# Simultaneous valuation of multiple instruments: A new approach to mapping

Donna Rowen[a]*, John Brazier[a], Aki Tsuchiya[a,b] and Mónica Hernandez[b]

**Background:** Previous methods of empirical mapping involve using regressions on patient or general population self-report data from datasets involving 2 or more instruments. However this approach relies on overlap in the descriptive systems of the measures, but key dimensions may not be present in both measures. Furthermore this assumes it is appropriate to use different instruments on the same population, which may not be the case for all patient groups. The aim of this study is to develop a new method of mapping using general population preferences for hypothetical health states defined by the descriptive systems of different measures.

**Methods:** The study is based on primary data collected by valuation interviews involving 13 health states defined by each of 6 instruments: EQ-5D (generic), SF-6D (generic), HUI2 (generic for children), AQL-5D (asthma specific), OPUS (social care specific), ICECAP (capabilities). Each interview involves 3 ranking and visual analogue (VAS) tasks with states from 3 different instruments where each task involves the simultaneous valuation of multiple instruments. The study includes 13 health and well-being states for each instrument that reflect a range of health state values according to the published health state values for each instrument and each health state is valued approximately 75-100 times. Descriptive statistics for health state values using VAS data are presented and compared to published value sets including the UK EQ-5D VAS tariff.

**Results:** The sample consists of 501 members of the UK general population with a reasonable spread of background characteristics (response rate=55%). The study achieved a completion rate of 99% for all states included in the rank and rating tasks and 94.8% of respondents have complete responses. Mean adjusted VAS values for the worst states defined by the descriptive systems range from 0.075 (SD=0.228) for EQ-5D to 0.289 (SD=0.816) for AQL-5D, best EQ-5D and OPUS states have values of 0.91 and 0.90 respectively. As expected, adjusted VAS health state values were different from the current value sets (since they largely used different valuation methods) and adjusted VAS values have a smaller range and spread than those in published value sets (e.g. EQ-5D UK TTO).

**Conclusions:** This pilot study shows the feasibility of a new method of mapping between measures using general population preferences rather than statistical association between them that is better able to take advantage of diversity in descriptive systems. This approach has an important role in evidence synthesis and cross programme comparisons in studies using different instruments.

**Key words:** Preference-based measures of health; quality of life; mapping; Visual Analogue Scale

[a] Health Economics and Decision Science, University of Sheffield, United Kingdom

[b] Department of Economics, University of Sheffield, United Kingdom

* Correspondence to: Donna Rowen, Health Economics and Decision Science, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA, UK. Telephone: +44114 222 0728. Fax: +44114 272 4095. Email: d.rowen@sheffield.ac.uk

## 1. Introduction

Economic evaluation using preference-based measures of health to generate Quality Adjusted Life Years (QALYs) is being increasingly used to inform health policy. The QALY measure combines both quantity and quality of life into a single measure and quality of life is measured using a preference-based measure of health. Recent years has seen the increasing proliferation of preference-based measures of health. These fall into three categories:

1) Generic measures for adults such as the EQ-5D (Dolan, 1997), the HUI3 (Feeny et al, 2002), the QWB scale (Kaplan and Anderson, 1988) and the SF-6D (Brazier et al, 2002)
2) Measures designed for specific groups, such as those for older people (Coast et al, 2006), children (Torrance et al, 1996; Stevens, 2008) and social care (Ryan et al, 2006).
3) Condition-specific measures designed for specific medical conditions, such as those for asthma (Revicki et al, 1998; Yang et al, 2007) and urinary incontinence (Yang et al, 2008).

If all of these measures are preference-based, measured on an interval scale, with the upper anchor at full health (=1) and the lower anchor at 0 (assuming it is equivalent to dead) then theoretically all instruments should be comparable to each other, where the value of a health state for a patient is identical regardless of the instrument used. However this is not true in practice. The increasing number of generic and condition specific preference-based measures of health have been shown to generate different scores in the same population (Brazier et al, 2002; Longworth and Bryan, 2003; O'Brien et al, 2003). This can be attributed to differences in their descriptive systems and valuation methods (Brazier et al, 2004; Tsuchiya et al, 2006). This creates a major problem for researchers and policy makers undertaking evidence synthesis and cross programme comparisons across studies using different instruments.

One solution to cross programme comparison is to use one generic preference-based measure in all economic evaluation (Dowie, 2002). However, this is not possible for all groups of patients, for example children require special consideration regarding language comprehension and development (Stevens, 2008), and there may be special considerations for very elderly people (Coast et al, 2006) and those needing social care (Ryan et al., 2006). Furthermore generic measures of health are inappropriate or insensitive for many medical conditions (for example Barton et al, 2004; Espallargues et al, 2005; Harper et al, 1997; Kobelt et al, 1999). Even if these arguments are not accepted, the fact remains that many different measures including more than one generic measure have already been developed and used and this will continue where there is no common agreement on which measure to use.

