Agota Szende
Mark Oppe
Nancy Devlin
Editors

EuroQol Group Monographs

EQ-5D value sets

inventory, comparative review and user guide

Springer

EQ-5D

EQ-5D Value Sets:
Inventory, Comparative Review and User Guide

EQ-5D Value Sets: Inventory, Comparative Review and User Guide

Edited by

AGOTA SZENDE
Covance Market Access Services, Leeds, U.K.

MARK OPPE
Institute for Medical Technology Assessment,
Erasmus MC, Rotterdam, The Netherlands

and

NANCY DEVLIN
City University, London, U.K.

On behalf of:
The EuroQol Group's Task Force on Value Sets

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-5510-2 (HB)
ISBN-13 978-1-4020-5510-2 (HB)
ISBN-10 1-4020-5511-0 (e-book)
ISBN-13 978-1-4020-5511-9 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Printed on acid-free paper

For Alan Williams,
whose intellect and wisdom continues to inspire us,
and to whom each of the authors are deeply indebted.

EuroQol Group

- The EuroQol Group is a network of international multidisciplinary researchers committed to the measurement of health-related quality of life. The EuroQol Group originally consisted of researchers from Europe, but nowadays includes members from North America, Asia, Africa, and Australasia. The Group is responsible for the development of EQ-5D, a preference-based measure of health-related quality of life.

- The EQ-5D self-report questionnaire consists of the EQ-5D descriptive system that measures health-related quality of life on 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) and the EQ VAS – a 20 cm vertical visual analogue scale that generates a self-rating of health-related quality of life. EQ-5D is widely used in clinical trials, observational studies, and other health surveys.

- The EuroQol Group has two primary research interests. One focuses on empirical work using EQ-5D, and the other focuses on methodological work to develop EQ-5D.

- The EuroQol Group is a "living" organization that, through its members, continuously conducts research using EQ-5D. Research areas include valuation and population studies, experimenting with the EQ-5D descriptive system, computerized applications, interpretation of EQ-5D ratings, and social inequalities in health status measurement.

- The EuroQol Group's website (www.euroqol.org) contains information about the EuroQol Group, membership and research activities, details of EQ-5D development and current status.

CONTENTS

LIST OF CONTRIBUTORS

Frank de Charro
Erasmus University Rotterdam, The Netherlands

Nancy Devlin
City University, London, United Kingdom

Mark Oppe
Institute for Medical Technology Assessment, Erasmus MC, Rotterdam, The Netherlands

David Parkin
City University, London, United Kingdom

Rosalind Rabin
EuroQol Executive Office, Rotterdam, The Netherlands

Agota Szende
Covance Health Economics and Outcomes Services, Leeds, United Kingdom

Chapter 1

Introduction

Rosalind Rabin, Frank de Charro, Agota Szende

1.1 Purpose of this booklet

Governments and healthcare funders worldwide are making increasing use of economic evaluation to inform priority setting in health care. For various reasons, cost benefit analysis is usually rejected in favour of cost-effectiveness or cost-utility analyses, often involving the estimation of the incremental cost per Quality Adjusted Life Year (QALY) gained (Drummond et al, 2005). The estimation of QALYs gained requires valuations for all relevant health states on a scale anchored at 1 = Full health and 0 = Dead.

The EQ-5D is widely used in this context and a number of value sets are available for all the health states generated by the EQ-5D descriptive system. These can be readily applied to health outcomes measured as EQ-5D profiles. EQ-5D has become one of the valuation approaches recommended by several reimbursement authorities and academic bodies in European countries (e.g. The Netherlands, Norway, Italy, Hungary, Poland, Portugal, UK), North America (e.g. Canada), and elsewhere (e.g. New Zealand).

The EuroQol Group frequently receives requests for advice regarding EQ-5D valuation data. Those seeking to apply EQ-5D valuations in economic evaluation want to know about the availability of EQ-5D value sets and how they can obtain them. They also seek specific guidance about *which* of the available value sets they should use for their purposes.

The Group has recently published a book devoted to the EQ-net project. Funded by the European Union under the Biomed II scheme, the project aimed to further develop the EQ-5D in the key areas of valuation, application, translation, and communication.

A considerable portion of the book focused on harmonising and integrating the results of the various EuroQol European valuation projects during the last 10 years (Brooks et al, 2003).

The purpose of the current booklet is to build on the work of the EQ-net project by drawing together and presenting, in a manner accessible to potential users of EQ-5D, the results of international research efforts to value EQ-5D states. Unlike the EQ-net book, this booklet focuses on both European and non-European valuation work. More specifically, our aims are: (1) to provide a comprehensive inventory of existing EQ-5D value sets, together with documentation and a comparative commentary on how these have been elicited and estimated; (2) to provide a basis for 'best practice' guidance to potential users of these value sets regarding the choice of value set for their particular application.

The EuoQol Group first met in 1987 with the aim of developing a standardized, non-disease-specific instrument for describing and valuing health-related quality of life (EuroQol Group, 1990). Originally a 6-dimensional questionnaire, the current 5-dimensional format was developed in 1991. Nowadays the EQ-5D self-report questionnaire (commonly known as EQ-5D), consists of 2 pages comprising the EQ-5D descriptive system (page 2) and a visual analogue scale - the EQ VAS (page 3). The descriptive system comprises 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has 3 levels: no problems, some problems, extreme problems. The respondent is asked to indicate his/her health state by marking the box against the most appropriate statement in each of the 5 dimensions. This decision results in a 1-digit number expressing the level selected for that dimension. The digits for 5 dimensions can be combined in a 5-digit number describing the respondent's health state.

The EQ VAS records the respondent's self-rated health on a vertical, 20 cm visual analogue scale where the endpoints are labelled 'best imaginable health state' and 'worst imaginable health state'. This information can be used as a quantitative measure of health outcome as judged by the individual respondents.

The EQ-5D self-report questionnaire is designed for self-completion by respondents and is ideally suited for use in postal surveys, in clinics, and in face-to-face interviews. It is cognitively undemanding, taking only a few minutes to complete. Instructions to respondents are included in the questionnaire. Applicable to a wide range of health conditions and treatments, the EQ-5D provides a descriptive profile and a single index value for health status that can be used in the clinical and economic evaluation of health care as well as in population health surveys (Figure 1).

A distinction should be made between the EQ-5D self-report questionnaire for measuring health outcome and the EQ-5D valuation questionnaire (designed to collect valuations for health states defined by the EQ-5D descriptive system). The latter contains a technique for valuing health states using the EQ-5D VAS rating scale - a vertical 20 cm visual analogue scale with the end points labelled 'best imaginable health state' at the top and 'worst imaginable health state' at the bottom having numeric values of 100 and 0 respectively. The EuroQol Group decided on a subset of health states for valuation in 1990 and values have been elicited from general population samples and from patients in several countries. For various reasons however, valuation studies have not always adhered to the standard approach – for example many valuation studies carried out by EuroQol Group members also used the Time Trade-Off (TTO) technique.

Figure 1: The EQ-5D self-report questionnaire

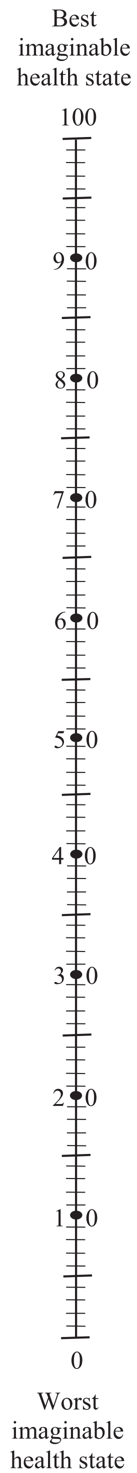
By placing a tick in one box in each group below, please indicate which statements best describe your own health state today.

Mobility	
I have no problems in walking about	❑
I have some problems in walking about	❑
I am confined to bed	❑
Self-Care	
I have no problems with self-care	❑
I have some problems washing or dressing myself	❑
I am unable to wash or dress myself	❑
Usual Activities (*e.g. work, study, housework, family or leisure activities)*	
I have no problems with performing my usual activities	❑
I have some problems with performing my usual activities	❑
I am unable to perform my usual activities	❑
Pain/Discomfort	
I have no pain or discomfort	❑
I have moderate pain or discomfort	❑
I have extreme pain or discomfort	❑
Anxiety/Depression	
I am not anxious or depressed	❑
I am moderately anxious or depressed	❑
I am extremely anxious or depressed	❑

To help people say how good or bad a health state is, we have drawn a scale (rather like a thermometer) on which the best state you can imagine is marked 100 and the worst state you can imagine is marked 0.

We would like you to indicate on this scale how good or bad your own health is today, in your opinion. Please do this by drawing a line from the box below to whichever point on the scale indicates how good or bad your health state is today.

Your own health state today



1.2 Valuation surveys

The "value sets" included in this guide are those where valuations have been produced for all 243 EQ-5D states. We have excluded from this inventory, a large number of EQ-5D valuation studies that do not provide valuations for *all* possible EQ-5D health states. For some countries more than one value set was derived. For these countries we have included the most recent in this booklet. Therefore, a few of the older value sets have been omitted. Value sets (previously referred to as "tariffs") may be produced either by valuing all states in a particular study ('saturation' studies) or, much more commonly, by valuing a selection of EQ-5D states and, using econometric techniques, extrapolating over the full set of states. As well as documenting the value sets, we offer a brief review of each survey. The valuation surveys included in this booklet are summarized in Table 1.

1.3 The structure of the booklet

The rest of this booklet consists of 3 further chapters. Chapters 2 and 3 provide comparative reviews of the Time Trade-Off and Visual Analogue Scale value sets. Chapter 4 provides guidance to users of EQ-5D value sets. Annexes 1 and 2 provide inventories of Time Trade-Off and Visual Analogue Scale valuation surveys. A fuller account of this work was presented at the EuroQol Group's 21st annual meeting in Chicago, USA, in 2004 and can be accessed via the EuroQol Group's website. Researchers planning to conduct new valuation surveys with EQ-5D should contact the EuroQol Executive Office (userinformationservice@euroqol.org).

For researchers unfamiliar with the advice of their national regulatory bodies, the EuroQol Group's website provides guidelines regarding national recommendations on valuing health states (www.euroqol.org). These guidelines are routinely updated. The content of this information however is intended to be summary information only as the information is liable to change. All users should contact their local regulatory bodies for up-to-date information.

Table 1: Summary of EQ-5D valuation survey characteristics

Country	N	Valuation method	Reference
Belgium	722	VAS	Cleemput, 2003
Denmark	1686	VAS	Wittrup-Jensen et al, 2002
Denmark	1332	TTO	Wittrup-Jensen et al, 2002
Europe	8709	VAS	Greiner et al, 2003
Finland	1634	VAS	Ohinmaa et al, 1996
Germany	339	VAS	Claes et al, 1999
Germany	339	TTO	Greiner et al, 2005
Japan	621	TTO	Tsuchiya et al, 2002
Netherlands	309	TTO	Lamers et al, accepted for publication
New Zealand	1360	VAS	Devlin et al, 2000
Slovenia	733	VAS	Prevolnik-Rupel et al, 2001
Spain	300	VAS	Badia et al, 1998
Spain	1000	TTO	Badia et al, 2001
UK	3395	VAS	MVH Group, 1995
UK	3395	TTO	MVH Group, 1995; Dolan, 1997
USA	4048	TTO	Shaw et al, 2003
Zimbabwe	2440	TTO	Jelsma et al, 2000

Chapter 2

Comparative review of Time Trade-Off value sets

Agota Szende, Mark Oppe, Frank de Charro

2.1 Introduction

The Time Trade-Off (TTO) method has played an important role in generating value sets for the EQ-5D. As one of the most widely accepted preference elicitation methods for health states (Torrance, 1986) and the method of choice in the first large-scale EQ-5D valuation study (Dolan, 1997), the TTO has become the preferred approach among EuroQol Group researchers who wanted to use a choice-based elicitation technique to value EQ-5D health states in their own countries.

The first TTO study (i.e. UK-MVH) (Dolan, 1997), carefully reviewed the potential advantages and disadvantages of the then existing alternative preference elicitation techniques. The TTO method was selected mainly based on evidence on the validity and reliability of this approach over the Standard Gamble (SG) technique. The TTO-based value set gained from this first UK study became widely used in validation and other clinical studies in many different therapeutic areas and patient populations. The accumulated scientific evidence using this TTO value set helped the acceptance of the EQ-5D by various reimbursement authorities, particularly by NICE in the UK. In response to the demand for value sets that reflect the views of populations in different countries, researchers started to perform additional valuation studies. This section describes these studies and reports the findings in a standard format to facilitate comparisons (Annex 1).

2.2 Survey characteristics

Eight TTO studies have been conducted to elicit preference weightings for the EQ-5D, including the first such study organised in the UK in 1993. Other European countries included Spain, Germany, Denmark, and the Netherlands. From elsewhere, valuation studies have been conducted in Japan, Zimbabwe, and more recently the USA.

Sampling procedures

All studies surveyed representative samples of adults from the general populations of the respective countries, although it was only the Danish, the German, and the first UK study in which the sampling covered the population of the whole country. The Dutch study covered the Rijnmond area, the Japanese sample covered 3 prefectures of Japan, the Spanish study covered one health care district of Barcelona, and the Zimbabwean sample was drawn from a list of residents of one suburb of Harare. All studies nevertheless weighted the final sample for the demographic structure of the country's whole population. Over 50% of contacted people were willing to participate in all surveys.

Sample size

The sample size exceeded n=500 in most surveys. Exceptions were the German study (n=339) and the Dutch study (n=309). The sample sizes used were proved to be sufficient to achieve statistically significant results, with the exception of the German study in which 4 of the estimating coefficients of the final utility algorithm were not statistically significant. In this case the value set was estimated without these coefficients.

2.3 Data collection methods, materials, and procedures

Health states valued

All studies had in common that they selected a number of health states with a mixture of severity levels for direct evaluation. The preference ratings for these health states served as the basis for eliciting an evaluation algorithm to estimate utilities for all possible health states. Studies varied though in the total number of health states selected for direct valuation and the number of health states each respondent was asked to value. Two main approaches were used. The first approach was used in the first UK investigation and then, without or with some variations, in Denmark, Germany, Spain, and Zimbabwe. This approach is based on the selection of a total number of 43 (or less) health states from which a smaller number of health states (typically 13) are valued in sub-samples of respondents. In the second approach, used by the Dutch study and the Japanese study, the same 17 health states were valued by each respondent in the survey. All these health states were anchored on the Full health (11111) and the Dead states, and a utility function was estimated based on mean ratings for the 17 states.

Data collection procedures

All valuation tasks were conducted during face-to-face interviews by trained interviewers. The Danish study used a computer-assisted method. The majority of studies administered the EQ-5D health questionnaire prior to the TTO tasks in order to familiarise participants with health state descriptions and also to describe study participants. Normally, a ranking exercise was used to elicit participants' preferences between health states. This was followed by the actual TTO exercises.

TTO props

In all studies, a specifically designed double-sided board was used for the valuation of each health state. One side was relevant for health states better than Dead, the other for those worse than Dead. Respondents were led by a process of 'bracketing' to find their point of indifference between alternatives. Respondents were asked to select a length of time in Full health that they regarded as equivalent to **10 years** in the target state. In the case of states regarded as worse than Dead, the choice was between dying immediately vs. spending a length of time (x) in the target state followed by **(10-x) years** in the 11111 state. In some cases, e.g. the German study, respondents were allowed to trade time in weeks instead of years.

Estimation of direct utility scores

States regarded better than Dead were anchored on the Full health and Dead scale: X/10. States regarded worse than Dead were calculated as X/10-1, so scores were bounded by -1. In the US model the scores for the states worse than Dead were bounded by -1 using a linear transformation: X/39.

2.4 Valuation models

Exclusion criteria for valuation data

Studies showed some variations in how they handled exclusion criteria for valuation data. Approaches to exclusion involved the following strategies: exclusion of 'extreme' values which deviated substantially from median scores; if less than 3 states were valued; if all states were valued the same; if all states were valued worse than Dead; if ratings reflected serious logical inconsistencies.

Model characteristics

In general, all valuation models are built on the assumption that health utilities are additive. The German study experimented with a multiplicative model together with the additive model.

For practical reasons, 'disutilities' (1-utility) were used in the model estimations, except for Zimbabwe. In all studies, 2 variables were used in each dimension in order to allow for different utility increments between levels 1 and 2 as compared to the increment between levels 2 and 3. The intercept (also called constant or sometimes N2 [1]) included in the models was interpreted as 'extra' disutility associated with any deviation from Full health. An additional interaction term (N3) was used in order to allow for measuring the 'extra' disutility when reporting severe (level 3) problems on at least 1 EQ-5D dimension. The Danish, the Japanese, and the Zimbabwean studies

---

[1] Some variations were found across study reports with respect to what people called the N2 term. Some studies used N2 for an additional interaction variable coded as 1 if any dimension was at level 2; 0 otherwise. In other studies the N2 variable was coded as 1 if any dimension was at level 2 *or* 3; 0 otherwise. Since none of the final value set models included an extra interaction term for problems on levels 2 or 3, we have simply used N2 as the constant term in this inventory.

did not use this latter term in the model. The US study used a model with 3 interaction terms that replaced the intercept and the N3 term.

