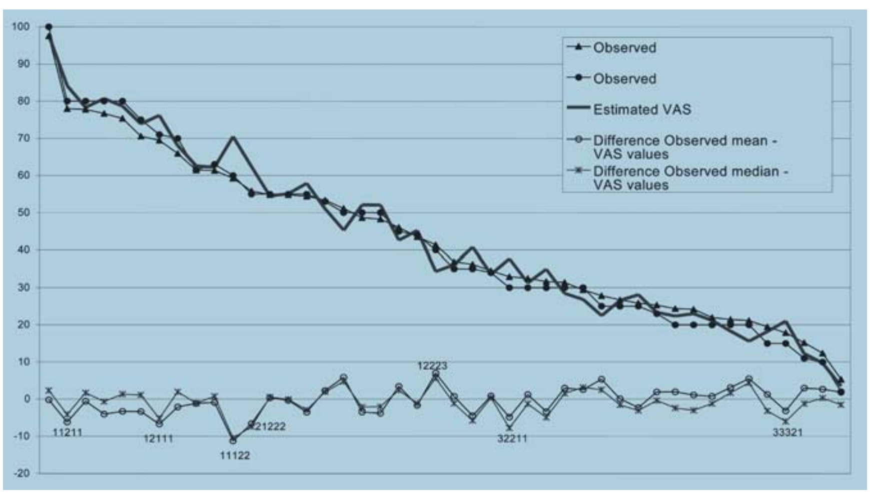

Fig. 1 ▲ **Comparison of the estimated with the mean and median observed VAS values**

Table 4

Observed mean and median values for the sub set of standard health states. Estimated and rescaled estimated values and differences between the observed mean and estimated values

Observed and estimated VAS values				Rescaled estimated values Rescaled with		Differences between observed			
						Mean and rescaled estimated values, rescaled with			Median and rescaled estimated values, rescaled
State	Mean	Median	Estimated	Mean value of dead	Median value of dead	Estimated values	Mean dead	Median dead	Median dead
11111	97.5	100.0	97.7	100.0	100.0	–0.2	–2.5	–2.5	0.0
11211	78.0	80.0	84.1	84.6	85.9	–6.2	–6.6	–7.9	–5.9
11121	77.8	80.0	78.3	77.9	79.8	–0.5	–0.2	–2.0	0.2
21111	76.7	80.0	80.7	80.6	82.2	–4.0	–4.0	–5.6	–2.2
11112	75.3	80.0	78.6	78.3	80.1	–3.3	–3.0	–4.8	–0.1
12111	69.4	71.0	76.2	75.5	77.5	–6.7	–6.0	–8.1	–6.5
11122	59.3	60.0	70.5	69.0	71.6	–11.2	–9.7	–12.3	–11.6
21232	36.9	35.0	36.1	29.8	35.7	0.8	7.1	1.2	–0.7
32211	32.9	30.0	37.8	31.7	37.4	–4.8	1.3	–4.5	–7.4
22323	25.3	23.0	23.3	15.2	22.3	1.9	10.1	3.0	0.7
22233	24.3	20.0	22.4	14.1	21.3	2.0	10.2	3.0	–1.3
33321	17.9	15.0	21.0	12.5	19.9	–3.1	5.3	–2.0	–4.9
33333	5.3	2.0	3.5	–7.4	1.6	1.8	12.7	3.8	0.4
Sum of absolute differences						46.6	78.6	60.6	42.0
Mean sum of absolute differences						3.6	6.0	4.7	3.2

The results of the rescaling procedure are shown in **Table 4** for the subset of the 13 core states, which were valued in nearly all studies. The estimated value of 97.7 for health state 11111 was very close to the rescaled value of 100. Rescaling using the mean for dead had considerable impact on scores under 50. The results of the rescaling procedure using the median value for dead were very close to the original non-rescaled values.

Table 4 also shows the differences between the observed mean and rescaled estimated values and the observed median and estimated values that were rescaled with the median value of dead. The sum of absolute differences of the rescaled values was lowest for the values that were rescaled with the median value for dead.