Previous attempts at mapping between instruments have involved existing datasets that have two or more measures used alongside each other, where regression analysis is used to estimate a statistical relationship between the indices generated by the measures or their descriptive systems (for example, Brazier et al, 2005; Franks et al, 2004; Gray et al, 2006; Nichol et al, 2001; Tsuchiya et al, 2002). This method assumes it is appropriate to use different instruments on the same population, which as described above may not be the case for all patient groups and all conditions. This approach also relies on a degree of overlap in the descriptive systems of the measures, but key dimensions may not be present in both measures. Therefore trying to map between measures in this way seems to miss the point. Instead what is needed is a means of relating the responses on one measure to another using a common metric and preserving the advantages that the descriptive system may bring.

One possibility is to do time trade-off (TTO) or standard gamble (SG) on multiple instruments in one sitting, for example see Tsuchiya et al. (2006). However, given limited resources and increasing interest in using ordinal methods a ranking task was used alongside a visual analogue scale (VAS) in our current study. In addition, strictly speaking, TTO and SG are valuation methods where respondents deal with one state at a time, and therefore there is no direct head-to-head comparison of states from different instruments. In this respect, ranking and VAS seem to be the real challenge. In this paper we present preliminary results from an interview survey suggesting that people can handle hypothetical states taken from multiple instruments in VAS exercises, thus providing data for preference-based mapping. The paper begins by presenting a brief description of the instruments involved in this study, outlining the methodology of the valuation study and the dataset, and then presents the VAS results. The results of the ranking exercise are still being analysed. The discussion seeks to explore whether this approach has the potential to enable preference-based mapping between preference-based measures of health and quality of life.

*1.1 Measures of health and quality of life*

The study involves 6 preference-based measures of health and quality of life: EQ-5D (generic), SF-6D (generic), HUI2 (generic for children), AQL-5D (asthma specific), OPUS (social care specific), ICECAP (capabilities). The choice of measures reflects a range of different types of measures that are currently in use or nearing use in the UK. These are summarised in table 1.

The EQ-5D is the most widely used generic preference-based measure of health-related quality of life which produces utility scores anchored at 0 for dead and 1 for perfect health (Dolan, 1997). The SF-6D is a generic preference-based single index measure for health that can be estimated using SF-36 and SF-12 data (Brazier and Roberts, 2004). The SF-6D is derived from a selection of SF-36 items and produces utility scores anchored at 0 for dead and 1 for perfect health. The HUI2 is a generic preference-based measure of health for children (Torrance et al. 1996) and produces utility scores anchored at 0 for dead and 1 for perfect health. The AQL-5D is a condition-specific preference-based measure of health for asthma derived from the Asthma Quality of Life Questionnaire (AQLQ) (Yang, 2007) and produces utility scores anchored at 0 for dead and 1 for perfect health.

ICECAP is a preference-based measure of capability for older people in the UK (Grewal et al, 2006). Utility scores are anchored at 0 for zero capabilities and 1 for perfect health. OPUS (older persons' utility scale) is a preference-based social care outcome measure for older people (Ryan et al, 2006). The descriptive system has 5 dimensions (food and nutrition, personal care, safety, social participation, control over daily living) each with 3 levels (no unmet needs, low unmet needs, high unmet needs). Quality weights exist for a utility index excluding the safety dimension and for an index with a redefinition of the safety dimension from the original valuation study (Ryan et al, 2006). Here we use OPUS excluding the safety dimension. Utility scores are anchored at 0 for the worst outcome (high unmet needs on all dimensions) and 1 for the best outcome (no unmet needs on all dimensions).

- Insert table 1 here –

## 2. Methods

The aim of the study is to develop a preference-based method of mapping between preference-based instruments. Respondents preferences are elicited for hypothetical health and well-being states defined by different descriptive systems. This means that the relationship between different instruments is determined directly by people's preferences for different states and not by associations in self-reported values. In the valuation study a sample of health states defined by each of 6 instruments presented above (EQ-5D, SF-6D, HUI2, AQL-5D, OPUS and ICECAP) are valued by a representative sample of the general population.

### *2.1 Study design*

The study has three stages. The first stage is a valuation study to determine people's preferences for health and well-being states described by different descriptive systems. The second stage is to estimate the relationship between the states included in the valuation study. The final stage is to estimate the relationship (i) between the states included in the valuation study and the published value set for each instrument; and (ii) across all instruments in order to enable preference-based mapping between instruments.

### *2.2 Valuation task*

For use in economic evaluation the health state valuation technique should be choice based (Drummond et al, 2005). Conventionally this has included cardinal methods such as standard gamble (SG) and time-trade off (TTO), yet over recent years there has been an increasing interest in using ordinal methods, such as ranking or pair wise comparisons (Salomon, 2003, 2007; McCabe et al, 2006). In a conventional ranking task respondents are asked to order a set of states from best to worst. Traditionally ranking has been used as a warm up exercise prior to preference elicitation using VAS, SG or TTO.