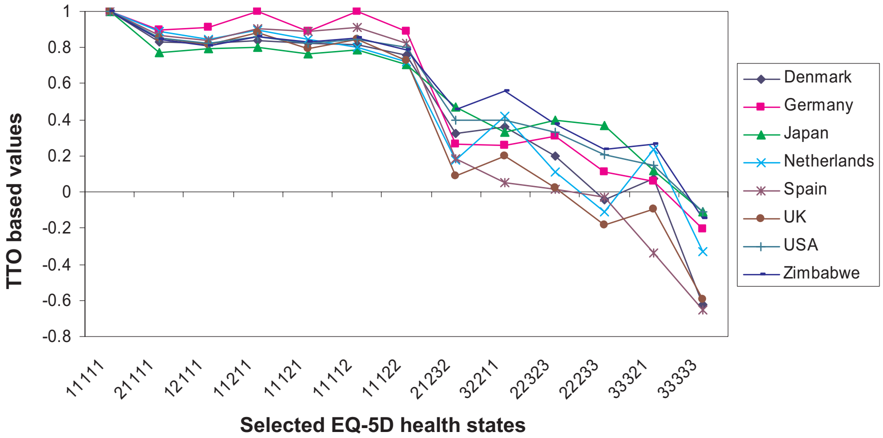
In the majority of studies a generalised least-squares regression technique was used based on individual-level (rather than aggregate-level) data. A random effects specification was normally applied to address the problem that the scores for health states generated by an individual were related.

2.5 Value sets

The estimated valuation algorithms and results on their goodness of fit for each study are presented in Tables 2A-2B. Due to some differences in the models, the direct comparison of estimating coefficients across surveys is somewhat limited. Nevertheless, some similarities across country scores can be observed. It is striking that almost across all countries and EQ-5D dimensions, the incremental disutility between problems reported on levels 2 and 3 was larger than the incremental disutility between levels 1 and 2. This finding was irrespective of whether the models included the N3 term or not. This observation reflects that people generally value level 3 problems on the EQ-5D proportionately worse than those on level 2. In comparing dimensions (based on level 3 coefficients and assuming no other problems), people in all countries either valued problems with mobility the worst (Denmark, Germany, Japan, Spain, USA) or problems with pain/discomfort (UK, Zimbabwe). Problems with usual activities were valued as least bad in all countries without exception.

In cross-country comparison, the differences in estimates for constant, interaction, and EQ-5D variables in the estimating formulas lead to significant differences in values associated with individual health states. For example, the utility scores for the health state '21232' represented a fourfold difference between the lowest value of 0.088 (UK-MVH) and the highest value of 0.472 (Japan). Other surveys reflected a range of values in between: The Netherlands 0.174; Spain 0.185; Germany 0.262; Denmark 0.321; USA 0.397; and Zimbabwe 0.453. Differences in utility scores between levels of problems in individual EQ-5D dimensions (taking everything else as constant) were generally below 0.1. Figure 2 summarizes valuations of selected EQ-5D health states across surveys.

Figure 2: Mean values of selected EQ-5D health states across TTO-based surveys



It would need further evidence to comment on how meaningful these differences are when considered in the clinical evaluation of alternative treatments, i.e. when estimating incremental improvements in health status over time when alternative medications are used. Based on the relatively large differences found between value sets, one cannot exclude the possibility that using different value sets on the same clinical trial data can alter the final conclusions of a study in certain cases.

Goodness of fit analyses proved that direct utilities from the surveys fitted model-derived utilities relatively well in all studies, with R^2 results of over 0.4. Goodness of fit was not only measured using R^2. Measures using the differences between predicted values and observed values such as mean absolute error and correlations were also measured for goodness of fit. No clear relationship was found between goodness of fit and model type with respect to whether the N3 term was included in the estimating model.

Table 2A: Coefficients for the estimation of the EQ-5D index values based on TTO valuation studies

	Germany	Netherlands	Spain	UK - MVH
Full health (11111)	1	1	1	1
At least one 2 or 3 (Constant)	- 0.001	- 0.071	- 0.024	-0.081
At least one 3 (N3)	- 0.323	- 0.234	- 0.291	-0.269
Mobility = 2	- 0.099	- 0.036	- 0.106	-0.069
Mobility = 3	- 0.327	- 0.161	- 0.430	-0.314
Self care = 2	- 0.087	- 0.082	- 0.134	-0.104
Self care = 3	- 0.174	- 0.152	- 0.309	-0.214
Usual activities = 2	–	- 0.032	- 0.071	-0.036
Usual activities = 3	–	- 0.057	- 0.195	-0.094
Pain/discomfort = 2	- 0.112	- 0.086	- 0.089	-0.123
Pain/discomfort = 3	- 0.315	- 0.329	- 0.261	-0.386
Anxiety/depression = 2	–	- 0.124	- 0.062	-0.071
Anxiety/depression = 3	- 0.065	- 0.325	- 0.144	-0.236
R^2	0.43	0.38	0.60	0.46

Table 2B: Coefficients for the estimation of the EQ-5D index values based on TTO valuation studies

	Denmark	Japan	USA*	Zimbabwe
Full health (11111)	1	1	1	1
At least one 2 or 3 (Constant)	- 0.114	- 0.152	–	- 0.100
At least one 3 (N3)	–	–	–	–
Mobility = 2	- 0.053	- 0.075	- 0.146	- 0.056
Mobility = 3	- 0.411	- 0.418	- 0.558	-0.204
Self care = 2	- 0.063	- 0.054	- 0.175	-0.092
Self care = 3	- 0.192	- 0.102	- 0.471	-0.231
Usual activities = 2	- 0.048	- 0.044	- 0.140	-0.043
Usual activities = 3	- 0.144	- 0.133	- 0.374	-0.135
Pain/discomfort = 2	- 0.062	- 0.080	- 0.173	-0.067
Pain/discomfort = 3	- 0.396	- 0.194	- 0.537	-0.302
Anxiety/depression = 2	- 0.068	- 0.063	- 0.156	-0.046
Anxiety/depression = 3	- 0.367	- 0.112	- 0.450	-0.173
D1	–	–	+ 0.140	–
I2-square	–	–	- 0.011	–
I3	–	–	+ 0.122	–
I3-square	–	–	+ 0.015	–
R^2	0.66	0.40	N/A	0.51

* D1: ordinal variable that represented the number of movements away from Full health beyond the first (ranging from 0 to 4); I3: ordinal variable that represented the number of dimensions at level 3 beyond the first. The square of the I3 term to allow for non-linearity in its association with the dependent variable. The square of I2, an ordinal variable that represented the number of dimensions at level 2 beyond the first.

Chapter 3

Comparative review of Visual Analogue Scale value sets

Mark Oppe, Agota Szende, Frank de Charro

3.1 Introduction

In addition to the Time Trade-Off approach, the Visual Analogue Scale (VAS) has become the other widely used method to elicit preferences for the EQ-5D. Its less demanding nature, easy administration, and favourable evidence regarding its psychometric characteristics made the VAS attractive to include in large postal surveys in more than 10 countries.

VAS valuation studies have been carried out for numerous reasons in the past and not all of them resulted in a full value set. These included comparisons between different versions of the EQ-5D and research on handling of missing data or inconsistencies. There were also examples of experiments in which researchers tried to elicit preferences for EQ-5D health states based on the EQ-5D health questionnaire by linking patient responses on the EQ VAS with reported problems on the 5 dimensions of the EQ-5D. Also, studies that omitted valuing the state Dead state could not be regarded as full valuation studies.

This inventory includes those studies that carried out a full valuation procedure and were designed to reflect the opinion of the general population. It should be noted however, that data from certain European studies that did not follow a full VAS valuation approach, such as Sweden, Germany, and the Netherlands, were still included in the European Value Set.

3.2 Survey characteristics

Nine VAS studies have been carried out to elicit preference weighting for the EQ-5D using a full valuation approach (Annex 2). The first study was carried out in Finland in 1992. Other European countries included Belgium, Denmark, Germany, Slovenia, Spain and the UK. Non-European countries included New Zealand. A European VAS value set was constructed using data from 11 valuation studies in 6 countries: Finland (1), Germany (3), The Netherlands (1), Spain (3), Sweden (1) and the UK (2). Although not all the studies included were representative of the country in which they were carried out and although data from a number of other European countries was not available, there is enough data from different European regions to make the European VAS dataset moderately representative for Europe.

The 11 studies were carried out in the period from January 1991 - March 1998. The survey settings varied between the studies so both postal surveys and interview-based surveys were included in the European dataset. The pooled data set consisted of valuations from 8709 respondents.

Sampling procedures

Most studies used a representative sample of the adult population of the entire country. The exceptions were the Belgian and Spanish studies. The Belgian study included a sample of Flemish respondents and the Spanish study included a sample of adults from one primary care district in Catalonia. The German, Spanish and UK studies were interview based. The other studies used postal surveys. In the postal surveys between 24% and 65% of contacted people were willing to participate. Response rates were higher in those countries that sent out 2 reminder letters (Finland 64.5%), made telephone contact (Denmark 53%), or sent out the reminder letter with a duplicate questionnaire (New Zealand 50%).

Sample size

The sample size exceeded n=500 in most surveys. The sample sizes used were proved to be sufficient to achieve statistically significant results in at least one of the models tested in each study.

3.3 Data collection methods, materials, and procedures

Health states valued

The standard EQ-5D valuation questionnaire includes 16 health states, also referred to as the common core, plus Dead. These health states represent a mixture of severity levels in the 5 dimensions and are included in all full valuation studies:

The common core states

11111 a	22233
11111 b	22323
11112	32211
11121	33321
11122	33333 a
11211	33333 b
12111	Unconscious
21111	Dead a
21232	Dead b

As can be seen, the states 11111, 33333 and Dead are included twice (a and b). This is because these states are presented once on both pages so that the remaining 6 states on the page can be scaled according to these anchor points. There were 2 studies (Denmark and Slovenia) that asked all respondents to value the common core states. The rest of the studies valued more states (between 24 and 46) using different combinations of states in sub-samples of study participants. The directly elicited ratings for these health states served as the basis for developing an algorithm to value all possible EQ-5D health states.

Data collection procedures

The majority of studies administered the EQ-5D health questionnaire prior to the VAS valuation exercises. The 3 interview based studies (i.e. the German, Spanish and UK studies) administered the EQ-5D health questionnaire plus a ranking exercise before participants' preferences between health states were elicited using the VAS method. This was then followed by a TTO rating task.

VAS props

Most studies used the standard EQ-5D Visual Analogue Scale. This is a 20 cm (thermometer like) vertical rating scale. The end-points of the scale are 'best imaginable health' (= 100) and 'worst imaginable health' (= 0). Numbers are shown on the scale for every 10, large tick marks for every 5 and small tick marks for every point on the scale. The valuation part of the EQ-5D typically has 2 pages. Each page has 2 columns with 4 health states and the VAS shown between the two columns. Respondents are asked to rate the 8 health states on each page by drawing lines from the health state to a point on the VAS reflecting their value for the health state. After valuing the 16 states respondents are asked to indicate the value of Dead on the VAS on both pages.

The standard EQ-5D valuation questionnaire asks the respondents to value the health states under the assumption that the health states last for 1 year. What happens after that year is unknown. The German, Spanish and UK studies used a 10-year period rather than a 1-year period.

Calculation of direct utility scores

Because the use of health state valuations in Cost-Utility Analysis requires valuations to be anchored at 0 for Dead and 1 for Full health, the data collected using the VAS valuation task has to be rescaled from 'best imaginable health' = 100 and 'worst imaginable health' = 0 to a scale where 11111 = 1 and Dead = 0. There are a number of ways to do this. The first, and most theoretically sound, is to rescale the values using the values for 11111 and Dead at respondent level. That is, a respondent's value for a particular state is rescaled using the values for 11111 and Dead from the same respondent (or indeed from the same page of the questionnaire). The second is to rescale the values based on aggregate data, using either mean or median values for 11111 and Dead from the whole sample. This rescaling on aggregate data is sometimes preferred to rescaling on individual data, because it does not require the values for 11111 and Dead from every respondent. Given persistent problems with missing valuations for Dead (it is not uncommon for up to one third of the responses to VAS surveys to have missing values for Dead) this approach avoids these valuations being discarded as unusable. For example in the common European Value Set study rescaling was performed on aggregate data so that valuation data from studies that did not include Dead could also be used. When to rescale, either before modeling the VAS valuation data or after, also has an impact on the model obtained and the exclusions that are necessary in order to obtain a good value set.

Most studies were rescaled at the individual level data, with the exception of Europe and Finland who rescaled at the aggregate level. Spain experimented with both approaches. The equation to rescale raw VAS valuation data is the following:

$$X_{rescaled} = (X_{raw} - Dead_{raw}) / (11111_{raw} - Dead_{raw})$$

Depending on the timing and level (i.e. individual/aggregate) the 'raw' values can be means, medians or individual values.

3.4 Valuation models

Exclusion criteria for valuation data

Studies showed some variations in how they handled exclusion criteria for valuation data. Approaches to exclusion involved the following strategies: exclusion of ‘extreme’ values which deviated substantially from median scores; if less than 3 states were valued; if all states were valued the same; if all states were valued worse than Dead; if ratings reflected serious logical inconsistencies.

Model characteristics

Almost all the VAS valuation models were built on the assumption that health utilities are additive. Several of the studies, such as the Belgian, Danish, and New Zealand, used ‘disutilities’ (1-utility) in the model estimations. The German VAS model was a multiplicative model.

The Belgian, Danish, New Zealand, Slovenian and Spanish studies used 1 variable for each dimension of the EQ-5D. This means that the difference in utilities between levels 1 and 2 was the same as that between levels 2 and 3. The European, Finnish, German and UK models used separate variables for level 2 and level 3 problems, allowing for different utility increments between levels. The intercept included in the models was interpreted as ‘extra’ disutility associated with any deviation from Full health. An interaction term (N3) was used in the Belgian, European, New Zealand, Spanish and UK models in order to allow for measuring ‘extra’ disutility when reporting severe (level 3) problems on at least 1 dimension.

In the majority of studies a generalised least-squares regression technique was used. A random effects specification was normally applied to address the problem that the scores for health states generated by an individual were related. The Finnish study experimented with the use of logistic transformations of data in order to normalise the observed preference values.

Selection of final model

The majority of studies developed different models and tested them to find the most appropriate to recommend for future use. Selecting the model that represents the data best is based on any or all of a number of criteria. These criteria include parsimony, statistical significance of the variables in the model, producing a logically consistent model and goodness of fit of the model. This inventory included those final models that were recommended by the authors of the original study. The exception was the Slovenian study where there was deviation from the original recommendation based on the above criteria. The Finnish study recommended the use of two basically equally well performing models. Only 1 of these 2 models was included in this inventory as the final value set, based on a slight difference in the logical consistency criteria.

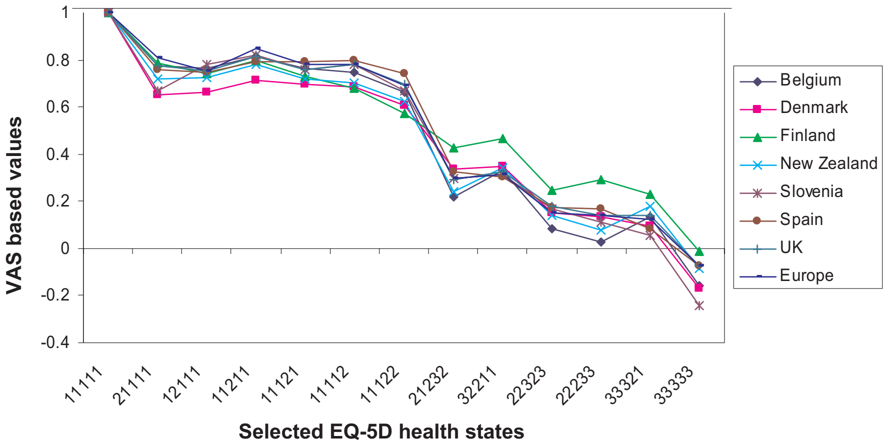
3.5 Value sets

The estimated valuation algorithms and results on their goodness of fit for each study are presented in Tables 3A-3B. It has to be noted that due to some differences in the models, the direct comparison of estimating coefficients across surveys is somewhat limited. Without sufficient comparative information on model types, it is difficult to determine if differences in value sets reflect differences in the actual views of the populations or if they are a consequence of the different analytic procedures.

In 5 of the 9 countries (Denmark, Finland, Germany, Slovenia and the UK), problems with mobility were regarded as least desirable among the 5 dimensions (when considering level 3 problems and no other problems exist). This was also the case in the European value set. Belgium and New Zealand regarded anxiety/depression as the worst dimension, while this was the dimension Spanish people worried about the least. Not unlike the TTO-based value sets, the majority of countries regarded problems with usual activities the least bad.