Discussion

Values for health states defined by the EQ-5D descriptive system were collected in 11 studies conducted in six European countries. A total of some 83,000 observations were successfully modelled using multilevel regression. All coefficients that resulted had the expected signs and the majority of the variance was explained mainly by the difference in the health states involved (R^2=74.5%). Of the total explained variance only 1.6% was due to differences in study design. There was considerable variation between respondents, some of whom systematically valued health lower or higher than was suggested by the model results. This pattern of valuation has been seen previously and is sometimes associated with different experiences of ill health.

A multi-level (random effects) model was used to include individual and study differences other than pre-ranking. The same data were analysed using a weighted ordinary least squares (OLS) model where the weights expressed the differences in volume of data per study. Estimated values of the health states calculated with the multi-level model were close to those estimated via the weighted OLS regression. The average absolute difference between the estimated values was 1.2 points over a range from 98 to 4. This observation is consistent with the findings of Busschbach et al. [21].

The main effects regression model includes only the scores on the descriptive system of the EQ-5D. An additional dummy variable was used to capture an important difference in study design. In three studies respondents were allowed to rank health states *before* rating them on a VAS scale. Although this method is more resource intensive both in terms of interview time and respondent burden, it does enable individuals to gain familiarity with the health states and to think longer about the valuation task. In the studies performed by the EuroQol Group ranking was permitted in studies that were interview-based. However, for pragmatic reasons it is not always feasible to use interviewers to collect data on the value of health states. Within the limits of the present data set it is not possible to make an objective judgement on the general issue of whether pre-ranking is preferable in studies collecting VAS values of health states.

Ranking and interview-based methods generate values which are higher for better health states and lower for worse health states. Approximately one-half of the valuations in the database were gen-

Table 5

Number of valuations per sample in the analysis after exclusions (*subtotal* the total of the core states in the analysis)

	Finland	Ger-1	Ger-2	Ger-3	Neth	Sp-1	Sp-2	Sp-3r	Sweden	UK-1	UK-2	Total
11111	908	183	93	331	524	597	266	287	369	95	3,174	6,827
11112	509	170	93	146	275	597	266	121	357	92	1,241	3,867
11121	495	170	92	166	272	597	266	119	362	94	1,279	3,912
11122	493	172	92	91	271	596	266	69	364	93	787	3,294
11211	500	171	91	162	270	597	266	112	366	94	1,286	3,915
12111	513	164	92	151	274	595	266	113	352	94	1,252	3,866
21111	519	–	94	147	266	596	266	109	355	92	1,268	3,712
21232	494	171	92	87	270	596	264	83	359	94	797	3,307
22233	483	165	94	–	271	596	266	76	359	93	792	3,195
22323	515	165	93	97	276	–	–	70	349	92	777	2,434
32211	512	166	92	135	274	–	266	66	353	92	781	2,737
33321	486	171	94	100	270	–	266	72	358	93	783	2,693
33333	890	173	94	333	515	597	266	286	361	93	3,160	6,768
Subtotal	7,317	2,041	1,206	1,946	4,028	5,964	3,190	1,583	4,664	1,211	17,377	50,527
11113	–	–	–	143	–	–	–	59	–	–	787	989
11131	–	–	–	109	–	–	–	76	–	–	773	958
11133	–	–	–	90	–	–	–	75	–	–	801	966
11312	174	–	–	155	–	–	–	61	–	–	791	1,181
12121	136	172	–	96	–	–	–	69	–	–	795	1,268
12211	175	–	–	94	–	–	–	88	–	–	797	1,154
12222	173	–	–	81	–	–	–	68	–		801	1,123
12223	136	–	–	96			–	66	–	–	789	1,087
13212	–	–	–	97	–	–	–	66	–	–	789	952
13311	–	–	–	–	–	–	–	75	–	–	780	855
13332	162	–	–	–	–	–	–	91	–	–	780	1,033
21133	132	–	–	79	–	–	–	77	–	–	796	1,084
21222	158	–	–	100	–	–	–	68	–	–	785	1,111
21312	–	–	–	–	–	–	–	67	–	–	785	852
21323	–	–	–	86	–	–	–	71	–	–	785	942
22112	159	–	–	92	–	–	–	80	–	–	785	1,116
22121	172	–	–	–	–	–	–	71	–	–	794	1,037
22122	161	–	–	91	–	–	–	69	–	–	775	1,096
22222	157	–	–	142	–	–	–	86	–	–	787	1,172
22322	–	–	–	–	–	594	265	–	–	–	–	859
22331	173	–	–	83	–	–	–	73	–	–	789	1,118
23232	160	–	–	140	–	–	–	69	–	–	782	1,151
23313	–	–	–	–	–	–	–	62	–	–	791	853
23321	158	–	–	86	–	–	–	74	–	–	790	1,108
32223	161		–	146	–	–	–	72	–	–	793	1,172
32232	162	–	–	96	–	–	–	73	–	–	799	1,130
32313	157	–	–	–	–	–	–	73	–	–	812	1,042
32331	–	–	–	93	–	–	–	65	–	–	793	951
33212	–	–	–	95	–	–	–	63	–	–	789	947
33232		–	–	93	–	–	–	64	–	–	791	948
33323	158	–	–	82	–	–	–	75	–	–	813	1,128
Total	10,341	2,213	1,206	4,411	4,028	6,558	3,455	3,729	4,664	1,211	41,094	82,910