The aim of this study is different to the usual valuation study; here we aim to determine the relationship between instruments, and hence we do not wish to produce or reproduce a value set for any of the instruments. Therefore the main advantages of using ranking for this study is that ranking is cognitively less complex when eliciting preferences over health states from different instruments, enables a direct comparison of health states from different instruments and enables preference elicitation for a greater number of health states per respondent than SG or TTO. Ranked data is often analysed using a rank ordered logit model. However, two assumptions of this model are likely to be rejected by our data: the independence of irrelevant alternatives and the independence of unobservables in repeated choices. These assumptions can be easily relaxed by using a mixed logit model. In addition, commonly used software does not place any restrictions on the estimated parameter values and therefore they need to be rescaled to the full health-dead scale typically used.

A conventional VAS task was also used where respondents are asked to rate health states on a scale from 0 to 100 using a vertical line on a page, where 0 is the 'worst imaginable state' and 100 is the 'best imaginable state'. There are concerns with the use of VAS including end point bias (Torrance et al, 2001) and a concern that is does not generate 'preferences' (Brazier et al, 2003). However these problems are less relevant for this study as a value set will not be produced. VAS provides a cardinal scale that sets of states have been valued against.

*2.3 Valuation study design*

In order to determine the relationship between instruments and to estimate this relationship each respondent values health states from 2 or more descriptive systems. One aim of the valuation study design was to reduce any systematic bias brought about through the combination of states that were valued by each individual both within a valuation task and across all valuation tasks for that individual. The study was designed so that each instrument appeared with each other instrument an equal number of times.

Each interview involved 3 ranking and VAS tasks with states from 3 different instruments and these are combined across tasks so that each task contains states from 2 instruments. The same health and well-being states are used for the rank and VAS tasks. Each task contains 8 states across 2 instruments: one mild state, one moderate state and the worst state for one instrument, one mild state, one moderate state and the worst state for a second instrument, and a generic 'best state' and 'dead'. Each interview involves 15 health and well-being states (5 for each instrument: two mild states, two moderate states and the worst state). Each time an instrument is included in a valuation task the worst state defined by that instrument is valued.

There are 20 variations of the interview as there are 20 different combinations of 3 instruments from the available 6 instruments using the binomial coefficient equation n!/(r!(n-r)!). Each variation of the interview is different because of the health and well-being states that are used in the ranking and VAS tasks, and hence each variation has a different 'card bloc', making 20 card blocs in total. Each instrument appears in 10 out of the 20 card blocs and hence 10 variations of the interview. The interviews are designed so that each health state is valued approximately 75-100 times (with the exception of 50 times for 4 EQ-5D states) and 500 times for each worst state.

Each card bloc involves 3 instruments with one 'card set' for each instrument, and each card set consists of 4 unique states of various severity plus the worst state as the fifth state. Each card set is used in 3 or 4 blocs[1] and across all blocs any given card set appears at least once in the first ranking task, at least once in the second ranking task and at least once in the third ranking task. This is to minimise any data variation due to respondents changing their understanding, consideration and concentration across the duration of the interview.

Figure 1 shows the task design for 6 out of the 20 card blocs. Instruments are labelled A, B, C, D, E and F and health states for each instrument are numbered 1 to 13, where 1 represents the worst state of each instrument.

- Insert figure 1 here -

[1] With the exception of instrument A where each card set is used in 2 or 3 card blocs due to the larger number of states and hence card sets.

*2.4 Selection of health states*

Twelve health and well-being states for each instrument plus the worst state for each instrument are included in the valuation study (except for EQ-5D with 15 states plus the worst state). The states were selected for each instrument (with the exception of EQ-5D) to reflect a range of health state values according to the published value set (or most recent version for AQL-5D and ICECAP) for each instrument, whilst guaranteeing that there is a variety of levels for each dimension included in the chosen health states. Sixteen states were selected for the EQ-5D using an orthogonal design in SPSS to generate an orthogonal array of states that enable the estimation of an additive function to value all states (this was not possible for all of the instruments since the orthogonal array was too large, for example 49 states for the SF-6D alone). All health and well-being states were checked to guarantee that the combination of levels and dimensions were feasible and realistic and hence appropriate for use in ranking and VAS tasks undertaken by members of the general population.

Instrument specific full health was not used for all instruments as instead the generic state 'best state' was used. This approach was chosen as the aim is not to focus upon differences or deficiencies in descriptive systems per se but to focus upon differences in values given to health states with different health problems. The instrument specific best state was only included for two instruments: EQ-5D since it is specified by orthogonal array and OPUS.

*2.5 The interviews*

All interviews were conducted by trained and experienced interviewers in the respondent's own home. The interviewers were undertaken by the Centre for Research and Evaluation (CRE) at Sheffield Hallam University who have done numerous valuation surveys including the UK valuation of HUI2 (McCabe et al.; 2005) and AQL-5D (Yang et al.; 2007). Respondents were not offered any financial reward for their participation.

The interview began with the respondent being asked to report their own health using the descriptive system for the EQ-5D and subsequently all instruments that the individual would value during the interview. This familiarised the respondent with the idea of describing states and the items and levels in the descriptive system for each of the instruments involved in their interview.