In cross-country comparisons, given the differences in estimates for constant, interaction, and EQ-5D variables in the models across countries, the differences in values associated with individual health states can be significant. For example, the utility scores for the health state ‘21232’ represented a twofold difference between the lowest value of 0.216 (Belgium) and the highest value of 0.424 (Finland). Other countries reflected a range of values in between: New Zealand 0.239; Germany 0.294; UK 0.294; Slovenia 0.297; the European Value Set 0.298; Spain 0.323; and Denmark 0.338). Differences in utility scores between levels of problems in individual EQ-5D dimensions (taking everything else as constant) were generally below 0.1 across VAS value sets. Figure 3 illustrates mean values for selected EQ-5D health states across various VAS-based surveys.

Figure 3: Mean values of selected EQ-5D health states in VAS-based surveys



Goodness of fit analyses proved that direct utilities from the surveys fitted model-derived utilities relatively well in most studies, with R^2 results of over 0.6. No clear relationship was found between goodness of fit and model type.

In countries where both TTO and VAS valuations were available, the value sets reflected lower scores for the VAS-based approach for mild health states, and higher scores for the VAS-based approach for severe health states (see Figure 4).

Figure 4: Comparison of TTO and VAS-based value sets in selected countries

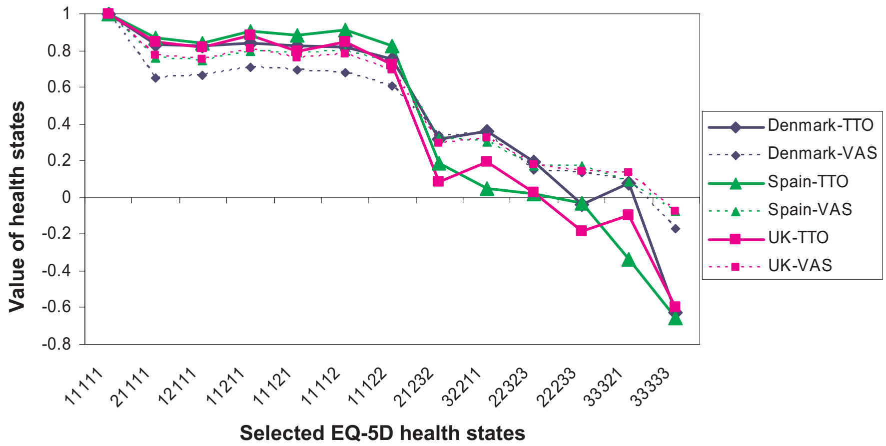


Table 3A: Coefficients for the estimation of the EQ-5D index values based on VAS valuation studies

	Belgium	**Europe**[†]	**New Zealand**	**Spain**	**UK**
Full health (11111)	1	1	1	1	1
At least one 2 or 3 (Constant)	-0.152	-0.128	-0.204	-0.150	-0.155
At least one 3 (N3)	-0.256	-0.229	-0.217	-0.212	-0.215
Mobility = 2	-0.074	-0.066	-0.075	-0.090	-0.071
Mobility = 3	-0.148	-0.183	-0.151	-0.179	-0.182
Self care = 2	-0.083	-0.117	-0.071	-0.101	-0.093
Self care = 3	-0.166	-0.156	-0.143	-0.202	-0.145
Usual activities = 2	-0.031	-0.026	-0.014	-0.055	-0.031
Usual activities = 3	-0.062	-0.086	-0.027	-0.110	-0.081
Pain/discomfort = 2	-0.084	-0.093	-0.080	-0.060	-0.084
Pain/discomfort = 3	-0.168	-0.164	-0.160	-0.119	-0.171
Anxiety/depression = 2	-0.103	-0.089	-0.092	-0.051	-0.063
Anxiety/depression = 3	-0.206	-0.129	-0.184	-0.102	-0.124
R^2	–	0.75	0.70	0.97	0.47

[†] The values of the European VAS model have been rescaled with the mean value of Dead

Table 3B: Coefficients for the estimation of the EQ-5D index values based on VAS valuation studies

	Denmark	Finland	Germany‡	Slovenia
Full health (11111)	1	1	1	1
At least one 2 or 3 (Constant)	-0.225	-0.158	0.926	-0.128
At least one 3 (N3)	–	–	–	–
Mobility = 2	-0.126	-0.058	0.945	-0.206
Mobility = 3	-0.252	-0.230	0.393	-0.412
Self care = 2	-0.112	-0.098	0.808	-0.093
Self care = 3	-0.224	-0.143	0.470	-0.186
Usual activities = 2	-0.064	-0.047	0.880	-0.054
Usual activities = 3	-0.128	-0.131	0.554	-0.108
Pain/discomfort = 2	-0.078	-0.111	0.975	-0.111
Pain/discomfort = 3	-0.156	-0.153	0.467	-0.222
Anxiety/depression = 2	-0.091	-0.160	0.817	-0.093
Anxiety/depression = 3	-0.182	-0.196	0.468	-0.186
R^2	0.82	0.74	0.72	0.65

‡ The German VAS model is a multiplicative model. See annex 2 for more details on the German model.

Chapter 4

Guidance to users of EQ-5D value sets

Nancy Devlin, David Parkin

4.1 Introduction

One of the most common questions asked of the EuroQol Group by those wishing to use EQ-5D value sets in economic evaluation is '*which* value set should I use?' The aim of this chapter is to provide advice on this question, and to guide potential users through the issues that are pertinent to choosing which of the value sets described earlier to use.

More technically, this chapter deals with the conversion of EQ-5D profiles to a single index by means of weights which are attached to the profiles' levels and dimensions. Although there are many possible uses for such indexes, the chapter mainly deals with their use for economic evaluation and the consequent need for weights – called value sets – which have the required properties in that context. However, we start by describing some general principles for selecting weights and possible uses of the resulting indexes.

An EQ-5D profile is a set of observations about a person according to the EQ-5D descriptive system – the level that they are assigned to in each of the 5 dimensions. An index is calculated from a profile by applying a formula that essentially attaches weights to each of these levels in each dimension. An important consideration in choosing a set of weights is that the resulting index should have properties that its intended use requires. Possible uses of an index include:

- Summarising EQ-5D profiles for statistical analysis.
- Describing the health of a population.
- Comparing population health between different regions, countries or other populations; or over time.

- Describing severity of illness amongst patients.
- Assessing population or patient priorities for treatment.
- Assessing the impact of health care interventions (such as pharmaceuticals) or policies on health.
- Assessing the cost-effectiveness of health care interventions or policies.

The first and last of these are discussed in more detail, with some general observations about other uses.

Summarising a profile for statistical analysis

There are important advantages in being able to summarise and represent a health profile by a single number – for example, it simplifies statistical analysis. However, it should be borne in mind that there is no “neutral” set of weights that can be used for this purpose. All sets of weights explicitly or implicitly compare each level of each dimension with every other and attach relative importance to them. No set of weights is “objective”: they all embody judgements about both what is meant by importance and also the appropriate source of information for assessing relative importance. It is therefore not possible to offer generalised guidance about which set of weights should be used if the sole purpose is summarising profiles for descriptive or inferential statistical analysis. Users should consider the wider purpose for which the summary will be used. If there is no one purpose, rather just a desire to provide information, then it may be better *not* to use an index, but to report the EQ-5D profiles themselves in some detail. This may also be preferable because an index provides less detailed information than a profile. Further, in some cases where a single number is required to represent health (for example, the generation of ‘population norms’ as reported by Kind et al, 1999 for the UK) it may be more appropriate to focus on the EQ VAS data provided by the relevant patients or populations themselves (i.e. their rating of their ‘own health today’ on the VAS), rather than applying social value sets to their EQ-5D profiles.

Assessing cost-effectiveness

The requirements for an index for use in cost-effectiveness analysis are mainly that it should provide an unambiguous measure of effectiveness – essentially, that higher scores on the index represent a better state of health and that the same differences between scores have the same level of importance. However, it is arguable that there is a further requirement if the effectiveness index is to be based on economic principles, such as those embodied in "cost-utility analysis" – essentially that the weights represent "values". Just as costs represent the volume of resources used weighted by the *value* of those resources, effectiveness, in the context of economic evaluation, should represent the volume of health "output" weighted by the *value* of that output.

For cost-utility analyses, the set of values will be used to calculate Quality Adjusted Life Years (QALYs) and for this purpose it is also essential that they are anchored between 0, representing states as bad as being Dead, and 1, representing Full health. In practice, this is the principal use of the value sets presented in later chapters.

There is a further requirement that, although not essential for all cost-effectiveness analyses, is an important one for the value sets described in this booklet. It is that the value sets should represent "social" valuations. The reasons for this, and alternative approaches that might be taken, are discussed in section 4.2.

Other uses

While cost-effectiveness analysis is the principal purpose of the value sets in this booklet, there are other uses for them. In principle, social value sets are used in any context where a single number is required to summarise or represent the social value of an EQ-5D state. For example, Derrett et al (2003) applied EQ-5D valuations to patients' EQ-5D profiles as a means of creating a ranking of patients on elective surgery waiting lists in terms of the severity of their condition and their suggested priority for treatment. The choice of value set in this and other uses should be based on some general principles. Essentially, the weights used should be appropriate to the proposed use of the index; they will not be "neutral", but reflect a definition of importance and of who shall judge importance; the resulting index should be unambiguous; and if the index is to be interpreted as a measure of value, the weights must also represent values. The considerations outlined in 4.2 below may also be relevant to these uses of the value sets.

4.2 Which value set should I use in an economic evaluation?

Below we set out the principal considerations that should guide your choice of value set for use in cost-effectiveness and cost-utility studies. Although we begin with some pragmatic approaches to this choice, selecting a value set inevitably involves confronting some important and complex theoretical issues regarding the generation and use of health state values. These issues affect *all* attempts to measure and value health states – not just the EQ-5D instrument.

Which decision-making process will the economic evaluation inform? Do decision-makers stipulate the nature of the valuations to be used?

In many cases, economic evaluation is performed for the express purpose of providing evidence for a formal decision-making process. For example, decisions and recommendations made by the National Institute of Clinical Excellence (NICE) in the UK and the Pharmaceutical Management Agency (PHARMAC) in New Zealand are both examples of public sector bodies that routinely use economic evaluations to make decisions about health care services. Where an evaluation is being performed to inform such decisions, the first consideration for those wanting to select a value set is an entirely pragmatic one: does the relevant decision maker specify any requirements or preferences regarding which value set should be used?

For example, NICE (2004) states that values used in economic evidence submitted to it "should be based on public preferences elicited using a choice-based method" (p.25). All of the value sets reported in this inventory are elicited from the general public. "Choice-based methods" in practice refers to Standard Gamble (SG) and Time Trade-Off approaches, a requirement whose theoretical grounds are discussed below. There are currently no EQ-5D value sets generated using SG, so the implication of

NICE's stated requirement is that those submitting evidence to it should value EQ-5D states using a TTO-based value set. Further, NICE stipulates that, for its purposes, the 'public' whose preferences are relevant should be consistent with its remit, namely the population of England and Wales (NICE, 2001). This further suggests that the York Measurement and Valuation of Health (MVH) TTO value set (Dolan, 1997), which is the only set based on a representative sample of the UK population, is currently most likely to be viewed as appropriate.

The EuroQol website (www.euroqol.org) provides a summary of health care decision making bodies internationally, and their stated requirements regarding the valuation of health states. In the circumstances where an evaluation is being performed that might simultaneously be submitted to more than one decision-making body, the choice of value set should be guided by each set of requirements. For example, where economic evaluation is being performed alongside a multi-country clinical trial, the value set relevant to each country should be applied to the effectiveness data generated and reported to decision makers in each country.

Where the user of the economic evaluation is not identified or where the relevant decision maker does not have any stated requirements or preferences about the nature of the valuations used in evidence submitted to it, the decision about how to proceed is left to the analyst.

Broadly speaking, there are three main considerations: **relevance** to the decision-making context (do the values reflect the geographical and economic context in which resource allocation decisions are made? Whose values are considered relevant to decision-making?); **empirical characteristics** (are the methods used to elicit and model the value set well described and appropriate?) and, more controversially, **theoretical properties** (is the method used to calculate the values defensible on theoretical grounds?).

Relevance

A general principle is that the value set used should be consistent with the decision-making context. An important consideration is whether the value set is appropriate for the population whose health status is being measured, which has two considerations. Whose values should be used? What is the appropriate source for such values?

The question of whose values should be used has been widely debated and there are a number of possible answers to that. Here, we will mainly consider the arguments for a "social" value set, which is meant to represent the values that the general public holds. In essence, these "social" valuations are generated from members of the general public being asked to imagine states that may be hypothetical to them, and to value them from the perspective of being in those states. An alternative is to use patients' values. One argument for doing this is that the views and preferences of patients who are actually experiencing the states are more well-informed. Pragmatically, it may be that the only source of values that is available is from the patients whose health states are being analysed, or that in some applications these are regarded by the relevant decision makers as being the most appropriate. In designing a study which collects EQ-5D data from patients in a context where there are no relevant value sets, it may be desirable to include an assessment of patients' values as an adjunct to the main analysis. However, there are a number of concerns about using patients' values in the context of, for example, economic evaluation. Differences between patients' and the general publics' valuation of states are widely observed, and there are a number of reasons for this, such as members of the general public valuing a state 'too low' where they cannot predict what their experience in that state would be, and their not taking into account patients' adaptation to that state (Brazier et al, 2004).

The *normative* argument usually advanced for using social valuations in economic evaluation is as follows. Broadly speaking, the purpose of any economic evaluation is to assess the 'value for money' of alternative uses of scarce healthcare resources. Where the context of these decisions is the public sector, it is generally argued that the valuation of health states used in the assessment of 'benefit' should reflect, as closely as possible, the preferences of the relevant general public (Weinstein et al, 1996). This is both because it is the general public who are funding healthcare, via taxes; and because the general public are potential users of the healthcare system and can provide valuations 'behind a veil of ignorance' (Dolan, 1999).

A slightly different issue is the question of which group of people these values are obtained from. A social value set could be obtained from a representative sample of the general public. However, it may be that there is no such set but that decision makers find it acceptable to use the values of other groups, such as informed members of the public, patients or experts as proxies. It may even be that decision makers prefer such values to represent social valuations. Here, we will assume that a representative sample is preferred.

Given that the weights used to quality-adjust length of life are generally meant to reflect the preferences of local taxpayers and potential recipients of healthcare, local (i.e. country-specific) value sets should be used where they are available. There has been some debate about whether or not it is appropriate to use the values from sub-groups of the population rather than the population as a whole – for example, the values of women or older people for conditions which only affect them (Sculpher and Gafni, 2001; Robinson and Parkin, 2002; Sculpher and Gafni, 2002) – but there is currently no consensus on this issue.

In most countries, value sets are available for *either* VAS or TTO; there are many countries with no available EQ-5D value set at all e.g. in Europe, France; and in North America, Canada. However, an EQ-5D value set now exists for at least one country in every continent, making the selection of a value set based on geographical proximity more feasible. A VAS-based value set exists for Europe which might reasonably be used to approximate health preferences across the EU countries (Greiner et al, 2003). The UK does have value sets for some sub-groups of the UK population (MVH 1995 contains 32 separate value sets, including sets for various age groups), although these are not widely promoted.

A final issue regarding relevance is the point in time that value sets were generated. Just as there are important differences in health state values between countries (as is evident in the value sets we report in the remainder of this booklet), it is possible there may be differences in a country's values over time. This would arise if preferences regarding health are not stable, as is normally assumed in economics, but change over time (Bridges, 2003), perhaps because of changing experience of and expectations about health. Although we currently have no evidence to suggest this is the case for

EQ-5D valuations, a more recent value set is preferable to an older one, providing they are equally relevant in other ways, and are otherwise comparable on the empirical and theoretical grounds discussed next.

Empirical characteristics

Although most value sets included in this inventory have been published in peer-reviewed journals, and therefore meet the scientific standards of those journals, there are no officially endorsed 'EuroQol Group value set' products as such. While the EuroQol Group promulgates a standard instrument for measuring the EQ-5D profile and EQ VAS for self-rated health, there has been somewhat less consistency in the instruments and research protocols used in the elicitation of valuations for *hypothetical* EQ-5D states. Although VAS is the valuation approach endorsed by the EuroQol Group, there has also been considerable research using TTO. Each value set has been produced by individual researchers. Researchers have employed slightly different instruments, using different protocols, and analysed their data using different econometric procedures – and the resulting value sets will reflect this. *Caveat emptor* applies for potential users.

We recommend that users familiarise themselves with the characteristics of the value sets before selecting one for use. Obvious questions to ask include: was there a reasonable response rate? Is the sample representative of the general public? Is there any cause for concern about data quality? (For example, were there high rates of missing or implausible valuations?).

In particular, users should think carefully about characteristics of the value set that may be important given the specific clinical context to which they want to apply the value sets. For example, if the condition under consideration involves very severe states, the means by which values for states worse than Dead have been calculated, rescaled or 'bounded' in the value set will be of particular relevance. If the health states pertaining to the condition of interest are experienced for long durations, how this relates to the duration of states described in the valuation exercise should be considered, given the possible effect of "maximum endurable time" (Sutherland et al, 1982) on valuations and, in TTO, the assumption of "constant proportionality" (Dolan and Stalmeier, 2003). If the treatment under consideration involves marginal improvements from very good health states to Full health, the way in which the constant term has been handled in modeling will effect the estimated improvement in quality adjusted life years.