Table 6

Mean, median and estimated values; differences between mean or median and estimates (VAS values of health states valued in more than one study)

State	Observed VAS values		Estimated VAS values	Observed difference	
	Mean	Median		Mean estimation	Median estimation
11111	97.5	100.0	97.7	–0.16	2.34
11211	78.0	80.0	84.1	–6.17	–4.14
11121	77.8	80.0	78.3	–0.54	1.70
21111	76.7	80.0	80.7	–4.01	–0.67
11112	75.3	80.0	78.6	–3.29	1.36
12211	70.5	75.0	73.9	–3.32	1.14
12111	69.4	71.0	76.2	–6.72	–5.17
12121	65.9	70.0	68.0	–2.10	1.98
22112	61.5	61.5	62.6	–1.13	–1.08
22121	61.3	63.0	62.2	–0.92	0.76
11122	59.3	60.0	70.5	–11.23	–10.49
21222	55.8	55.0	62.4	–6.58	–7.40
22122	54.8	55.0	54.4	0.37	0.57
11113	54.8	55.0	55.1	–0.30	–0.08
12222	54.4	55.0	57.9	–3.50	–2.90
11312	53.3	53.0	51.0	2.28	1.96
21312	51.1	50.0	45.3	5.86	4.74
22222	48.6	50.0	52.1	–3.48	–2.12
11131	48.2	50.0	52.0	–3.80	–2.04
13212	46.0	45.0	42.6	3.45	2.40
13311	43.5	44.0	45.2	–1.63	–1.18
12223	41.4	40.0	34.3	7.04	5.66
21232	36.9	35.0	36.1	0.76	–1.14
11133	36.2	35.0	40.7	–4.49	–5.73
21323	34.5	34.0	33.6	0.86	0.39
32211	32.9	30.0	37.8	–4.83	–7.77
23321	32.5	30.0	31.3	1.24	–1.25
21133	31.6	30.0	35.0	–3.32	–4.95
22331	31.4	30.0	28.4	2.97	1.56
22322	29.4	30.0	26.8	2.56	3.17
23232	27.8	25.0	22.5	5.35	2.53
33212	26.7	25.0	26.6	0.12	–1.57
23313	25.8	25.0	28.1	–2.26	–3.09
22323	25.3	23.0	23.3	1.95	–0.33
22233	24.3	20.0	22.4	1.97	–2.36
13332	24.2	20.0	23.0	1.13	–3.02
32313	22.0	20.0	21.2	0.73	–1.23
32223	21.4	20.0	18.3	3.12	1.69
32232	21.1	20.0	15.6	5.53	4.39
32331	19.5	15.0	18.2	1.27	–3.19
33321	17.9	15.0	21.0	–3.14	–6.00
33232	15.2	11.0	12.2	2.97	–1.22
33323	12.4	10.0	9.7	2.68	0.31
33333	5.3	2.0	3.5	1.82	–1.49
Sum of absolute differences				132.95	120.26
Mean absolute difference				3.02	2.73

erated in the more structured interview-based studies that incorporated pre-ranking. These studies reported lower levels of internal respondent inconsistency. This leads to the conclusion that ranking before valuation probably enhances the quality of the data. Researchers might wish to consider this finding in the design of any future valuation studies. Whether pre-ranking is technically possible in a postal survey is a topic currently under investigation within the EuroQol Group.