In the next stage of the interview the respondent was asked to rank a shuffled set of 8 cards (task 1) as described above. The respondent was asked to place the cards in order of how good or bad they think they are. Respondents were asked to imagine that they themselves were actually in each state and that it is going to last for the rest of their life without changing. The respondent was then asked to rate these same states (without reshuffling the cards) using the VAS, allowing respondents to change the ordering from the previous ranking. The respondent was then asked to repeat the ranking and VAS tasks twice with different sets of cards (tasks 2 and 3).

The final stage of the interview involved self-completion questions on background characteristics and how difficult the respondent found the interview.

*2.6 Selection of respondents*

Respondents were from the geographical areas in the North of England including urban and rural areas with a mix of socio-economic characteristics. Streets were sampled from the selected areas and all willing participants within quota were interviewed. Letters were sent to households informing them the interviewers will be in their area. Interviewers then visited houses and interviewed all people who were willing to participate and within quota from those addresses.

*2.7 Analysis*

Raw VAS ratings measured on the 0 to 100 scale are rescaled for each task completed using the equation outlined in Brazier et al. (2003):

$$A_j = \frac{R_j - R(dead)}{R(best) - R(dead)} \qquad (1)$$

where $A_j$ represents the adjusted VAS rating for each health state $j = 1, \quad 2, \quad ..., \quad J$, ***R(dead)*** represents the raw rating given to 'dead', $R_j$ represents the raw rating given to health state ***j*** and ***R(best)*** represents the raw rating given to the best health state.[2] This rescales the values such that the highest valued state equals 1 and dead equals 0, hence states can have a value worse than dead.

Duration, completion, understanding, difficulty and effort are reported for a selection of card blocs. Mean and standard deviation (SD) of adjusted VAS score for highest and worst states by instrument are presented. Mean, SD, median and interquartile range (IQR) are reported for raw and adjusted VAS data for all states, and adjusted VAS values are compared to published value sets. For the EQ-5D predicted adjusted VAS score is compared to the published VAS value set.

**3. The data**

There were 502 successfully conducted interviews, a response rate of 55% for suitable respondents answering their door at time of interview. The study achieved a completion rate of 99% for all states included in the rank and rating tasks (140 rank values and 178 VAS values missing out of 12,024 values) and 94.8% (475/501) of respondents had complete responses. Two respondents (0.4%) had no rank and rating responses and hence are excluded from the analysis and one respondent is excluded for unusable responses. Characteristics of all respondents included in the analysis are presented in table 2 and compared to the general population in South Yorkshire and England. The sample has a higher proportion of retired individuals and home owners.

- Insert table 2 here -

[2] One alternative rescaling method of VAS data is the Parducci Range-Frequency model (see Parducci and Wedell (1986)). However this requires the arbitrary setting of crucial weighting parameters and hence has not been used here. The MVH EQ-5D study adjusted VAS values using equation (1) where *R(best)* represents the raw rating given to state 11111, generic full health. This adjustment cannot be used here as we do not have instrument specific best state for all instruments.

## 4. Results

### *4.1 Feasibility*

Table 3 presents duration of interview, self-reported difficulty of questions and understanding and effort and concentration as reported by the interviewer for the most difficult and easiest card blocs. Mean duration varies from 32 to 41 minutes across all blocs and missing VAS responses range from 0 to 33 for each card bloc, with missing values seeming unrelated to difficulty. Responses for difficulty, understanding and effort and concentration are similar for all card blocs, suggesting little variation in feasibility across all blocs. There is no clear pattern of difficult combinations of instruments, as one of the easiest and both of the most difficult blocs contain AQL-5D and ICECAP.

- Insert table 3 here -

### *4.2 VAS results*

Table 4 shows the worst and highest states by instrument indicating that the best state value for each instrument is lower than 1 and the worst state value varies by instrument (range=0.075-0.289). Instrument specific full health is less than 1 for EQ-5D (0.909) and OPUS (0.899). Table 5 presents observed raw and adjusted VAS scores (where dead=0) for all health states included in the valuation study. Perhaps unsurprisingly all instruments have a wide range of VAS values, where the AQL-5D (condition specific measure) has a smaller range and this is consistent with the published value set utility range using TTO. The results suggest important differences in the value range of each of the instruments (with mean worst state values ranging from 0.075 to 0.289). The inter-quartile range suggests that adjusted values vary across individuals and potentially across card blocs.

We modelled the VAS utility value on the vector of all health states and background characteristics using a maximum likelihood random effects model. No background characteristics variables were significant and hence to minimise error the mean values for health states are reported here rather than predicted values.