Theoretical properties of alternative valuation methods

There are several alternative methods for valuing health states, including magnitude estimation, paired comparisons, rating scales such as a VAS, SG and TTO. These may have different theoretical properties, which may be an important factor in the choice of valuation states. However, we will only discuss VAS, SG and TTO, as these cover the range of theoretical issues that are important in this context. Although these issues are discussed here to help you decide which EQ-5D value set to choose, these same issues are faced by researchers regardless of which descriptive system is used to measure health.

Early in its development, the EuroQol Group adopted the VAS as its standard valuation method (Kind, 2003). The main reason for this was that self-completion questionnaires were seen as the only practical means of obtaining large valuation data sets and the VAS was the most suited to such a survey instrument. The VAS has continued to be the most widely used approach in EuroQol Group valuation studies. In contrast, the health economics literature seems to have reached something approaching a consensus that VAS valuations are inferior to other approaches, such as SG and TTO, on theoretical grounds (e.g. Brazier et al, 1999). This poses a dilemma for potential users of EQ-5D value sets, particularly where value sets that are leading candidates for use on relevance or empirical grounds are generated using VAS techniques.

Discussion of these issues is complicated by semantic disagreements. Does a 'QALY' refer to any attempt to quality-adjust life years – or does it specifically refer to length of life weighted by the *utility* experienced in each state? (Richardson, 1994; Parkin and Devlin, 2006). Does 'utility' refer to any attempt to elicit preferences on a 0-1 scale – or can valuations only be referred to as utilities if they meet the utility-under-uncertainty requirements of Von Neumann-Morgenstern (VNM) expected utility theory? (Drummond et al, 2005). Is Cost-Utility Analysis (CUA) a term that can only be used to describe economic evaluations that use utilities to estimate QALYs? If so, what term do we apply to an economic evaluation which reports cost-per-QALY evidence where the QALYs have been calculated using valuations that are *not* VNM utilities?

In brief, the theoretical grounds for selecting values with which to quality-adjust life years is often described in the health economics literature as follows. The Standard Gamble is the only technique that elicits preferences under conditions of uncertainty, thus meeting the requirements of VNM utility theory, and therefore SG is commonly suggested to have primacy on theoretical grounds. TTO was developed as a more pragmatic means of eliciting health state valuations that have similar empirical properties to SG valuations. Thus, while they are *not* utilities in the sense of utility-under-uncertainty, they generate valuations that are empirically similar. Further, both SG and TTO involve eliciting valuations by observing the trade-offs participants state they are prepared to accept between risk and quality of life (SG) and length and quality of life (TTO) when presented with a series of hypothetical scenarios. The use of trade-offs is appealing to economists: theories of value (e.g., Hicks, 1943) generally ascribe the value of a thing as the amount of another thing one is prepared to forego to obtain it. Thus both SG and TTO are often described as 'choice-based' and, incorrectly, given both approaches are *stated* preference techniques, as 'revealing preferences' (see, for example, Drummond et al 2005, p. 145). VAS is described as 'choice-less', as participants are simply asked to state their values.

The key objections to the use of VAS in CUA are that it is alleged that VAS lacks a theoretical foundation and is not related to the underlying theory of QALYs; VAS involves no choice or trade-off, and has no basis in economic theory; and VAS values are only appropriate for problems that involve certainty, and therefore have limited applicability in health care. Despite the apparent unanimity of opinion favouring SG and TTO over VAS, there are contrasting views on each of these issues (Parkin and Devlin, 2006; Naylor and Llewellyn-Thomas, 1998).

First, the underlying theoretical foundation of QALYs derives from extra-welfarism (Sen, 1977) – the idea that social choices should not be based only on individuals' utilities, but rather that something else be maximised – usually, in this context, health (Culyer, 1991). Viewed in this way the QALY is no more than a convenient device to combine length and quality of life into a single metric, which replaces utility as the objective function to be maximised (Parkin and Devlin, 2006). The theoretical foundations of QALYs therefore do not require that quality of life be valued using a particular measurement method. Second, VAS, TTO and SG are *all* techniques for eliciting *stated* preferences. While SG and TTO seek these preferences indirectly i.e., by the use in both of a 'numeraire', tradeoffs against which are used to value health states, this introduces a variety of biases associated with the characteristics of the numeraire in each case – risk (risk aversion) and length of life (time preference and maximal endurable time) for SG and TTO respectively. VAS valuation arguably employs a more direct approach: participants can consider, compare and assign relative values to all states under consideration – rather than valuing each separately as in SG and TTO. However, VAS too is associated with various biases – for example, end-of-scale and spacing out bias (Drummond et al, 2005). Finally, theoretical arguments about risk and uncertainty in favour of the SG are valid only at the individual level – not at a social decision-making level (Drummond et al, 2005).

Thus the theoretical and empirical case for favouring any one method of health state valuation over another is far from clear-cut. In practice, there are currently no EQ-5D value sets generated from SG methods, so for users the choice is between TTO and VAS.

Uncertainty over values: sensitivity analysis

There may not be any one value set which is unequivocally 'the best' for a given application on each of the criteria discussed above. This suggests that the EQ-5D values used in economic evaluation would appropriately be considered as uncertain parameters. Given that there may be more than one value set appropriate for use, analysts should examine whether the choice of value set makes any substantive difference to results and conclusions i.e., the analyst should treat the values in an economic evaluation as uncertain parameters which, just as with other non-stochastic uncertain variables such as the discount rate, should be subject to sensitivity analysis. Currently this is not a common practice– but it is readily done and would improve confidence in results.

There are clearly differences between the actual values for each EQ-5D state reported by the various value sets described in this inventory – but the magnitude of these differences, and their implications for estimates of quality adjusted life years, is not always obvious. For example, if one value set contains values that are systematically higher (or lower) than another for the health states relevant to a given therapy, these differences may 'wash out' in economic evaluation, which focuses on the incremental *change* in health resulting from that therapy.

4.3 Can I use EQ-5D values if I have not included the EQ-5D in my study as a health outcome measure?

The recommended approach in designing a study where one of the required outcomes is a single index of health is to use the EQ-5D instrument, obtain data on the EQ-5D profiles of the relevant patients or populations, and to apply to these an appropriate EQ-5D value set. However, it may be that analysts wish to generate values where EQ-5D data have not been collected but other health state instruments, such as condition specific or clinical measures, have been used. It is possible, under certain conditions, to use EQ-5D value sets for this purpose, though there will be some loss of precision in the estimates of values.

What is required is a mapping between the health state measure which has been used and the EQ-5D, so that an equivalence is reached between profiles described by the two. This could be generated by the analyst, or taken from published studies. The methods of achieving a mapping range from expert judgement to patient surveys.

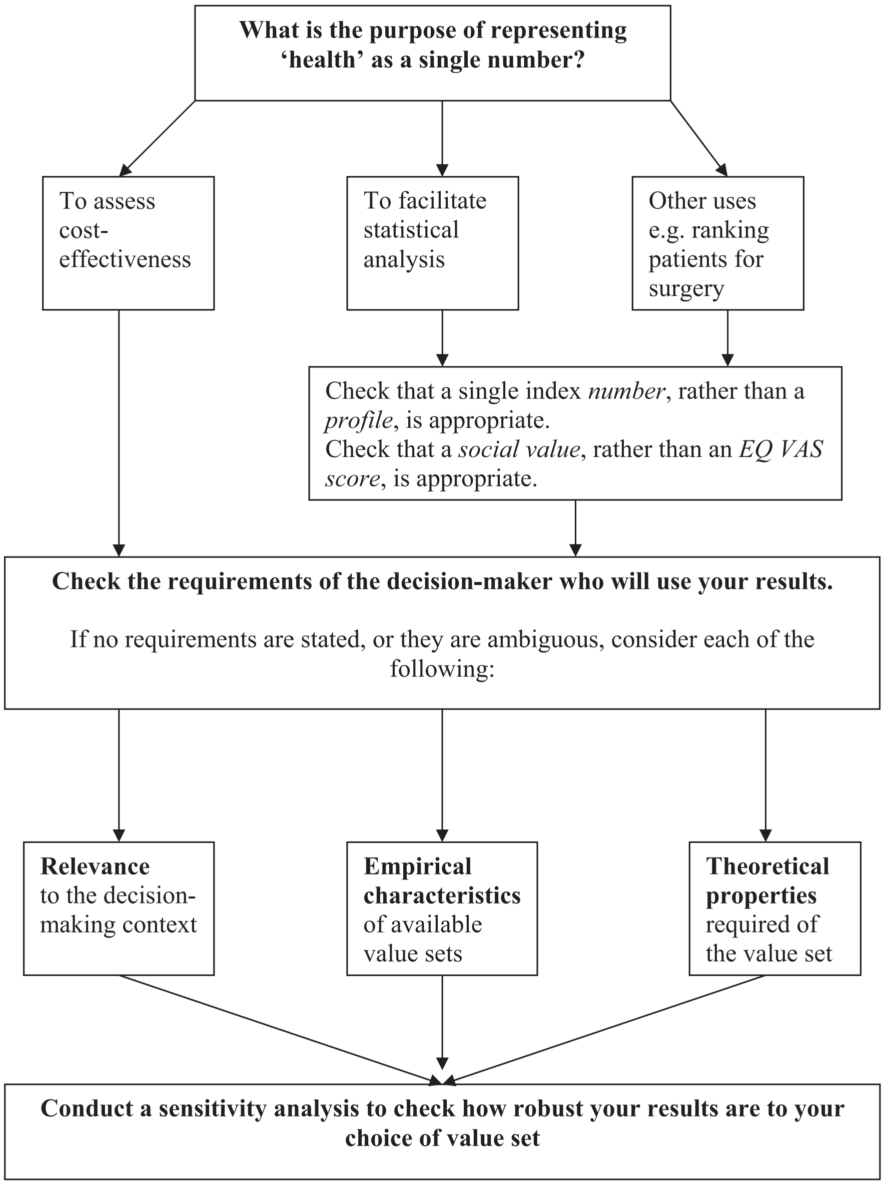
This approach has been widely used, for example in the economic evaluation of beta interferon therapy in multiple sclerosis (Parkin et al, 1998; Chilcott et al, 2003). However, it must be recognised that mapping may involve subjective judgements; that the valuations obtained are very much dependent on the quality of the mapping technique used; and are almost certain to involve a loss of precision by the addition of an additional source of approximation to “true” values.

4.4 Conclusions

Figure 5 provides an overview of the considerations that should determine your choice between the EQ-5D value sets presented in the following chapters. There is no simple answer to the question of which value set to use: the answer depends on the specific nature of the research application, the sort of decisions it informs, and the context in which the evidence from your research will be used.

In some cases, which value set to use will be determined by the stated requirements of those using the evidence to inform decision-making. Where this is not the case, we encourage potential users of EQ-5D value sets carefully to consider each of the practical and theoretical issues discussed in this chapter. Where there remains uncertainty over which value set to use, we recommend that researchers should report the sensitivity of their results and conclusions to the use of alternative value sets. If results are not substantially affected by the choice of value set, this increases confidence in the findings. If results and conclusions are contingent on the specific value set used, this is important to convey to those who will use your research.

Figure 5: Which EQ-5D value set should I use?



References

Badia X, Roset M, Monserrat S, Herdman M. The Spanish VAS tariff based on valuation of EQ-5D health states from the general population. In: Rabin RE et al, editors. EuroQol Plenary meeting Rotterdam 1997, 2-3 October. Discussion papers. Centre for Health Policy & Law, Erasmus University, Rotterdam, 1998: 93-114.

Badia X, Roset R, Herdman, M, Kind P. A comparison of GB and Spanish general population time trade-off values for EQ-5D health states. Med Decis Making 2001; 21(1): 7-16.

Brazier J, Deverill M, Green C, Harper R, Booth A. A review of the use of health status measures in economic evaluation. NHS R&D HTA programme, Health Technol Assess 1999; 3(9):i-164.

Brazier J, Akehurst R, Brennan A, Dolan P, Klaxton C, McCabe C, O'Hagan T, Sculpher M, Tsuchiya A. Should patients have a greater role in valuing health states: whose well-being is it anyway? Paper presented at the CES/HESG conference, Paris, January 2004.

Bridges J. Stated preference methods in health care evaluation: an emerging methodological paradigm in health economics. Appl Health Econ Health Policy 2003; 2(4):213-224.

Brooks R, Rabin R, de Charro F, editors. The measurement and valuation of health status using EQ-5D: a European perspective. Kluwer Academic Publishers 2003.

Chilcott J, McCabe C, Tappenden P, O'Hagan A, Cooper N, Abrams K, Claxton K. Modeling the cost-effectiveness of interferon beta and glatiramer acetate in the management of multiple sclerosis. BMJ 2003; 326(7388):522; discussion 522.

Claes C, Greiner W, Uber A, Graf von der Schulenburg JM. An interview-based comparison of the TTO and VAS values given to EuroQol states of health by the general German population. In: Greiner W, J-M. Graf v.d. Schulenburg, Piercy J, editors. EuroQol Plenary Meeting, 1-2 October 1998. Discussion papers. Centre for Health Economics and Health Systems Research, University of Hannover, Germany. Uni-Verlag Witte, 1999; 13-39.

Cleemput I. Economic evaluation in renal transplantation: outcome assessment and cost-utility of non-compliance. Acco, Leuven 2003.

Culyer AJ. The normative economics of health care finance and provision. In: McGuire A, Fenn P, Mayhew K, editors. Providing health care. Oxford University Press 1991.

Derrett S, Devlin N, Hansen P, Herbison P. Prioritising patients for elective surgery. A prospective study of clinical priority assessment criteria in New Zealand. Int J Technol Assess Health Care 2003; 19(1):91-105.

Devlin NJ, Hansen P, Kind P, Williams A. Logical inconsistencies in survey respondents' health state valuations - a methodological challenge for estimating social tariffs. Health Econ 2003; 12(7):529-544.

Dolan P. Modeling valuations for EuroQol health states. Med Care 1997; 35(11):1095-1108.

Dolan P. Whose preferences count? Med Decis Making 1999; 19(4):482-486.

Dolan P, Stalmeier P. The validity of time trade-off values in calculating QALYs: constant proportional time trade-off versus the proportional heuristic. J Health Econ 2003; 22(3):445-458.

Drummond MF, Sculpher M, Torrance G, O'Brien B, Stoddart G. Methods for the economic evaluation of health care programmes. Oxford University Press, 3rd edition 2005.

EuroQol Group. EuroQoL-a new facility for the measurement of health-related quality of life. Health Policy 1990;16(3):199-208).

Greiner W, Weijnen T, Nieuwenhuizen M, et al. A single European currency for EQ-5D health states. Results from a six country study. Eur J Health Econ 2003; 4(3):222-231.

Greiner W, Claes C, Buschbach JJV, Graf von der Schulenburg JM. Validating the EQ-5D with time trade off for the German population. Eur J Health Econ 2005; 6(2):124-130.

Hicks JR. The four consumer's surpluses. Rev Econ Stud 1943; 11(1):31-41.

Jelsma J, Hansen K, De Weerdt W, De Cock P, Kind P. How do Zimbabweans value health states? Popul Health Metr 2003; 1(1):11.

Kind P. Guidelines for value sets in economic and non-economic studies using EQ-5D. Chapter 4 in: Brooks R, Rabin R, de Charro F, editors. The measurement and valuation of health status using EQ-5D: a European perspective. Kluwer 2003.

Kind P, Hardman G, Macran S. UK Population norms for EQ-5D. York Centre for Health Economics, Discussion Paper, University of York 1999.

Lamers LM, McDonnell J, Stalmeier PFM, Krabbe PFM, Busschbach JJV. The Dutch Tariff: Results and arguments for an effective design for national EQ-5D valuation studies. Accepted for publication in Health Economics.

MVH Group. The Measurement and Valuation of Health. Final report on the modeling of valuation tariffs. York Centre for Health Economics 1995.

Naylor CD, Llewellyn-Thomas HA. Utilities and preferences for health states: time for a pragmatic approach? J Health Serv Res Policy 1998; 3(3):129-131.

National Institute of Clinical Excellence. Guidance for Manufacturers and Sponsors. Technology Appraisal Process Series No.5, 2001 (http://www.nice.org.uk).

National Institute for Clinical Excellence. Guide to methods of technology appraisal 2004 (http://www.nice.org.uk/pdf/TAP_Methods.pdf).

Ohinmaa A, Eija H, Sintonen H. Modeling EuroQol values of Finnish adult population. In: Badia X, Herdman M, Segura A, editors. EuroQol Plenary Meeting Barcelona 1995. Discussion Papers. Institut Universitari de Salut Publica de Catalunya 1996; 67-76.

Ohinmaa A, Sintonen H. Inconsistencies and modeling of the Finnish EuroQol (EQ-5D) preference values. In: Greiner W, J-M. Graf v.d. Schulenburg, Piercy J, editors. EuroQol Plenary Meeting, 1-2 October 1998. Discussion papers. Centre for Health Economics and Health Systems Research, University of Hannover, Germany. Uni-Verlag Witte 1999; 57-74.