For some specific applications health state values must be represented on a scale on which dead takes the value '0' and full health the value '1'. In the present analysis rescaling was applied to predicted VAS values using an average value for dead. This method is considered defensible since individual respondent's values for dead vary in a way that appears to be independent of the value of the health states. Where data are rescaled on a within-respondent basis, the explanatory power of the model is reduced. The R^2 value decreased from 0.73 for the regression on the original values to 0.57 for the regression on the individually rescaled values. The concept of 'not being alive' is controversial [22], and empirical study of the value associated with dead remains incomplete. Given these attendant uncertainties it seems reasonable to use average values for dead in the rescaling of health state values.

Moreover, rescaling on a within-respondent basis may exceptionally produces r values greater than 1 and less than –1. The rescaled values less than –1 can be extremely negative, causing a significant increase in the variance around mean health state values. All the rescaled health states showed a high negative skewness. To counter these effects the negative values were truncated at –1, which was an arbitrary cut-off point. Where separate values for dead are not collected as part of the original valuation survey, it clearly becomes impossible to consider rescaling at the level of the individual respondent. Non-availability of values for dead may also arise from respondent failure to complete this task in postal questionnaire surveys or from partial completion of interview protocols.

The restriction of range of the raw VAS values between 0 and 100 was expected to produce skewed response distributions at the end of the scale. This phenomenon was expected to interfere with the linear regression analysis because values in such an analysis are assumed to be normally distributed. To solve this problem a logistic transformation was applied in which upper and lower bounds were adjusted accordingly. This procedure reduced skewness at both ends of the scale. However, the regression on the adjusted values resulted in a much lower R^2 value than in the regression on the unadjusted data. Although the transformation resolved the problem of skewed response distributions, it therefore also changed the nature of the linear additive model into a kind of multiplicative model. It can therefore be assumed that the underlying structure of the VAS valuations was indeed additive, and that the applied transformation made the situation worse. It was therefore decided to maintain the use of the linear regression model.

The results of the multi-level regression generally support the expected characteristics of the value set derived for EQ-5D health states. All the coefficients have the correct signs, while the values are plausible and all differ from '0'. For all the dimensions of the EQ-5D level 3 on the descriptive system has a stronger negative effect on the estimated VAS values than level 2. Given these characteristics of the value set and the highR^2 value in the regression analyses, the overall conclusion is that it is possible to describe this European VAS data set, gathered by the EuroQol Group, by a single common model. The resulting VAS value set can be applied to generate a weighted health status index for all the potential 243 EQ-5D health states. This will be especially valuable for multi-national and technology assessment studies and other applications where the comparability of data is important.

This contribution reports the analysis of pooled data from 11 European studies. The studies differed in terms of design, response rates and data quality. Furthermore, the composition of achieved samples was not always representative of the general population within each country. Hence the pooled data set may only partially capture societal preferences. Furthermore, data were available only for a subset of European countries. However, as far we know, this is the first study to report the values for standardised health states within a European context. The resulting model is internally consistent with plausible coefficients for EQ-5D health states. It is reasonable to hope that this model will perform equally well when data from other national population surveys are included. Evidence of such robustness would suggest that deriving a single set of European values of EQ-5D health states is indeed a practical proposition.

Corresponding author

Dr. Wolfgang Greiner

Institut für Versicherungsbetriebslehre,
Universität Hannover,
Königsworther Platz 1, 30167 Hannover, Germany
e-mail: wg@ivlb.uni-hannover.de

Acknowledgements

This research, performed on data collected previously, was supported by a grant from the European Community: Biomed II programme PL 963118, Concerted Action. The authors acknowledge the contributions of the participants of the Biomed EQ-net project, whose effort and expertise made it possible to produce this paper. We particularly thank Stefan Björk (Novo Nordisk A/S, Bagsvaerd, Denmark), Marie-Louise Essink-Bot (Department of Public Health, Erasmus University Rotterdam, The Netherlands) and Idoia Gaminde (Departamento de Salud, Gobierno de Navarra, Pamplona, Spain) for providing data. We also thank Rosalind Rabin who edited the English. Earlier versions of this manuscript were presented at the Biomed EQ-net scientific meetings.