- Insert tables 4 and 5 and figure 2 here -

Figure 2 plots adjusted VAS values and the published value sets for each instrument, indicating the different relationships for each instrument. Note that the published value sets are typically not based on VAS. It shows that EQ-5D VAS values from our study are mostly higher than the published TTO-based values. Overall VAS values have a smaller range and spread than the published value set utility scores using other valuation methods. The results indicate that for milder health and well-being states the VAS value is often lower than the published value set utility score, whereas for more severe states the VAS value is often higher than the published value set utility score. This relationship is present for each instrument with the exception of the AQL-5D where all VAS values are lower than the published value set utility scores, and EQ-5D where most VAS values are higher than the published TTO and VAS value set utility scores (as shown in table 5).

### *4.3 EQ-5D results*

The orthogonal array used to select the EQ-5D states in the valuation study enables the estimation of an additive function to value all EQ-5D states. Figure 3 shows the relationship between predicted adjusted VAS score and the MVH EQ-5D VAS published value set for all EQ-5D states. Predicted adjusted VAS score is estimated using a maximum likelihood random effects model using the model specified in Dolan (1997) using the adjusted VAS data. All main effects coefficients are of the expected sign and all are significant at the 1% level with the exception of the dummy variable representing usual activities at the most severe level. The 'N3' term, a dummy variable for states with at least one dimension at the most severe level, is significant but smaller than that in the MVH EQ-5D VAS published value set. Figure 3 shows that predicted adjusted VAS scores are similar to the MVH VAS value set for mild states, yet overpredict for the majority of states with at least one dimension at the most severe level (ICC = 0.899).

- Insert figure 3 here -

**5. Discussion**

Respondents were able to value states from different instruments alongside each other using VAS and ranking methods and extremely high completion rates were achieved at the task level. The results indicate the feasibility of VAS as a means of comparison of health states across different instruments. One advantage of VAS or ranking over TTO or SG is that respondents see multiple states at the same time.

Each card bloc appears feasible and comparable in terms of completion, understanding, difficulty and effort. Yet the combination of instruments in a given task may affect VAS values. Variation in range by instrument seems to reflect differences in the highest state that is valued and the severity of the worst state for each instrument. Differences in highest state valued is due in part to the valuation study design, but differences in values for the worst state are due to preferences and the severity of the instrument. In particular, the EQ-5D VAS model appears to be systematically different from the MVH EQ-5D VAS model (see Figure 3), though ICC is high. Of interest is identifying the impact of combining EQ-5D states with states from other instruments.

The next part of the study will involve analysis on the rank data to estimate the relationship between the health states included in the valuation study. Previous work on ranking data has used a conditional logit model to estimate values for health states. Under the assumptions of the conditional logit model, a ranking task can be seen as being the same as a sequence of independent choices in smaller and smaller sets of alternatives. However, in reality these choices are rarely independent since choices made by each individual tend to be related due to unobservable factors. Another related problem of this modelling methodology is that the conditional logit model exhibits Independence of Irrelevant Alternatives (IIA). We will estimate a mixed logit model on ranking data to take into account the panel structure of the data and at the same time relax the IIA property.

The final stage of the study involves the use of these results to map between the six measures of health and quality of life to estimate a full value set using the VAS data. This will involve estimating the relationship between the 13 health states valued for each instrument and the original value set. This will be done using the VAS and rank data collected in the valuation study. This mapping uses preferences rather than statistical association and is better able to take advantage of diversity in descriptive systems for different measures. This will enable the integration of evidence from a larger range of studies for economic evaluation and hence enable better cost effectiveness models to be produced.

## 6. Conclusions

The rationale of the study is to develop a preference-based method for mapping between preference-based measures of health and quality of life. Our valuation study uses a general population sample to value six health and quality of life instruments using rank and VAS valuation techniques. Our VAS results indicate that a preference-based method is a feasible method for mapping between preference-based measures of health and quality of life.

## Acknowledgements

We would like to thank Professor Coast and Professor Netten for the use of ICECAP and OPUS respectively. We would also like to thank all the interviewees who took part in the survey. This study is funded by a project grant from the MRC. The usual disclaimer applies.

**References**


Barton, G.R., Bankart, J., Davis, A.C., Summerfield, Q.A., Comparing utility scores before and after hearing aid provision: Results According to the EQ-5D, HUI3 and SF-6D. *Applied Health Economics and Health Policy* 3:103-105, 2004.

Brazier, J., Greene, C., McCabe, C., and Stevens, K., Use of visual analog scales in economic evaluation. Expert Review of Pharmacoeconomics Outcomes Research 3:293-302, 2003.

Brazier. J., Roberts, J., The estimation of a Preference-Based Measure of Health From the SF-12. *Medical Care* 42:851-859, 2004.

Brazier. J., Roberts, J., Deverill, M., The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics* 21:271-292, 2002.

Brazier. J., Roberts, J., Tsuchiya, A. and Busschbach, J., A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Economics* 13:873-884, 2004.

Brooks, R., Euroqol: the current state of play. *Health Policy* 37:54-72, 1996.

Coast, J., Flynn, T., Grewal, I., Natarajan, L., Lewis, J., Sprotson, K., Developing an index of capability for health and social policy evaluation for older people: theoretical and methodological challenges. *HESG, January 2006.*

Dolan, P., Modeling Valuations for EuroQol Health States. *Medical Care* 35:1095-1108, 1997.