Parkin D, McNamee P, Jacoby A, Miller P, Thomas S, Bates D. A cost-utility analysis of interferon beta for multiple sclerosis. Health Technol Assess 1998; 2(4): i-58.

Parkin D, Devlin N. Is there a case for using visual analogue scale valuations in cost-utility analysis? Health Econ 2006 [in press].

Prevolnik Rupel V, Rebolj M. The Slovenian VAS Tariff based on valuations of EQ-5D health states from the general population. In: Cabasés JM, Gaminde I, editors. Proceedings of the 17th Plenary Meeting of the EuroQol Group. Universidad Pública de Navarra 2001; 23-47.

Richardson J. Cost-Utility Analysis: what should be measured? Soc Sci Med 1994; 39(1):7-21.

Robinson A, Parkin D. Recognising diversity in public preferences: the use of preference sub-groups in cost-effectiveness analysis. A response to Sculpher and Gafni. Health Econ 2002; 11(7):649.

Sculpher M, Gafni A. Recognising diversity in public preferences: the use of preference sub-groups in cost-effectiveness analysis. Health Econ 2001; 10(4):317-324.

Sculpher M, Gafni A. Recognising diversity in public preferences: the use of preference sub-groups in cost-effectiveness analysis. Authors Reply. Health Econ 2002; 11(7):653.

Sen A. Social choice theory: a re-examination. Econometrica 1977; 45(1):53-90.

Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. Med Care 2005; 43(3):203-220.

Sutherland HJ, Llewellyn-Thomas H, Boyd NF, Till JE. Attitudes toward quality of survival. The concept of "maximal endurable time". Med Decis Making 1982; 2(3):299-309.

Torrance GW. Measurement of health state utilities for economic appraisal. J Health Econ 1986; 5(1):1-30.

Tsuchiya A, Ikeda S, Ikegami N, et al. Estimating an EQ-5D population value set: The case of Japan. Health Econ 2002; 11(4):341-353.

Weinstein, MC, Siegel J, Gold MR, Kamlet MS, Russell LB. Recommendations of the panel on cost-effectiveness in health and medicine. JAMA 1996; 276(15):1253-1258.

Wittrup-Jensen KU, Lauridsen JT, Gudex C, Brooks R, Pedersen KM. Estimating Danish EQ-5D tariffs using TTO and VAS. In: Norinder A, Pedersen K, Roos P, editors. Proceedings of the 18th Plenary Meeting of the EuroQol Group. IHE, The Swedish Institute for Health Economics 2002; 257-292.

Annex 1:

Inventory of Time Trade-Off valuation surveys

Danish TTO value set		Example: the value for health state 21232	
Full health (11111)	1	Full health	= 1
At least one 2 or 3 (constant)	- 0.114	Minus constant	- 0.114
At least one 3 (N3)	–	Minus N3	–
Mobility = 2	- 0.053	Minus MO level 2	- 0.053
Mobility = 3	- 0.411		
Self care = 2	- 0.063	Minus SC level 1	- 0.000
Self care = 3	- 0.192		
Usual activities = 2	- 0.048	Minus UA level 2	- 0.048
Usual activities = 3	- 0.144		
Pain/discomfort = 2	- 0.062		
Pain/discomfort = 3	- 0.396	Minus PD level 3	- 0.396
Anxiety/depression = 2	- 0.068	Minus AD level 2	- 0.068
Anxiety/depression = 3	- 0.367		
		State 21232	= 0.321

DATA COLLECTION

Data were collected in an interviewer setting during the Spring of 2000. 4075 addresses were contacted. 1421 were not at home after 3 attempts and 1321 refused to participate. A total of 1332 respondents completed the interview, that is 50% of the people who were contacted.

EXCLUSION CRITERIA

No exclusions were reported.

SAMPLE CHARACTERISTICS

Of the 1332 respondents included in the analyses, 42.0% were male and 58.0% were female. The age distribution of the respondents was:

18-29 yrs	15.8%
30-59 yrs	55.1%
≥ 60 yrs	29.1%

VALUATION METHODS

A computer assisted interview method was used. Each respondent was asked to do the following:

1. EQ-5D descriptive system and EQ VAS.
2. Ranking exercise.
3. Valuation of the ranked health states on a VAS.
4. TTO exercise.

In the TTO exercise, each respondent valued 16 states.

- 22222 and 33333.
- 2 out of 5 'mild' states (21111, 12111, 11211, 11121, 11112).
- 8 other states (used in earlier EuroQol Group investigations).
- 2 'hold out'-states (as Danish diabetes or heart disease patients described their own health).

States 11111 and Dead served as anchor points. A total of 46 states were directly valued, including all common core states.

The TTO valuations were transformed to lie on the interval [-1,1]. States regarded as better than Dead were calculated as t/10, where t is the number of years in 11111. States regarded as worse than Dead were calculated as -t/10.

A random effects model was used to derive an additive utility function. Disutilities (1-S) were used in the model estimations. 12 different models were tested. The resulting Danish value set was based on a main effects model including:

- A variable for each of the 5 dimensions (0 = no problems, 1 = some problems, 2 = extreme problems).
- A dummy variable for each of the 5 dimensions (1 = extreme problems, 0 = otherwise).
- The intercept, representing any deviation from Full health.

The R^2 value of the model was 0.66.

VALUE SET

The mathematical representation of the model for health state X is:

X = 1 - 0.114 - 0.053 MO2 - 0.411 MO3 - 0.063 SC2 - 0.192 SC3 - 0.048 UA2 - 0.144 UA3 - 0.062 PD2 - 0.396 PD3 - 0.068 AD2 - 0.367 AD3

A Danish value set that includes the N3 factor (used when any dimension is at level 3) is also available from the EuroQol Executive Office.

REFERENCES

Wittrup-Jensen KU, Lauridsen JT, Gudex C, Brooks R, Pedersen KM. Estimating Danish EQ-5D tariffs using TTO and VAS. In: Norinder A, Pedersen K, Roos P, editors. Proceedings of the 18^{th} Plenary Meeting of the EuroQol Group. IHE, The Swedish Institute for Health Economics, 2002; 257-292.

German TTO value set		**Example: the value for health state 21232**	
Full health (11111)	1	Full health	= 1
At least one 2 or 3 (constant)	- 0.001	Minus constant	- 0.001
At least one 3 (N3)	- 0.323	Minus N3	- 0.323
Mobility = 2	- 0.099	Minus MO level 2	- 0.099
Mobility = 3	- 0.327		
Self care = 2	- 0.087	Minus SC level 1	- 0.000
Self care = 3	- 0.174		
Usual activities = 2	–	Minus UA level 2	–
Usual activities = 3	–		
Pain/discomfort = 2	- 0.112		
Pain/discomfort = 3	- 0.315	Minus PD level 3	- 0.315
Anxiety/depression = 2	–	Minus AD level 2	–
Anxiety/depression = 3	- 0.065		
		State 21232	= 0.262

DATA COLLECTION

The German valuation study was a combined VAS / TTO study. Data were collected in an interviewer setting from October 1997 - March 1998. 380 addresses were randomly selected from a telephone directory and distributed to interviewers. 41 respondents could not be contacted and 5 interviews had to be aborted because the respondent could not master the task.

EXCLUSION CRITERIA

Extreme values were excluded: 68 values from 18 different health states deviated so much from the median scores for these states that they were considered to be faulty and were excluded from the analysis.

SAMPLE CHARACTERISTICS

Of the 339 respondents included in the analyses, 55.3% were male and 44.7% were female. The age distribution of the respondents was:

15-25 yrs	7.7 %
25-45 yrs	31.1 %
45-65 yrs	40.2 %
65+ yrs	21.0 %

VALUATION METHODS

Each respondent was asked to do the following:

1. EQ-5D descriptive system.
2. Ranking exercise.
3. Valuation of the ranked health states on a VAS.
4. EQ VAS.
5. TTO exercise.

In the TTO exercise, each respondent valued 13 states. States 11111 and Dead served as anchor points. Unconscious was not included in the survey. A total of 35 health states were valued.

The base TTO values were transformed according to the following set of rules:

States better than Dead: (t = number of years in 11111)		States worse than Dead: (x = number of years in target state)	
t > 5	→ (t - 0.5) / 10	x > 5	→ [(x + 0.5) / 10] -1
t < 5	→ (t + 0.5) / 10	x < 5	→ [(x - 0.5) / 10] -1
t = 0, 5 or 10	→ t / 10	x = 0, 5 or 10	→ [x / 10] –1

If t was indicated in weeks → (9 + t/52) / 10

A linear additive model specification was used. The model consists of 1 variable for each of the 5 dimensions (0 = no problems, 1 = some problems, 2 = extreme problems), plus a dummy variable for each of the 5 dimensions (1 = extreme problems, 0 = otherwise), plus a constant and a N3 term.

The model yields value sets based on the following formula:
Constant - coefficient*value of dimension (i.e. 0, 1 or 2) - coefficient*value of dimension at level 3 (i.e. 0 or 1) - N3

Since not all of the coefficients of the model were statistically significant, the final model for the German TTO data value set was estimated using only the statistically significant variables.

The R^2 value of the model was 0.51.

VALUE SET

The mathematical representation of the model for health state X is:

X = 1 - 0.001 - 0.099 MO2 - (2*0.099 + 0.129) MO3 - 0.087 SC2 - (2*0.087) SC3 - 0.112 PD2 - (2*0.112 + 0.091) PD3 - 0.065 AD3 - 0.323 N3

REFERENCES

Greiner W, Claes C, Buschbach JJV, Graf von der Schulenburg JM. Validating the EQ-5D with time trade off for the German population. Eur J Health Econ 2005; 6(2):124-130.

Japanese TTO value set		**Example: the value for health state 21232**	
Full health (11111)	1	Full health	= 1
At least one 2 or 3 (constant)	- 0.152	Minus constant	- 0.152
At least one 3 (N3)	–	Minus N3	–
Mobility = 2	- 0.075	Minus MO level 2	- 0.075
Mobility = 3	- 0.418		
Self care = 2	- 0.054	Minus SC level 1	- 0.000
Self care = 3	- 0.102		
Usual activities = 2	- 0.044	Minus UA level 2	- 0.044
Usual activities = 3	- 0.133		
Pain/discomfort = 2	- 0.080		
Pain/discomfort = 3	- 0.194	Minus PD level 3	- 0.194
Anxiety/depression = 2	- 0.063	Minus AD level 2	- 0.063
Anxiety/depression = 3	- 0.112		
		State 21232	= 0.472

DATA COLLECTION

Data were collected in August and September of 1998. People aged 20 and above were sampled in 3 prefectures. After exclusions, the data were weighted for age and gender according to the Japanese national population. 972 people were selected, of which 621 agreed to take part in the survey (63.9%).

EXCLUSION CRITERIA

a. Completely missing TTO data (n = 57).
b. < 3 states valued (n = 3).
c. All states valued the same (n = 18).
d. All states valued worse than Dead (n = 1).

Thus 78 of the 621 respondents were excluded and 543 remained for analysis.

SAMPLE CHARACTERISTICS

Of the 543 respondents included in the analyses, 57.6% were male and 42.4% were female. The mean age of the respondents was 48.14 years.

VALUATION METHODS

Each respondent was asked to do the following:

1. EQ-5D descriptive system and EQ VAS.
2. VAS valuation of 14 health states.
3. Ranking of 19 health states.
4. TTO exercise.

States 11111 and Dead served as anchor points and were therefore not part of the TTO task. Each respondent valued the same 17 states:

11112	11312	32211
11113	12111	32223
11121	13311	32313
11131	21111	33323
11133	22222	33333
11211	23232	

The 19 health states that were part of the ranking exercise in part 2 were the 17 health states used in the TTO exercise, with the additional states of 11111 and Dead.

The TTO valuations were transformed to lie on the interval [-1,1]. States regarded better than Dead were calculated as t/10, where t is the number of years in 11111. States regarded worse than Dead were calculated as -t/10.

A choice was made to carry out the main analysis by OLS regressions without accounting for respondent effects. Several models were tested. A main effects model was selected as the final model.

The R^2 value of the model was 0.400.

VALUE SET

The mathematical representation of the main effects model for health state X is:

X = 1 - 0.152 - 0.075 MO2 - 0.418 MO3 - 0.054 SC2 - 0.102 SC3 - 0.044 UA2 - 0.133 UA3 - 0.080 PD2 - 0.194 PD3 - 0.063 AD2 - 0.112 AD3

REFERENCES

Tsuchiya A, Ikeda S, Ikegami N, et al. Estimating an EQ-5D population value set: The case of Japan. Health Economics 2002; 11(4):341-353.

Dutch TTO value set		Example: the value for health state 21232	
Full health (11111)	1	Full health	= 1
At least one 2 or 3 (constant)	- 0.071	Minus constant	- 0.071
At least one 3 (N3)	- 0.234	Minus N3	- 0.234
Mobility = 2	- 0.036	Minus MO level 2	- 0.036
Mobility = 3	- 0.161		
Self care = 2	- 0.082	Minus SC level 1	- 0.000
Self care = 3	- 0.152		
Usual activities = 2	- 0.032	Minus UA level 2	- 0.032
Usual activities = 3	- 0.057		
Pain/discomfort = 2	- 0.086		
Pain/discomfort = 3	- 0.329	Minus PD level 3	- 0.329
Anxiety/depression = 2	- 0.124	Minus AD level 2	- 0.124
Anxiety/depression = 3	- 0.325		
		State 21232	= 0.174

DATA COLLECTION

A marketing research company recruited respondents between the ages of 18 and 75 from the Rijnmond area. Quota sampling was used to achieve a sample that was representative of the Dutch population with regard to age and gender. Face-to-face interviews were conducted in the summer of 2003. In total, 309 respondents were interviewed - 2 discontinued the interview before the TTO valuation task and data from 2 respondents were lost due to technical problems.

EXCLUSION CRITERIA

Data from 7 respondents were excluded because they valued each state as 1. After exclusions, data from 298 respondents remained for analysis.

SAMPLE CHARACTERISTICS

Of the 298 respondents included in the analysis, 51% were male and 49% were female. The age distribution of the respondents was:

18-24 yrs 12.1%

25-34 yrs 19.8%

35-44 yrs 21.5%

45-54 yrs 20.8%

55-64 yrs 16.4%

65-74 yrs 9.4%

VALUATION METHODS

Each respondent was asked to do the following:

1. EQ-5D descriptive system and EQ VAS.
2. Ranking exercise.
3. Valuation of the ranked health states on a VAS.
4. TTO exercise.

The TTO exercise was performed on a computer. A graphic computer program replaced the TTO boards. The program presented the health states to be valued in random order. States 11111 and Dead served as anchor points and were therefore not part of the TTO task. Each respondent valued the same 17 states:

11112	11211	21111	32223
11113	11312	22222	32313
11121	12111	23232	33323
11131	13311	32211	33333
11133			

The TTO valuations were transformed to lie on the interval [-1,1]. States regarded better than Dead were calculated as t/10, where t is the number of years in 11111. States regarded worse than Dead were calculated as -t/10.

A random effects model was used to derive an additive utility function. Disutilities (1-S) were used in the model estimations. A number of different models were tested. The resulting Dutch value set was based on a N3 model including:

- Two dummy variables for level 2 and level 3 in each dimension.
- The intercept, representing any deviation from Full health.
- A N3 dummy indicating whether 1 or more of the dimensions was on level 3.

The R^2 value of the model was 0.38.

VALUE SET

The mathematical representation of the model for health state X is:

X = 1 - 0.071 - 0.036† MO2 - 0.161 MO3 - 0.082 SC2 - 0.152 SC3 - 0.032† UA2 - 0.057 UA3 - 0.086 PD2 - 0.329 PD3 - 0.124 AD2 - 0.325 AD3 - 0.234 N3

† not statistically significant different from 0 ($p > 0.05$).

REFERENCES

Lamers LM, McDonnell J, Stalmeier PFM, Krabbe PFM, Busschbach JJV. "The Dutch Tariff: Results and arguments for an effective design for national EQ-5D valuation studies". Accepted for publication in Health Economics.

Spanish TTO value set		Example: the value for health state 21232	
Full health (11111)	1	Full health	= 1
At least one 2 or 3 (constant)	- 0.024	Minus constant	- 0.024
At least one 3 (N3)	- 0.291	Minus N3	- 0.291
Mobility = 2	- 0.106	Minus MO level 2	- 0.106
Mobility = 3	- 0.430		
Self care = 2	- 0.134	Minus SC level 1	- 0.000
Self care = 3	- 0.309		
Usual activities = 2	- 0.071	Minus UA level 2	- 0.071
Usual activities = 3	- 0.195		
Pain/discomfort = 2	- 0.089		
Pain/discomfort = 3	- 0.261	Minus PD level 3	- 0.261
Anxiety/depression = 2	- 0.062	Minus AD level 2	- 0.062
Anxiety/depression = 3	- 0.144		
		State 21232	= 0.185

DATA COLLECTION

Valuation data were collected within one primary health care district on the outskirts of Barcelona in 1997. Respondents were selected using quota sampling to match the Spanish general population by age and gender according to the 1991 census. 1930 individuals were contacted to obtain a sample of a 1000 individuals.

EXCLUSION CRITERIA

25 questionnaires were excluded principally due to serious inconsistencies in the valuation of the health states.