Appendix

Tables 5 and 6 present, respectively, the number of valuations per sample and the mean, median and estimated values, with the differences between mean or median and estimates

References

1. Brooks R (1996) EuroQol: the current state of play. Health Policy 37:53–72
2. EuroQol Group (1990)EuroQol-a new facility for the measurement of health-related quality of life. Health Policy 16:199–208
3. Johnson J, Ohinmaa A, et al (2000) Comparison of Finnish and US-based visual analog scale valuations of the EQ-5D measure. Med Decis Making 20:281–289
4. Agt HM von, Essink-Bot ML, Krabbe PF, Bonsel GJ (1994) Test-retest reliability of health state valuations collected with the EuroQol questionnaire. Soc Sci Med 39:1537–1544
5. Essink-Bot ML, Stouthard MEA, Bonsel GJ (1993) Generalizability of valuations on health states collected with the EuroQol-questionnaire. Health Econ 12:237–346
6. Gudex C (1992) Are we lacking a dimension of energy in the EuroQol instrument? In: Björk S (ed) EuroQol conference proceedings. Lund, October 1991. Lund: IHE, pp 61–82
7. Gudex C, Dolan P, Kind P, Williams A (1996) Health state valuations from the general public using the visual analogue scale. Qual Life Res 6:521–531
8. Williams A (1995) The measurement and valuation of health: a chronicle. Discussion paper no 136. York: Centre for Health Economics
9. Gudex C, Dolan P, Kind P, Thomas R, Williams AH (1997) Valuing health states. Interviews with the general public. Eur J Public Health 7:441–448
10. Ohinmaa A, Eija H, Sintonen H (1996) Modelling EuroQol values of Finnish adult population. In: Badia X, Herdman M, Segura A (eds) EuroQol 1995 Barcelona Plenary Meeting, 3–6 October. Barcelona: Institut de Salut Pública de Catalunya, pp 67–76
11. Badia X, Fernández E, Segura A (1995) Influence of sociodemographic and health status variables on valuation of health states in a Spanish population. Eur J Public Health 5:87–93
12. Gaminde I, Cabasés J (1996) Measuring valuations for health states amongst the general population in Navarra (Spain). In: Badia X, Herdman M, Segura A (eds) EuroQol 1995 Barcelona Plenary Meeting, 3–6 October. Barcelona: Institut de Salut Pública de Catalunya, pp 113–122
13. Badia X, Roset M, Herdman M, Kind P (2001) A comparison of United Kingdom and Spanish general population time trade-off values for EQ-5D health states. Med Decis Making 21:7–16
14. Björk S, Norinder A (1999) The weighting exercise for the Swedish version of the EuroQol. Health Econ 8:117–126
15. Schulenburg J. Graf von der, Claes C, Greiner W, Uber A (1998) Die deutsche Version des EuroQol-Fragebogens. Z Gesundheitswiss 6:3–30
16. Claes C, Greiner W, Uber A, von der Schulenburg JM Graf (1998) The new German version of the EuroQol quality of life questionnaire. In: Rabin RE, Busschbach JJV, de Charro FT, Essink-Bot ML, Bonsel GJ (eds) EuroQol Plenary Meeting, Rotterdam 1997, 2–3 October. Rotterdam: Centre for Health Policy and Law, Erasmus University, pp 93–114
17. Claes C, Greiner W, Uber A, von der Schulenberg J-M Graf (1998) An interview-bases comparison of the TTO and VAS values given to EuroQol states of health by the general German population. In: Greiner W, von der Schulenburg JM Graf, Piercy J (eds) EuroQol Plenary Meeting, Hanover 1998, 1–2 October. Hanover: Uni-Verlag Witte, pp 13–38
18. Ohinmaa A, Sintonen H (1999) Inconsistencies and modelling of the Finnish EuroQol (EQ-5D) preference values. In: Greiner W, von der Schulenburg JM Graf, Piercy J (eds) EuroQol plenary meeting, Hanover 1998, 1–2 October. Hanover: Uni-Verlag Witte, pp 57–74
19. Kreft I, de Leeuw J (1998) Introducing multilevel modeling. London: SAGE
20. Dolan P (1997) Modeling valuations for EuroQol health states. Med Care 35:1095–1108
21. Busschbach JJV, McDonell J, Essink-Bot M-L, van Hout BA (1999) Estimating parametric relationships between health description and health valuation with an application to the EuroQol EQ-5D. J Health Econ 18:551–571
22. Macran S, Kind P (2001)"Death„ and the valuation of health-related quality of life. Med Care 39:217–227