Dowie, J., Decision validity should determine whether a generic or condition-specific HRQOL measure is used in health care decisions. *Health Economics* 11:1-8, 2002.

Drummond, M.F., Schulpher, M.J., Torrance, G.W., O'Brien, B.J., Stoddart, G.L., *Methods for the Economic Evaluation of Health Care Programmes* Oxford: Oxford University Press, 2005.

Espallargues, M., Czoski-Murray, C., Bansback, N., Carlton, J., Lewis, G., Hughes, L., Brand, C., Brazier, J., The impact of Age Related Macular Degeneration on Health Status Utility Values. *Investigative Ophthalmology and Visual Science* 46:4016-4023, 2005.

Feeny, D., Furlong, W., Torrance, G.W., Goldsmith, C.H., Zhu, Z., DePauw, S., Denton, M., Boyle, M., Multi-attribute and single-attribute utility functions for the Health Utilities Index Mark 3 system. *Medical Care* 40,113-128, 2002.

Franks, P., Lubetkin, E. I., Gold Marthe R, Tancredi, D. J., Haomiao, J., Mapping the SF-12 to the EuroQol EQ-5D Index in a National US Sample. *Medical Decision Making* 24:247-254, 2004.

Gray, A. M., Rivero-Arias, O., Clarke, P. M., Estimating the Association between SF-12 Responses and EQ-5D Utility Values by Response Mapping. *Medical Decision Making* 26:18-29, 2006.

Grewal, I., Lewis, J., Flynn, T.N., Brown, J., Bond, J., Coast, J., Developing attributes for a generic quality of life measure for older people: preferences or capabilities? *Social Science and Medicine* 62:1891-1901, 2006.

Harper, R., Brazier, J.E., Waterhouse, J.C., Walters, S.J., Jones, N.M.B., Howard, P. (1997) A comparison of outcome measures for patients with chronic obstructive pulmonary disease (COPD) in an outpatient setting. *Thorax* 52:879-887, 1997.

Kaplan, R.M., Anderson, J.P., A general health policy model: Update and applications. *Health Services Research* 23:203-235, 1998.

Longworth, L., Bryan, S., An empirical comparison of EQ-5D and SF-6D in liver transplant patients. *Health economics* 12:1061-1067, 2003.

Nichol, M.B., Sengupta, N., Globe, D.R., Evaluating Quality Adjusted Life Years: Estimation of the Health Utility Index (HUI2) from the SF-36. *Medical Decision Making* 21:105-112, 2001.

McCabe, C., Stevens, K., Roberts, J., Brazier, J., Health state values for the HUI 2 descriptive system: Results from a UK survey. *Health Economics* 14:231-244, 2005.

McCabe, C., Brazier, J., Gilks, P., Tsuchiya, A., Roberts, J., O'Hagan, A., Stevens, K., Using rank data to estimate health state utility models. *Journal of Health Economics* 25:418-431, 2006.

O'Brien, B.J., Spath, M., Blackhouse, G., Severens, J.L., Brazier, J.E., A view from the Bridge: agreement between the SF-6D utility algorithm and the Health Utilities Index. *Health Economics* 12:975-982, 2003.

Revicki, D.A., Leidy, N.K., Brennan-Deimer, F., Sorenson, S. and Togias, A., Integrating patients preferences into health outcomes assessment: The multi-attribute asthma symptom utility index. *Chest* 114:998-1007, 1998.

Ryan, M., Netten, A., Skatun, D. and Smith, P., Using discrete choice experiments to estimate a preference-based measure of outcome - An application to social care for older people. *Journal of Health Economics* 25:927-944, 2006.

Salomon, J.A., Using ordinal data to estimate cardinal valuations. In: *Measuring and Valuing Health Benefits for Economic Evaluation*, edited by Brazier, J., Ratcliffe, J., Salomon, J.A., Tsuchiya, A., Oxford: Oxford University Press, 2007.

Salomon, J.A., Reconsidering the use of rankings in the valuation of health states: a model for estimating cardinal values from ordinal data. *Population Health Metrics* 1 (12), 2003.

Stevens, K., Working with children to develop dimensions for a preference based generic paediatric health related quality of life measure. *Health Economics and Decision Sciences Discussion Paper,* 2008.

Torrance, G.W., Feeny, D.H., Furlong, W.J., Barr, R.D., Zhang, Y., Wang, Q., A multi-attribute utility function for a comprehensive health status classification system: Health Utilities Mark 2. *Medical Care* 34:702-722, 1996.

Torrance, G.W., Feeny, D.H., Furlong, W.J., Visual Analogue Scales: do they have a role in the measurement of preferences for health states? *Medical Decision Making* 21:329-334, 2001.

Tsuchiya, A., Brazier, J., Roberts, J., Comparison of valuation methods used to generate the EQ-5D and SF-6D value sets. *Journal of Health Economics* 25:334-346, 2006.