SAMPLE CHARACTERISTICS

Of the 975 respondents included in the analyses, 46% were male and 54% were female. The age distribution of the respondents was:

18-34 yrs	33%
35-49 yrs	24%
50-59 yrs	16%
60+ yrs	27%

VALUATION METHODS

Each respondent was asked to do the following:

1. The EQ-5D descriptive system and EQ VAS.
2. Ranking exercise.
3. TTO exercise.

In the TTO exercise, each respondent valued 13 states:

- 33333 and Unconscious.
- 2 from 5 'very mild' states (11112, 11121, etc).
- 3 from 12 'mild' states (11122, 11131, etc).
- 3 from 12 'moderate' states (13212, 32331, etc).
- 3 from 12 'severe' states (33232, 23232, etc).

States 11111 and Dead served as anchor points. 43 health states were valued in total.

The TTO valuations were transformed to lie on the interval [-1,1]. States regarded better than Dead were calculated as t/10, where t is the number of years in 11111. States regarded worse than Dead were calculated as -t/10.

Several random effects models were tested. The final model was selected based on goodness of fit, parsimony and consistency.

The R^2 value of the model was 0.60.

VALUE SET

The mathematical representation of the model for health state X is:

X = 1 - 0.024 - 0.106 MO2 - 0.430 MO3 - 0.134 SC2 - 0.309 SC3 - 0.071 UA2 - 0.195 UA3 - 0.089 PD2 - 0.261 PD3 - 0.062 AD2 - 0.144 AD3 - 0.291 N3

REFERENCES

Badia X, Roset R, Herdman, M, Kind P. A comparison of GB and Spanish general population time trade-off values for EQ-5D health states. Med Decis Making 2001; 21(1): 7-16.

UK TTO value set		**Example: the value for health state 21232**	
Full health (11111)	1	Full health	= 1
At least one 2 or 3 (constant)	- 0.081	Minus constant	- 0.081
At least one 3 (N3)	- 0.269	Minus N3	- 0.269
Mobility = 2	- 0.069	Minus MO level 2	- 0.069
Mobility = 3	- 0.314		
Self care = 2	- 0.104	Minus SC level 1	- 0.000
Self care = 3	- 0.214		
Usual activities = 2	- 0.036	Minus UA level 2	- 0.036
Usual activities = 3	- 0.094		
Pain/discomfort = 2	- 0.123		
Pain/discomfort = 3	- 0.386	Minus PD level 3	- 0.386
Anxiety/depression = 2	- 0.071	Minus AD level 2	- 0.071
Anxiety/depression = 3	- 0.236		
		State 21232	= 0.088

DATA COLLECTION

Data were collected in the period from August - December 1993. The sample was drawn from the national postcode address file (SCPR) and was representative of the non-institutionalised adult population of England, Scotland and Wales. 6080 addresses were drawn. 3395 (56%) were willing to be interviewed.

EXCLUSION CRITERIA

Respondents were excluded from the dataset according to the following criteria for both their VAS and their TTO valuations:

a. < 3 states valued.
b. All states valued the same.
c. 11111 and/or Dead not valued in the VAS valuation task.
d. Dead ≥ 11111 in the VAS valuation task.
e. Dead > all states other than 11111.
f. All data from one of the interviewers.
g. Any state missing from either VAS or TTO.

The total number of exclusions was 398. Therefore data from 2997 respondents were included in the analysis.

SAMPLE CHARACTERISTICS

Of the 2997 respondents included in the analysis, 43% were male and 57% were female. The age distribution of the respondents was:

18-24 yrs	9%
25-34 yrs	22%
35-49 yrs	25%
50-64 yrs	21%
65+ yrs	24%

VALUATION METHODS

Each respondent was asked to do the following:

1. EQ-5D descriptive system and EQ VAS.
2. Ranking exercise.
3. Valuation of the ranked health states on a VAS.
4. TTO exercise.

In the TTO exercise each respondent valued 13 states:

- 33333 and Unconscious.
- 2 from 5 'very mild' states (11112, 11121, etc).
- 3 from 12 'mild' states (11122, 11131, etc).
- 3 from 12 'moderate' states (13212, 32331, etc).
- 3 from 12 'severe' states (33232, 23232, etc).

States 11111 and Dead served as anchor points. 43 health states were valued in total. These states were:

11112	12222	22112	32223
11113	12223	22121	32232
11121	13212	22122	32313
11122	13311	22222	32331
11131	13332	22233	33212
11133	21111	22323	33232
11211	21133	22331	33321
11312	21222	23232	33323
12111	21232	23313	33333
12121	21312	23321	Unconscious
12211	21323	32211	

The TTO valuations were transformed to lie on the interval [-1,1]. States regarded better than Dead were calculated as t/10, where t is the number of years in 11111. States regarded worse than Dead were calculated as -t/10.

A generalised least-squares regression technique was used. The utility function was additive. Disutilities (1-S) were used in the model estimations. Two dummy variables were used for levels 2 and 3 in each dimension in order to allow for different utility increments between levels 1 and 2 as compared to the increment between levels 2 and 3.

Additional interaction terms between dummy variables across dimensions were used in order to measure 'extra' disutility when reporting problems on multiple dimensions. Individual-level (rather than aggregate-level) data analysis was performed. A random effects specification was used to address the problem that the scores for health states generated by an individual were related.

The R^2 value of the model was 0.46.

VALUE SET

The mathematical representation of the model for health state X is:

X = 1 – 0.081 – 0.069 MO2 – 0.314 MO3 – 0.104 SC2 – 0.214 SC3 – 0.036 UA2 – 0.094 UA3 – 0.123 PD2 – 0.386 PD3 – 0.071 AD2 – 0.236 AD3 – 0.269 N3

REFERENCES

MVH Group. The Measurement and Valuation of Health. Final report on the modeling of valuation tariffs. York Centre for Health Economics, 1995.

Dolan P. Modeling valuations for EuroQol health states. Med Care 1997; 35(11): 1095-108.

US TTO value set		Example: the value for health state 21232	
Full health (11111)	1	Full health	= 1
Mobility = 2	- 0.146	Minus MO level 2	- 0.146
Mobility = 3	- 0.558		
Self care = 2	- 0.175	Minus SC level 1	- 0.000
Self care = 3	- 0.471		
Usual activities = 2	- 0.140	Minus UA level 2	- 0.140
Usual activities = 3	- 0.374		
Pain/discomfort = 2	- 0.173		
Pain/discomfort = 3	- 0.537	Minus PD level 3	- 0.537
Anxiety/depression = 2	- 0.156	Minus AD level 2	- 0.156
Anxiety/depression = 3	- 0.450		
D1	+ 0.140	Plus D1	+ 3*0.140
I2-square	- 0.011	Minus I2-square	- 4*0.011
I3	+ 0.122	Plus I3	+ 0*0.122
I3-square	+ 0.015	Plus I3-square	+ 0*0.015
		State 21232	= 0.397

DATA COLLECTION

Data were collected between June 2002 and October 2002. A multi-stage probability sample was selected from the adult US household population using a sampling frame based on residential mailing lists and demographic data from the 2000 Census. The objective of the sampling design was to oversample Hispanics and non-Hispanic blacks while completing 4000 in-home interviews. Sampling weights were derived and applied to allow for appropriate extrapolation to the target population.

5237 persons were selected for interview. A total of 4048 interviews were completed, a non-weighted interview response rate of 77.3%

EXCLUSION CRITERIA

In total 11 exclusion criteria were used:

a. More than 1 missing TTO value (not including Unconscious) (n = 200).
b. All states valued the same (n = 27).
c. Subject valued 1 or more health states more than once (n = 17).
d. All states valued worse than Dead (n = 9).
e. Other exclusion criteria (n = 23).

The total number of exclusions was 275.

SAMPLE CHARACTERISTICS

Of the 3773 respondents included in the analysis, 49% were male and 51% were female. The mean age of the respondents was 44.5 years.

VALUATION METHODS

Each respondent was asked to do the following:

1. EQ-5D descriptive system.
2. Ranking exercise.
3. Value the health states on the VAS.
4. EQ VAS.
5. TTO exercise.

Each respondent was asked to value 13 health states in the TTO exercise. All respondents valued 33333 and Unconscious. States 11111 and Dead served as anchor points. 43 health states were valued in total. These states were:

11112	12222	22112	32223
11113	12223	22121	32232
11121	13212	22122	32313
11122	13311	22222	32331
11131	13332	22233	33212
11133	21111	22323	33232
11211	21133	22331	33321
11312	21222	23232	33323
12111	21232	23313	33333
12121	21312	23321	Unconscious
12211	21323	32211	

The TTO valuations were linearly transformed to lie on the interval [-1,1]. States regarded better than Dead were anchored on the Full health and Dead scale: t/10, where t is the number of years in 11111. States regarded worse than Dead were calculated as X/39, where X = -t/(10-t) (since the observed values ranged between 0 and -39, transformed values were obtained by dividing them by 39).

A number of different models and model specifications were tested. A split sample approach was used to evaluate the predictive accuracy of the models. The final model that was selected was a random effects model. The model included the following variables:

- Two dummy variables for levels 2 and 3 in each dimension.
- An ordinal variable D1 that represented the number of movements away from Full health beyond the first (ie it took on values ranging from 0 to 4).
- An ordinal variable I3 that represented the number of dimensions at level 3 beyond the first.
- The square of the I3 term to allow for non linearity in its association with the dependent variable.

- The square of I2, an ordinal variable that represented the number of dimensions at level 2 beyond the first.

The R^2 value of the final model was 0.38.

VALUE SET

The mathematical representation of the model for health state X is:

X = 1 - 0.146 MO2 - 0.558 MO3 - 0.175 SC2 - 0.471 SC3 - 0.140 UA2 - 0.374 UA3 - 0.173 PD2 - 0.537 PD3 - 0.156 AD2 - 0.450 AD3 + 0.140 D1 - 0.011 I2-square + 0.122 I3 + 0.015 I3-square.

REFERENCES

Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. Med Care 2005; 43(3): 203-220.

Zimbabwean TTO value set		**Example: the value for health state 21232**	
Full health (11111)	1	Full health	= 1
At least one 2 or 3 (constant)	- 0.100	Minus constant	- 0.100
At least one 3 (N3)	–	Minus N3	–
Mobility = 2	- 0.056	Minus MO level 2	- 0.056
Mobility = 3	- 0.204		
Self care = 2	- 0.092	Minus SC level 1	- 0.000
Self care = 3	- 0.231		
Usual activities = 2	- 0.043	Minus UA level 2	- 0.043
Usual activities = 3	- 0.135		
Pain/discomfort = 2	- 0.067		
Pain/discomfort = 3	- 0.302	Minus PD level 3	- 0.302
Anxiety/depression = 2	- 0.046	Minus AD level 2	- 0.046
Anxiety/depression = 3	- 0.173		
		State 21232	= 0.453

DATA COLLECTION

Data were collected in March 2000. 2488 residents of randomly selected small residential plots of land in Glenview, a suburb of Harare, were interviewed. The entrance criteria included completion of primary school education and a minimum age of 15. Of the 2488 residents that were contacted, 48 refused to participate. Therefore 2440 interviews were conducted, a response rate of 98%.

The interviews resulted in a sample in which women, younger people and those with a higher level of literacy were over-represented. There was a considerable interviewer effect.

EXCLUSION CRITERIA

56 respondents had incomplete data and were excluded. Replies from 201 respondents demonstrated inconsistency but were *included* in the analysis. After the exclusion of data from the 56 respondents with incomplete data, 2384 respondents remained for analysis.

SAMPLE CHARACTERISTICS

Of the 2384 respondents included in the analysis, 38% were male and 62% were female. The age distribution of the respondents was:

15-24 yrs	46%
25-34 yrs	33%
35-44 yrs	11%
45-54 yrs	7%
55+ yrs	3%

VALUATION METHODS

Each respondent was asked to do the following:

1. EQ-5D descriptive system and EQ VAS.
2. TTO exercise.

Each respondent valued 7 states which included 1 or 2 very mild, mild, moderate and severe states. All respondents also valued an 8th state: 33333. 11111 was valued, Unconscious and Dead were not valued (Dead was treated as an anchor point with a utility score of 0). A total of 38 states were valued.

The TTO valuations were transformed to lie on the interval [-1,1]. All states were calculated as t/10, where t is the number of years in 11111.

A split sample design was used in which 2/3 of the dataset was used to derive the model and 1/3 to validate the model.

A residual maximum likelihood linear mixed model was fitted to the data. Interviewer effect and subject nested within interviewer were fitted as random effects. The 3 levels of the 5 domains were fitted as fixed effects.

The chosen model was a main effects model.

The R^2 value of the model was 0.514.

VALUE SET

The mathematical representation of the model for health state X is:

X = 1 - 0.100 - 0.056 MO2 - 0.204 MO3 - 0.066 SC2 - 0.231 SC3 - 0.043 UA2 - 0.135 UA3 - 0.067 PD2 - 0.302 PD3 - 0.046 AD2 - 0.173 AD3

REFERENCES

Jelsma J, Hansen K, De Weerdt W, De Cock P, Kind P: How do Zimbabweans value health states? Popul Health Metr 2003; 1(1):11.

Annex 2:

Inventory of Visual Analogue Scale valuation surveys

Belgian VAS value set		**Example: the value for health state 21232**	
Full health (11111)	1	Full health	= 1
At least one 2 or 3 (constant)	- 0.152	Minus constant	- 0.152
At least one 3 (N3)	- 0.256	Minus N3	- 0.256
Mobility = 2	- 0.074	Minus MO level 2	- 0.074
Mobility = 3	- 0.148		
Self care = 2	- 0.083	Minus SC level 1	- 0.000
Self care = 3	- 0.166		
Usual activities = 2	- 0.031	Minus UA level 2	- 0.031
Usual activities = 3	- 0.062		
Pain/discomfort = 2	- 0.084		
Pain/discomfort = 3	- 0.168	Minus PD level 3	- 0.168
Anxiety/depression = 2	- 0.103	Minus AD level 2	- 0.103
Anxiety/depression = 3	- 0.206		
		State 21232	= 0.216

DATA COLLECTION

Data for the Belgian VAS valuation study were collected in the summer of 2001 in a postal survey. 2,754 Flemish language versions of the EQ-5D questionnaire were sent out. A reminder was sent two weeks after the initial mailing. The target population was respondents representative of the Flemish population of Belgium who were older than 18 years and of whom 50% were male. 967 questionnaires were returned. 245 were blank, so the response rate was 26.2%.

EXCLUSION CRITERIA

Data from 174 respondents was excluded from the sample based the following criteria:

a. < 3 states valued (n = 36).
b. all states valued the same (n = 0).
c. 11111 and/or Dead not valued (n = 126).
d. Dead ≥ 11111 (n = 12).

SAMPLE CHARACTERISTICS

Of the 548 respondents that were included in the final dataset, 44.9% were male and 55.1% were female. The mean age was 45.9 yrs.

VALUATION METHODS

Three different versions of the questionnaires were sent out and values for a total of 25 health states were obtained. 16 states had to be valued in each questionnaire using the standard 20cm vertical EQ-5D VAS. The duration of the time spent in the states was 1 year. The 25 states included in the survey were:

11111	12111	32223
11112	13311	32313
11113	21111	33321
11121	21232	33323
11122	22222	33333
11131	22233	Unconscious
11133	22323	Dead
11211	23232	
11312	32211	

Individuals' values for Dead were used to rescale their valuations to a scale with anchor points Full health = 1 and Dead = 0 as described above. A random effects model was used to derive an additive utility function. Disutilities (1-S) were used in the model estimations and 9 different models were tested. The resulting Belgian value set was based on a model including: a variable for each dimension, an intercept representing any deviation from Full health and an N3 dummy indicating whether 1 or more of the dimensions was on level 3.

The R^2 value of the model was not reported.

VALUE SET

The mathematical representation of the model for rescaled health state X is:

$X_{rescaled}$ = 1- 0.152 - 0.074 MO2 - 2 * 0.074 MO3 - 0.083 SC2 - 2 * 0.083 SC3 - 0.031 UA2 -2 * 0.031 UA3 - 0.084 PD2 - 2 * 0.084 PD3 - 0.103 AD2 - 2 * 0.103 AD3 - 0.256 N3

REFERENCES

Cleemput I. Economic evaluation in renal transplantation: outcome assessment and cost-utility of non-compliance. Acco, Leuven 2003.

Danish VAS value set		Example: the value for health state 21232	
Full health (11111)	1	Full health	= 1
At least one 2 or 3 (constant)	- 0.225	Minus constant	- 0.225
At least one 3 (N3)	–	Minus N3	–
Mobility = 2	- 0.126	Minus MO level 2	- 0.126
Mobility = 3	- 0.252		
Self care = 2	- 0.112	Minus SC level 1	- 0.000
Self care = 3	- 0.224		
Usual activities = 2	- 0.064	Minus UA level 2	- 0.064
Usual activities = 3	- 0.128		
Pain/discomfort = 2	- 0.078		
Pain/discomfort = 3	- 0.156	Minus PD level 3	- 0.156
Anxiety/depression = 2	- 0.091	Minus AD level 2	- 0.091
Anxiety/depression = 3	- 0.182		
		State 21232	= 0.338

DATA COLLECTION

The data for the Danish valuation study was collected in the winter of 1999/2000. A third party was commissioned to carry out the collection of the data from a representative sample of the Danish population. 6350 non-institutionalised persons aged 18 and over were contacted by telephone and asked whether they wanted to participate in the study. 1356 did not want to participate. Of the remaining 4996 people who were willing to participate, 1663 did not return the questionnaire. A total of 3331 (53%) completed questionnaires were returned of which 1686 included the standard EQ-5D VAS valuation exercise. 507 respondents were excluded from the dataset. Therefore the Danish VAS value set was based on data from 1179 respondents.

EXCLUSION CRITERIA

The exclusion criteria that were used were:

a. < 3 states valued (n=120).
b. all states valued the same (n=17).
c. 11111 and/or Dead not valued (n = 354).
d. Dead ≥ 11111 (n=16).

A total of 507 respondents were excluded from the dataset.

SAMPLE CHARACTERISTICS

Of the total number of respondents included in the final dataset, 42.1% of the respondents were male, and 57.1% were female. Compared to the Danish general population of 2000, women and respondents in the 30-59 years age group were over-represented. Age distribution of the respondents presented in 3 groups was as follows:

18-29 yrs	17.5%
30-59 yrs	61.7%
60+ yrs	20.9%

VALUATION METHODS

The 16 common core EQ-5D health states were included in the Danish VAS valuation study (including the state Unconscious). These states were divided over 2 pages, each containing 8 states. The states 11111, 33333 were included twice (once on both pages). In addition to the 16 states, respondents were asked to value Dead on both pages. The standard 20cm vertical EQ-5D VAS was used. The duration of the time spent in the states was 1 year.

Individuals' values for Dead were used to rescale their valuations to a scale with anchor points Full health = 1 and Dead = 0.

A random effects model was used to derive an additive utility function. Disutilities (1-S) were used in the model estimations and 12 different models were tested. The resulting Danish value set was based on a model including: a variable for each dimension and the intercept representing any deviation from Full health.

The R^2 value of the model was 0.82.

VALUE SET

The mathematical representation of the model for rescaled health state X is:

$X_{rescaled}$ = 1 – 0.225 – 0.126 MO2 – 2* 0.126 MO3 – 0.112 SC2 – 2*112 SC3 – 0.064 UA2 – 2*0.064 UA3 – 0.078 PD2 – 2* 0.078 PD3 – 0.091 AD2 – 2*0.091 AD3

A Danish value set that includes the N3 factor (used when any dimension is at level 3) is also available from the EuroQol Executive Office.

REFERENCES

Wittrup-Jensen KU, Lauridsen JT, Gudex C, Brooks R, Pedersen KM. Estimating Danish EQ-5D tariffs using TTO and VAS. In: Norinder A, Pedersen K, Roos P, editors. Proceedings of the 18th Plenary Meeting of the EuroQol Group. 2001. Copenhagen, Denmark. IHE, The Swedish Institute for Health Economics, 2002: 257-292.

European VAS value set*		Example: the value for health state 21232*	
Full health (11111)	1	Full health	= 1
At least one 2 or 3 (N2)	-0.1279	Minus N2	- 0.1279
At least one 3 (N3)	-0.2288	Minus N3	- 0.2288
Mobility = 2	-0.0659	Minus MO level 2	- 0.0659
Mobility = 3	-0.1829		
Self care = 2	-0.1173	Minus SC level 1	- 0.0000
Self care = 3	-0.1559		
Usual activities = 2	-0.0264	Minus UA level 2	- 0.0264
Usual activities = 3	-0.0860		
Pain/discomfort = 2	-0.0930		
Pain/discomfort = 3	-0.1637	Minus PD level 3	- 0.1637
Anxiety/depression = 2	-0.0891	Minus AD level 2	- 0.0891
Anxiety/depression = 3	-0.1290		
		State 21232*	= 0.2982

* The values have been rescaled to a scale with 11111 = 1 and Dead = 0. This was performed using the mean value of Dead.

DATA COLLECTION

The European VAS value set was constructed using data from 11 valuation studies in 6 countries: Finland (1), Germany (3), The Netherlands (1), Spain (3), Sweden (1) and the UK (2). Although not all of the studies included were representative of the country in which they were carried out and data from a number of other European countries was not available, there is sufficient data from different European regions to make the European VAS dataset moderately representative for Europe.

The 11 studies were carried out in the period from January 1991 - March 1998. The survey settings varied between the studies so both postal surveys and interview-based surveys were included in the European dataset. The pooled data set consisted of valuations from 8709 respondents, the response rates varied between studies. 1839 respondents were excluded from the dataset. Therefore the European VAS value set was based on data from 6870 respondents.

EXCLUSION CRITERIA

A total of 1839 respondents were excluded from the analyses, using the same exclusion criteria for data from all studies. These exclusion criteria were:

a. All states valued the same or fewer than 3 states valued (n=873).
b. More than 3 inconsistencies (n=966).

Because rescaling was performed on aggregate data, respondents who had missing values for 11111 and/or Dead or valued Dead ≥ 11111 were not excluded.

SAMPLE CHARACTERISTICS

Of the total number of respondents included in the final dataset, 46.1% were male, and 53.9 % were female. Age distribution of the respondents presented in 3 groups was as follows:

≤ 44 yrs	50.6%
45-64 yrs	29.2%
≥ 65 yrs	20.2%

VALUATION METHODS

Analyses were carried out on data from a total of 44 different health states. The number of health states per study ranged from 11 to 43 and they varied between studies. The standard 20cm vertical EQ-5D VAS was used in all studies. The duration of the health states that people were asked to value were either 1 year or 10 years, depending on the study. The studies with a ranked interview based study design (RID) were the studies with a duration of 10 years. A dummy variable for RID was included in the model.

Health states included in the European VAS study

11111	12211	21323	23321
11112	12222	22112	32211
11113	12223	22121	32223
11121	13212	22122	32232
11122	13311	22222	32313
11131	13332	22233	32331
11133	21111	22322	33212
11211	21133	22323	33232
11312	21222	22331	33321
12111	21232	23232	33323
12121	21312	23313	33333

Rescaling was performed on aggregate data. Rescaled value sets were calculated for both the mean and the median value for Dead with the following equation:

$X_{rescaled} = 100 * (X_{raw} – Dead_{mean\ or\ median}) / (11111_{raw} – Dead_{mean\ or\ median})$

where the mean value of Dead = 10 and the median value of Dead = 2.

Note that the estimated value for Full health (11111) was used in the formula and not the observed mean or median value for Full health.

A number of different models were tried. The model that was selected was a multi-level random effects model. The 3 levels in the model were the evaluations (level 1) nested within the respondents (level 2) nested within the studies (level3). The model included the following parameters:

- A dummy was included for ranked interview-based study designs (RID).
- A dummy N2 for any move away from Full health.
- A dummy N3 for 1 or more dimensions on level 3.
- Interaction terms of RID and N2 (RN2) and RID and N3 (RN3).
- 10 dummy variables, 2 for each dimension. 1 for level 2 and 1 for level 3.

The R^2 value of the model was 0.745.

VALUE SET

The complete model was of the form:

$X_{raw} = C + c_1RID + c_2N2 + c_3N3 + c_4RN2 + c_5RN3 + c_6MO2 + c_7MO3 + c_8SC2 + c_9SC3 + c_{10}UA2 + c_{11}UA3 + c_{12}PD2 + c_{13}PD3 + c_{14}AD2 + c_{15}AD3$

After the model was estimated the parameter estimates of C and c_1 were added, as were the estimates of c_2 and c_4 and those of c_3 and c_5. This reduced the number of parameters from 16 to 13.

$X_{raw} = 97.66 - 11.21\ N2 - 20.06\ N3 - 5.78\ MO2 - 16.03\ MO3 - 10.28\ SC2 - 13.67\ SC3 - 2.31\ UA2 - 7.54\ UA3 - 8.15\ PD2 - 14.35\ PD3 - 7.81\ AD2 - 11.31\ AD3$

The raw values X_{raw} can then be rescaled with the equation mentioned above. Alternatively, the model itself can be rescaled. Using the mean value for Dead the mathematical representation of the rescaled model is:

$X_{rescaled} = 1 - 0.1279\ N2 - 0.2288\ N3 - 0.0659\ MO2 - 0.1829\ MO3 - 0.1173\ SC2 - 0.1559\ SC3 - 0.0264\ UA2 - 0.0860\ UA3 - 0.0930\ PD2 - 0.1637\ PD3 - 0.0891\ AD2 - 0.1290\ AD3$

REFERENCES

Greiner W, Weijnen T, Nieuwenhuizen M, et al. A single European currency for EQ-5D health states. Results from a six country study. Eur J Health Econ 2003; 4(3):222-231.

Finnish VAS value set		**Example: the value for health state 21232**	
Full health (11111)	1	Full health	= 1
At least one 2 or 3 (constant)	- 0.158	Minus constant	- 0.158
At least one 3 (N3)	–	Minus N3	–
Mobility = 2	- 0.058	Minus MO level 2	- 0.058
Mobility = 3	- 0.230		
Self care = 2	- 0.098	Minus SC level 1	- 0.000
Self care = 3	- 0.143		
Usual activities = 2	- 0.047	Minus UA level 2	- 0.047
Usual activities = 3	- 0.131		
Pain/discomfort = 2	- 0.111		
Pain/discomfort = 3	- 0.153	Minus PD level 3	- 0.153
Anxiety/depression = 2	- 0.160	Minus AD level 2	- 0.160
Anxiety/depression = 3	- 0.196		
		State 21232	= 0.424

DATA COLLECTION

The Finnish data were collected in November 1992. 4000 randomly selected persons over 16 years of age received the questionnaire by mail. Two reminders were mailed approximately two weeks apart. The sample was divided over 17 sub-samples who received a different questionnaire. The study is based on 11 of the sub-samples (2530 persons). Because the elderly population was over-sampled, a weighting for age and sex was used. Of the 2530 persons that were contacted, 1634 persons returned the questionnaire resulting in a response rate of 64.5 %.

EXCLUSION CRITERIA

Data from 362 respondents was excluded due to the following exclusion criteria:

a. < 3 states valued (n=192).
b. all states valued the same (n=53).
c. 11111 not valued (n=72).
d. Dead ≥ 11111 (n=41).

Different exclusion strategies were tested with respect to the handling of inconsistencies. It was found that including respondents with 3 or less inconsistencies did not significantly influence the modeling. Therefore the models are based on a sub-sample containing only persons with 3 or fewer inconsistent values, which resulted in an additional 344 exclusions. The total number of respondents that were excluded from the Finnish dataset was 706 and hence the modeling was done on data from 928 persons.

SAMPLE CHARACTERISTICS

Of the total number of respondents included in the final dataset, 46.2 % were male, 53.8 % were female and the mean age was 42.5 yrs.

VALUATION METHODS

Each respondent valued a subset of the 46 states with a duration of 1 year on the standard 20cm VAS. The 46 health states that were included in the study were:

11111	12122	22111	22331	33122
11112	12211	22112	23232	33321
11121	12222	22121	23321	33323
11122	12223	22122	31231	33333
11211	13233	22211	31323	Dead
11233	13332	22212	32132	Unconscious
11312	21111	22221	32211	
12111	21133	22222	32223	
12112	21222	22233	32232	
12121	21232	22323	32313	

The respondents were asked to value the states 11111, 33333 and Dead twice.

An OLS model was used to derive an additive utility function. 8 different models were tested. The models utilised both individual data and logistic transforms of individual data.

The final model was an additive model based on the logit transformations of the individual data. This model included separate dummy variables for levels 2 and 3 on each dimension. In addition, dummy variables were included in the model to indicate any move from Full health, the value of Dead, and the value of Unconscious.

The R^2 value of the model was 0.743.

Rescaling to 11111 = 1 and Dead = 0 was performed on the parameter estimates of the anchor points (i.e. on the aggregate data). In the model, a logit transformation was used on the individual data with the purpose of normalising the preference values. The formula used is: $g = ln\ (M / (1-M))$, where M stands for a health state and g for the transformed state.

VALUE SET

The complete model was of the form:

$X_{rescaled}$ = 1.000 – 0.158 – 0.058 MO2 – 0.230 MO3 – 0.098 SC2 – 0.143 SC3 – 0.047 UA2 – 0.131 UA3 – 0.111 PD2 – 0.153 PD3 – 0.160 AD2 – 0.196 AD3 – 0.869 Unconscious - 0.842 Dead

REFERENCES

Ohinmaa A, Eija H, Sintonen H. Modeling EuroQol values of Finnish adult population. In: Badia X, Herdman M, Segura A, editors. EuroQol Plenary Meeting Barcelona 1995. Discussion Papers. Institut Universitari de Salut Publica de Catalunya, 1996; 67-76.

Ohinmaa A, Sintonen H. Inconsistencies and modeling of the Finnish EuroQol (EQ-5D) preference values. In: Greiner W, J-M. Graf v.d. Schulenburg, Piercy J, editors. EuroQol Plenary Meeting, 1-2 October 1998. Discussion papers. Centre for Health Economics and Health Systems Research, University of Hannover, Germany. Uni-Verlag Witte, 1999; 57-74.

German VAS value set*		Example: the value for health state 21232	
Full health (11111)	1	Full health	= 1
At least one 2 or 3 (constant)	0.9256	Times constant	0.9256
At least one 3 (N3)	–		
Mobility = 2	0.9447	Times MO level 2	0.9447
Mobility = 3	0.3927		
Self care = 2	0.8080	Times SC level 1	1.0000
Self care = 3	0.4702		
Usual activities = 2	0.8803	Times UA level 2	0.8803
Usual activities = 3	0.5538		
Pain/discomfort = 2	0.9745		
Pain/discomfort = 3	0.4671	Times PD level 3	0.4671
Anxiety/depression = 2	0.8174	Times AD level 2	0.8174
Anxiety/depression = 3	0.4682		
		State 21232	= 0.2939

* The model is a multiplicative model. This implies that when any of the dimensions are at level 1 the appropriate coefficient for that level is 1.

DATA COLLECTION

The second German valuation study was a combined VAS / TTO study. Data were collected in an interviewer setting from October 1997 to March 1998. 380 addresses were randomly selected from a telephone directory and distributed to interviewers. 41 respondents could not be contacted and 5 interviews had to be aborted because the respondent could not master the task.

EXCLUSION CRITERIA

Extreme values for health states were excluded: 45 values from 20 different health states deviated so much from the median scores for these states that they were considered to be faulty and were excluded from the analysis.

SAMPLE CHARACTERISTICS

Of the 339 respondents included in the analyses 55.3% were male and 44.7% were female. The age distribution of the respondents was:

15-25 yrs	7.7%
25-45 yrs	31.1%
45-65 yrs	40.2%
65+ yrs	21.0%

VALUATION METHODS

The interviews consisted of a ranking exercise followed by the valuation of health states on the VAS, the EQ-5D self-classifier and finally the TTO exercise. In the VAS valuation exercise, the standard 20cm vertical VAS was used and the duration of the health states that people were asked to value was 10 years in each health state (analogous to the time frame used in the standard TTO exercise).

A total of 36 health states were valued. Each respondent valued 14 states. States 11111 and 33333 were valued once; Unconscious and Dead were not valued.
The 36 states that were valued were:

11111	12111	21232	32211
11112	12121	21323	32223
11113	12211	22112	32232
11121	12222	22122	32331
11122	12223	22222	33212
11131	13212	22323	33232
11133	21111	22331	33321
11211	21133	23232	33323
11312	21222	23321	33333

No rescaling of the value set was reported.

OLS regression was used to derive a final model. The model is multiplicative. It consists of 1 variable for each of the 5 dimensions (0=no problems, 1=some problems, 2=extreme problems), plus a dummy variable for each of the 5 dimensions (1=extreme problems, 0=otherwise), plus a constant.

The R^2 value of the model was 0.72.

VALUE SET

The model is a multiplicative model in which the variables appear as exponents:

$$X = \beta_0 * \beta_{12}^{b12} * \beta_{13}^{b13} * \beta_{22}^{b22} * \beta_{23}^{b23} * \beta_{32}^{b32} * \beta_{33}^{b33} * \beta_{42}^{b42} * \beta_{43}^{b43} * \beta_{52}^{b52} * \beta_{53}^{b53}$$

β_0 = a constant associated with any move from Full health = 0.9256
β_{xy} = instalment (x = EQ-5D dimension, y = level)
b_{xy} = dummy variable (x = EQ-5D dimension, y = level)
b_{x2} = 1 if answering level is 2
b_{x3} = 1 if answering level is 3

The coefficients β_{xy} that were derived for the model are:

dimension	*level 1*	*level 2*	*level 3*
MO	1.00	0.9447	0.3927
SC	1.00	0.8080	0.4702
UA	1.00	0.8803	0.5538
PD	1.00	0.9745	0.4671
AD	1.00	0.8174	0.4682

Example: The value of health state 21232.

$$X_{21232} = \beta_0 * \beta_{12} * 1.00 * \beta_{32} * \beta_{43} * \beta_{52}$$
$$= 0.9256 * 0.9447 * 1.00 * 0.8803 * 0.4671 * 0.8174$$
$$= 0.2939$$

Since the survey did not include valuations for Dead, the values derived from the data cannot be rescaled to a scale where Dead = 0.