Tsuchiya, A., Brazier,J., McColl,E., Parkin, D., Deriving preference-based single indices from non-preference based condition-specific instruments: Converting AQLQ into EQ5D indices, *Sheffield Health Economics Group Discussion Paper*, 2002.

Yang, Y., Tsuchiya, A., Brazier, J., Young, T., Estimating a preference-based single index from the Asthma Quality of Life Questionnaire (AQLQ). *Health Economics and Decision Sciences Discussion Paper*, 2007.

Yang, Y., Brazier, J., Tsuchiya, A., Coyne, K., Estimating a Preference-Based Single Index from the Overactive Bladder Questionnaire. *Value in Health*, forthcoming, 2008.

**Table 1** Measures of health and quality of life

| *Instrument* | *Summary* | *Dimensions* | *Levels* | *Unique states* | *Reference* |
|---|---|---|---|---|---|
| EQ-5D | Generic | 5 dimensions: Mobility, self-care, usual activity, pain/discomfort and anxiety/depression | 3 levels: no problems, some problems, extreme problems | 243 | Brooks (1996) |
| SF-6D | Generic | 6 dimensions: Physical functioning, role limitations, social functioning, pain, mental health, vitality | Between 4 and 6 levels, depends on the dimensions | 18,000 | Brazier et al. (2002, 2004) |
| HUI2 | Generic for children | 7 dimensions: Sensation, mobility, emotion, cognition, self care, pain, fertility | 4 or 5 levels, depends on the dimensions | 8,000 | Torrance et al. (1996) |
| AQL-5D | Condition specific for asthma | 5 dimensions: Concern about asthma, shortness of breath, weather and pollution stimuli, sleep impact and activity limitations | 5 levels: no problems to extreme problems | 3,125 | Yang et al. (2007) |
| ICECAP | Capability measure for older people in UK | 5 dimensions: Attachment, security, role, enjoyment, control | 4 levels: all, a lot, a little, none | 1,024 | Grewal et al. (2006) |
| OPUS | Social care outcome measure for older people | 5 dimensions: Food and nutrition, personal care, safety, social participation, control over daily living | 3 levels: no unmet needs, low unmet needs, high unmet needs | 243 | Ryan et al. (2006) |

**Table 2** Characteristics of respondents

| | *Included respondents (n=499)* | *South Yorkshire*[3] | *England* |
|---|---|---|---|
| Mean age (s.d.) | 48 (18) | - | - |
| Female | 49.9% | 51.2% | 51.3% |
| Married/Partner | 69.1% | - | - |
| Employed or self-employed | 49.1% | 56.1% | 60.9% |
| Unemployed | 1.4% | 4.1% | 3.4% |
| Long-term sick | 3.6% | 7.7% | 5.3% |
| Full-time student | 4.2% | 7.5% | 7.3% |
| Retired | 30.5% | 14.4% | 13.5% |
| Own home outright or with a mortgage | 86.1% | 64.0% | 68.7% |
| Renting property | 13.9% | 36.0% | 31.3% |
| Secondary school is highest level of education | 43.9% | - | - |
| EQ-5D score (s.d.) | 0.86 (0.23) | - | - |
| Found valuation task difficult (judged by respondent) | 29.3% | - | - |
| Doubtful whether the respondent understood the tasks (judged by interviewer) | 4.1% | - | - |

[3] Statistics for South Yorkshire Health Authority and for England in the Census 2001. Questions used in this study and the census are not identical. The census includes persons aged 16 and above whereas this study only surveys persons aged 18 and above.

**Table 3** Duration, completion, difficulty, understanding and effort by card bloc

| | *Card bloc* | | | |
|---|---|---|---|---|
| | Most difficult blocs | | Easiest blocs | |
| | 17 | 8 | 7 | 14 |
| Instruments | AQL-5D<br>OPUS<br>ICECAP | SF-6D<br>ICECAP<br>AQL-5D | HUI2<br>OPUS<br>EQ-5D | AQL-5D<br>HUI2<br>ICECAP |
| *n* | 24 | 25 | 27 | 25 |
| Duration | 34 | 41 | 32 | 32 |
| Missing VAS responses | 0 | 1 | 33 | 5 |
| | | | | |
| *Difficulty of questions, self-reported* | | | | |
| Very difficult | 3 | 3 | | |
| Quite difficult | 8 | 4 | 4 | 4 |
| Neither | 8 | 4 | 4 | 10 |
| Fairly easy | 3 | 13 | 13 | 9 |
| Very easy | 2 | 1 | 4 | 1 |
| | | | | |
| *Understanding, reported by interviewer* | | | | |
| Understood and performed exercises easily | 17 | 15 | 23 | 16 |
| Some problems but seemed to understand in the end | 7 | 8 | 1 | 8 |
| Doubtful whether the respondent understood the exercises | 0 | 2 | 2 | |
| | | | | |
| *Effort and concentration, reported by interviewer* | | | | |
| Concentrated very hard | 15 | 16 | 19 | 12 |
| Concentrated fairly hard | 9 | 6 | 5 | 11 |
| Didn't concentrate very hard | 0 | 1 | 1 | |
| Concentrated at start but lost interest towards end | | 2 | 1 | 1 |