REFERENCES

Claes C, Greiner W, Uber A, Graf von der Schulenburg JM. An interview-based comparison of the TTO and VAS values given to EuroQol states of health by the general German population. In: Greiner W, J-M. Graf v.d. Schulenburg, Piercy J, editors. EuroQol Plenary Meeting, 1-2 October 1998. Discussion papers. Centre for Health Economics and Health Systems Research, University of Hannover, Germany. Uni-Verlag Witte, 1999; 13-39.

New Zealand VAS value set		Example: the value for health state 21232	
Full health (11111)	1	Full health	= 1
At least one 2 or 3 (constant)	- 0.2041	Minus constant	- 0.2041
At least one 3 (N3)	- 0.2165	Minus N3	- 0.2165
Mobility = 2	- 0.0753	Minus MO level 2	- 0.0753
Mobility = 3	- 0.1506		
Self care = 2	- 0.0714	Minus SC level 1	- 0.0000
Self care = 3	- 0.1428		
Usual activities = 2	- 0.0136	Minus UA level 2	- 0.0136
Usual activities = 3	- 0.0272		
Pain/discomfort = 2	- 0.0798		
Pain/discomfort = 3	- 0.1596	Minus PD level 3	- 0.1596
Anxiety/depression = 2	- 0.0920	Minus AD level 2	- 0.0920
Anxiety/depression = 3	- 0.1840		
		State 21232	= 0.2389

DATA COLLECTION

Questionnaires were mailed in January 1999 to a non-stratified random sample of 3000 members of the general public over 18 years. A reminder and duplicate questionnaire were sent two weeks later. The sample was obtained from the electoral register. An initial check confirmed the sample to be representative of the New Zealand population in terms of age, sex and Maori/non-Maori composition. 259 questionnaires were returned as deceased/incorrect addresses. 1360 completed questionnaires were returned, a response rate of 50%.

EXCLUSION CRITERIA

Of the 1360 returned questionnaires, 441 were excluded because of the following exclusion criteria:

a. < 3 states valued (n = 143).
b. all states valued the same (n = 46).
c. 11111 and/or Dead not valued (n = 237).
d. Dead > 11111 (n = 15).

A key feature of the New Zealand research was experimenting with alternative exclusion criteria for logical inconsistencies:

a) no exclusions.
b) include only those with 0-1 pairwise logical inconsistencies.

The two different exclusion strategies for inconsistencies resulted in 2 different sample sizes: n = 919 for the full sample and n = 396 for the sub-sample including respondents with 0 or 1 inconsistencies only.

SAMPLE CHARACTERISTICS

Age and gender for the full sample and sub-sample were distributed as follows:

	full sample (n = 919)	sub-sample (n = 396)
18-29 yrs	13.3 %	17.4 %
30-49 yrs	43.3 %	44.4 %
50-69 yrs	33.2 %	29.8 %
70+ yrs	9.4 %	8.1 %
Male	43.5 %	45.0 %
Female	56.1 %	54.3 %

VALUATION METHODS

Three versions of the questionnaire were sent out, containing 3 different sets of health states. One third of the sample each received one version of the questionnaire:
Version A consisted of the common core states, versions B and C consisted of a sub-set of the common core states plus a selection of further states. In total, 24 EQ-5D states, plus Dead and Unconscious, were valued using the standard 20 cm vertical VAS. Respondents were asked to value spending 1 year in each health state. 11111, 33333 and Dead were valued twice in each of the 3 versions of the questionnaire.

Health states valued in the New Zealand VAS valuation study:

11111	11211	22222	32313
11112	11312	22233	33321
11113	12111	22323	33323
11121	13311	23232	33333
11122	21111	32211	Unconscious
11131	21232	32223	Dead
11133			

Individuals' values for Dead used to rescale their valuations with the following formula: $X_{rescaled} = (X_{raw} - Dead_{raw}) / (11111_{raw} - Dead_{raw})$

The model used was a random effects model. Nine specifications were considered. Selection criteria used for the model specifications were that the value set estimated from a model must be logically consistent, and parsimony. Which specification was 'best' was dependent on the data set – with different specifications appropriate for the 2 data sets. Models were estimated using disutilities (1-S). The final model was a main effects model based on the n=396 sub-sample. It included a constant that represents any deviation from Full health plus one variable for each EQ-5D dimension, with level 1 = 0; level 2 = 1; level 3 = 2, plus an N3 variable that represents whether any of the dimensions are at level 3.

The R^2 value of the final model was 0.70.

VALUE SET

The complete model was of the form:

$X_{rescaled} = 1 – 0.2041 – 0.0753\ MO2 – 2*0.0753\ MO3 – 0.0714\ SC2 – 2*0.0714\ SC3 – 0.0136\ UA2 – 2*0.0136\ UA3 – 0.0798\ PD2 – 2*0.0798\ PD3 – 0.0920\ AD2 – 2*0.0920\ AD3 – 0.2165\ N3$

In their paper, the researchers who derived the value set recommend that analysts proposing to use the values reported here test the effects of using this value set and other value sets on their results in a sensitivity analysis.

REFERENCES

Devlin NJ, Hansen P, Kind P, Williams, A. Logical inconsistencies in survey respondents' health state valuations - a methodological challenge for estimating social tariffs. Health Econ 2003; 12(7):529-544.

Slovenia VAS value set		**Example: the value for health state 21232**	
Full health (11111)	1	Full health	= 1
At least one 2 or 3 (constant)	- 0.128	Minus constant	- 0.128
At least one 3 (N3)	–	Minus N3	–
Mobility = 2	- 0.206	Minus MO level 2	- 0.206
Mobility = 3	- 0.412		
Self care = 2	- 0.093	Minus SC level 1	- 0.000
Self care = 3	- 0.186		
Usual activities = 2	- 0.054	Minus UA level 2	- 0.054
Usual activities = 3	- 0.108		
Pain/discomfort = 2	- 0.111		
Pain/discomfort = 3	- 0.222	Minus PD level 3	- 0.222
Anxiety/depression = 2	- 0.093	Minus AD level 2	- 0.093
Anxiety/depression = 3	- 0.186		
		State 21232	= 0.297

DATA COLLECTION

Data collection for the Slovenian valuation study was gathered in April and May 2000. Questionnaires were sent out to 3000 randomly selected respondents with a minimum age of 18 years. No reminder was sent and 792 questionnaires were returned. Of these, 38 were empty and 21 did not include any valuations, therefore a usable response rate of 24.4%.

EXCLUSION CRITERIA

All respondents with 1 or more logical inconsistencies were excluded from the dataset (n = 363). Therefore the sample on which the analyses were carried out consisted of data from 370 respondents.

SAMPLE CHARACTERISTICS

Of the 370 respondents included in the analysis, 40.3% were male and 59.7% were female. Age of the respondents was distributed as follows:

18-25 yrs	23.0%
26-35 yrs	23.8%
36-45 yrs	19.5%
46-55 yrs	13.2%
56-65 yrs	10.8%
66+ yrs	9.7%

VALUATION METHODS

Respondents were asked to value spending 1 year in each of the 16 common core states on the standard 20cm vertical VAS plus the value of Dead. The states 11111, 33333 and Dead were valued twice.

Individuals' values for Dead were used to rescale their valuations to a Full health =1, Dead = 0 scale with the following formula:
$X_{rescaled} = (X_{raw} – Dead_{raw}) / (11111_{raw} – Dead_{raw})$

An OLS model was used to derive an additive utility function. 13 different models were tested.

The final model included a variable for each dimension (level 1 = 1, level 2 = 2 and level 3 = 3) and a constant representing the value of Full health.

The R^2 value of the model was 0.648.

VALUE SET

In order to make the model comparable to the other studies presented in this booklet, two adjustments were made. Firstly the values for the variables of the dimensions were changed from (1, 2, 3) to (0, 1, 2) and secondly the constant was changed from 'constant' to '1 - adjusted constant'. Note that neither of these modifications changes the model, just the presentation of the model. After the modifications the model has the form:

X = 1 – 0.128 – 0.206 MO2 – 2 * 0.206 MO3 – 0.093 SC2 – 2 * 0.093 SC3 – 0.054 UA2 – 2 * 0.054 UA3 – 0.111 PD2 – 2 * 0.111 PD3 – 0.093 AD2 – 2 * 0.093 AD3

REFERENCES

Prevolnik Rupel V, Rebolj M. The Slovenian VAS Tariff based on valuations of EQ-5D health states from the general population. In: Cabasés JM, Gaminde I, editors. Proceedings of the 17th Plenary Meeting of the EuroQol Group. Universidad Pública de Navarra 2001; 23-47.

Spanish VAS value set		Example: the value for health state 21232	
Full health (11111)	1	Full health	= 1
At least one 2 or 3 (constant)	- 0.1502	Minus constant	- 0.1502
At least one 3 (N3)	- 0.2119	Minus N3	- 0.2119
Mobility = 2	- 0.0897	Minus MO level 2	- 0.0897
Mobility = 3	- 0.1794		
Self care = 2	- 0.1012	Minus SC level 1	- 0.0000
Self care = 3	- 0.2024		
Usual activities = 2	- 0.0551	Minus UA level 2	- 0.0551
Usual activities = 3	- 0.1102		
Pain/discomfort = 2	- 0.0596		
Pain/discomfort = 3	- 0.1192	Minus PD level 3	- 0.1192
Anxiety/depression = 2	- 0.0512	Minus AD level 2	- 0.0512
Anxiety/depression = 3	- 0.1024		
		State 21232	= 0.3227

DATA COLLECTION

Data for this Catalan valuation study were collected October-December 1996. Two trained interviewers recruited a random sample of 300 individuals, both patients and non-patients attending a primary care setting in Cornellà de Llobregat (Catalonia). They were selected via age and sex quota sampling so as to be representative of the Catalan general population according to figures in the 1991 Catalan census.

EXCLUSION CRITERIA

6 questionnaires were excluded from the analysis because of inconsistent responses.

SAMPLE CHARACTERISTICS

Of the 294 respondents included in the analysis, 45.9% were male and 54.1% were female. Age of the respondents was distributed as follows:

< 24 yrs	16.0%
25-34 yrs	18.4%
35-44 yrs	16.7%
45-54 yrs	15.3%
55-64 yrs	15.0%
65+ yrs	18.7%

VALUATION METHODS

The interviews, in which the interviewers had a passive role, consisted of 4 stages:

1. The EQ-5D descriptive system and EQ VAS.
2. Ranking the 13 health states + Unconscious + Dead.
3. Valuation of the health states using the VAS.
4. Valuation of the health states using the TTO.

Stages 3 and 4 were administered in random order.

During the VAS valuation part of the interviews, values for a total of 45 health states were elicited. Each respondent valued a subset of 15 states with a duration of 10 years on the standard 20cm VAS. The health states were randomly chosen from the pool of 45 states according to the following rules:

- 11111, 33333, Dead and Unconscious.
- plus 2 from 5 'very mild' states (11112, 11121, etc).
- plus 3 from 12 'mild' states (11122, 11131, etc).
- plus 3 from 12 'moderate' states (13212, 32331, etc).
- plus 3 from 12 'severe' states (33232, 23232, etc).

The 45 health states valued were:

11111	12111	21133	22233	32313
11112	12121	21222	22323	32331
11113	12211	21232	22331	33212
11121	12222	21312	23232	33232
11122	12223	21323	23313	33321
11131	13212	22112	23321	33323
11133	13311	22121	32211	33333
11211	13332	22122	32223	Unconscious
11312	21111	22222	32232	Dead

2 sets of analyses were carried out, 1 in which individuals' values for Dead were used to rescale their valuations and 1 in which the rescaling was performed on aggregate data using the median values of the raw states. The formula used in both cases was:

$$X_{rescaled} = (X_{raw} - Dead_{raw}) / (11111_{raw} - Dead_{raw})$$

The effect of a number of logarithmic transformations on the data were also tested. Analyses at individual level were performed with Ordinary Least Squares, analyses on the aggregate data with Weighted Least Means. Eleven specifications were considered. The best model was selected on the basis of goodness of fit, parsimony and inconsistency. The model that was selected on these 3 criteria was one where rescaling was performed on the median values of the aggregate data. Logarithmic transformations were not used to obtain the final value set.

The R^2 value of the model was 0.969.

VALUE SET

The complete mathematical description of the model is:

$X_{rescaled}$ = 1 – 0.1502 – 0.0897 MO2 – 2 * 0.0897 MO3 – 0.1012 SC2 – 2 * 0.1012 SC3 – 0.0551 UA2 – 2 * 0.0551 UA3 – 0.0596 PD2 – 2 * 0.0596 PD3 –0.0512 AD2 – 2 * 0.0512 AD3 – 0.2119 N3

REFERENCES

Badia X, Roset M, Monserrat S, Herdman M. The Spanish VAS tariff based on valuation of EQ-5D health states from the general population. In: Rabin RE et al, editors. EuroQol Plenary meeting Rotterdam 1997, 2-3 October. Discussion papers. Centre for Health Policy & Law, Erasmus University, Rotterdam, 1998; 93-114.

UK VAS value set		**Example: the value for health state 21232**	
Full health (11111)	1	Full health	= 1
At least one 2 or 3 (constant)	- 0.155	Minus constant	- 0.155
At least one 3 (N3)	- 0.215	Minus N3	- 0.215
Mobility = 2	- 0.071	Minus MO level 2	- 0.071
Mobility = 3	- 0.182		
Self care = 2	- 0.093	Minus SC level 1	- 0.000
Self care = 3	- 0.145		
Usual activities = 2	- 0.031	Minus UA level 2	- 0.031
Usual activities = 3	- 0.081		
Pain/discomfort = 2	- 0.084		
Pain/discomfort = 3	- 0.171	Minus PD level 3	- 0.171
Anxiety/depression = 2	- 0.063	Minus AD level 2	- 0.063
Anxiety/depression = 3	- 0.124		
		State 21232	= 0.294

DATA COLLECTION

Data were collected in the period from August - December 1993. The sample was drawn from the national postcode address file (SCPR) and was representative for the non-institutionalised adult population of England, Scotland and Wales. 6080 addresses were drawn. 3395 (56%) were willing to be interviewed.

EXCLUSION CRITERIA

Respondents were excluded from the dataset according to the following criteria for both VAS and TTO valuations:

a. < 3 states valued.
b. All states valued the same.
c. 11111 and/or Dead not valued in the VAS valuation task.
d. Dead ≥ 11111 in the VAS valuation task.
e. Dead > all states other than 11111.
f. All data from one of the interviewers.
g. Any state missing from either VAS or TTO.

The total number of exclusions was 398. Therefore data from 2997 respondents was included in the analysis.

SAMPLE CHARACTERISTICS

Of the 2997 respondents included in the analysis, 43% were male and 57% were female. The age distribution of the respondents was:

18-24 yrs	9%
25-34 yrs	22%
35-49 yrs	25%
50-64 yrs	21%
65+ yrs	24%

VALUATION METHODS

Each respondent was asked to do the following:

1. EQ-5D descriptive system and EQ VAS.
2. Ranking exercise.
3. Valuation of the ranked health states on a VAS.
4. TTO exercise.

During the VAS valuation part of the interviews, values for a total of 45 health states were obtained. Each respondent valued a subset of 15 states with a duration of 10 years on the standard 20cm VAS:

- 11111, 33333, Unconscious and Dead.
- 2 from 5 'very mild' states (11112, 11121, etc).
- 3 from 12 'mild' states (11122, 11131, etc).
- 3 from 12 'moderate' states (13212, 32331, etc).
- 3 from 12 'severe' states (33232, 23232, etc).

The 45 health states that were valued were:

11111	12222	22121	32313
11112	12223	22122	32331
11113	13212	22222	33212
11121	13311	22233	33232
11122	13332	22323	33321
11131	21111	22331	33323
11133	21133	23232	33333
11211	21222	23313	Unconscious
11312	21232	23321	Dead
12111	21312	32211	
12121	21323	32223	
12211	22112	32232	

A generalised least-squares regression technique was used. The utility function was additive. Disutilities (1-S) were used in the model estimations. Two dummy variables were used for levels 2 and 3 in each dimension in order to allow for different utility increments between levels 1 and 2 as compared to the increment between levels 2 and 3.

Additional interaction terms between dummy variables across dimensions were used in order to measure 'extra' disutility when reporting problems on multiple dimensions. Individual-level (rather than aggregate-level) data analysis was performed. A random effects specification was used to address the problem that the scores for health states generated by an individual were related.

The R^2 value of the model was 0.47.

VALUE SET

The mathematical representation of the model for health state X is:

X = 1 – 0.155 – 0.071 MO2 – 0.182 MO3 – 0.093 SC2 – 0.145 SC3 – 0.031 UA2 – 0.081 UA3 – 0.084 PD2 – 0.171 PD3 – 0.063 AD2 – 0.124 AD3 – 0.215 N3

REFERENCES

MVH Group. The Measurement and Valuation of Health. Final report on the modeling of valuation tariffs. York Centre for Health Economics, 1995.