**Table 4** Adjusted VAS score of worst and highest states by instrument

| *Instrument* | *Highest state valued* | | *Worst state* | |
|---|---|---|---|---|
| | Health state | Mean (s.d.) | Health state | Mean (s.d.) |
| EQ-5D | 11111 | 0.909 (0.159) | 33333 | 0.075 (0.228) |
| SF-6D | 211111 | 0.860 (0.133) | 645655 | 0.266 (0.250) |
| HUI2 | 112222 | 0.706 (0.210) | 455445 | 0.077 (0.229) |
| AQL-5D | 13321 | 0.717 (0.205) | 55555 | 0.289 (0.816) |
| ICECAP | 12321 | 0.872 (0.114) | 44444 | 0.227 (0.275) |
| OPUS | 1111 | 0.899 (0.111) | 3333 | 0.186 (0.871) |

**Table 5** Health states with observed raw and adjusted VAS scores and published value set

| *Health state description* | *Number of valuations* | *Mean raw VAS* | *Standard deviation for raw VAS* | *Median for raw VAS* | *Interquartile range for raw VAS* | *Mean adjusted VAS* | *Standard deviation for adjusted VAS* | *Interquartile range for adjusted VAS* | *Median for adjusted VAS* | *VAS published value set utility value (where available[4]* |
|---|---|---|---|---|---|---|---|---|---|---|
| *EQ-5D* | | | | | | | | | | |
| 11111 | 74 | 90.82 | 15.71 | 95 | 90.00-100.00 | 0.909 | 0.159 | 0.900-1.000 | 0.95 | 1 |
| 11322 | 50 | 71.26 | 19.04 | 70 | 65.00-86.25 | 0.709 | 0.200 | 0.647-0.861 | 0.70 | 0.403 |
| 12311 | 76 | 69.09 | 20.11 | 72.5 | 55.00-85.00 | 0.676 | 0.213 | 0.540-0.848 | 0.70 | 0.457 |
| 21113 | 75 | 62.48 | 21.26 | 70 | 50.00-77.00 | 0.604 | 0.223 | 0.500-0.750 | 0.67 | 0.435 |
| 13211 | 74 | 61.69 | 22.22 | 60 | 50.00-80.00 | 0.604 | 0.234 | 0.498-0.800 | 0.60 | 0.455 |
| 11223 | 75 | 59.55 | 22.97 | 60 | 45.00-75.00 | 0.577 | 0.238 | 0.433-0.750 | 0.60 | 0.392 |
| 22212 | 98 | 58.25 | 21.94 | 60 | 40.00-75.00 | 0.570 | 0.231 | 0.400-0.750 | 0.60 | 0.587 |
| 21331 | 50 | 54.70 | 20.55 | 60 | 39.75-70.00 | 0.540 | 0.212 | 0.374-0.700 | 0.59 | 0.308 |
| 23121 | 75 | 51.29 | 22.80 | 50 | 35.00-70.00 | 0.492 | 0.248 | 0.330-0.688 | 0.50 | 0.330 |
| 13132 | 76 | 47.29 | 24.54 | 41.5 | 25.00-70.75 | 0.438 | 0.299 | 0.250-0.619 | 0.40 | 0.251 |
| 12133 | 47 | 42.91 | 26.44 | 40 | 20.00-67.00 | 0.381 | 0.384 | 0.211-0.612 | 0.37 | 0.243 |
| 31112 | 47 | 42.47 | 24.28 | 40 | 22.00-60.00 | 0.392 | 0.261 | 0.200-0.600 | 0.40 | 0.385 |
| 31231 | 50 | 36.04 | 21.88 | 30 | 19.75-50.00 | 0.347 | 0.224 | 0.166-0.493 | 0.30 | 0.247 |
| 32121 | 48 | 31.98 | 23.26 | 30 | 15.00-45.00 | 0.288 | 0.264 | 0.150-0.450 | 0.25 | 0.272 |
| 33313 | 75 | 23.23 | 21.06 | 15 | 10.00-30.00 | 0.197 | 0.235 | 0.053-0.300 | 0.15 | 0.099 |
| 33333 (worst state) | 496 | 11.89 | 15.77 | 5 | 1.00-15.00 | 0.075 | 0.227 | 0.000-0.150 | 0.05 | -0.072 |
| *SF-6D* | | | | | | | | | | |
| 211111 | 99 | 85.91 | 13.11 | 90 | 85.00-95.00 | 0.860 | 0.133 | 0.847-0.947 | 0.90 | |
| 211211 | 76 | 77.89 | 17.13 | 85 | 70.00-90.00 | 0.777 | 0.182 | 0.697-0.900 | 0.85 | |
| 112221 | 74 | 75.52 | 20.55 | 80 | 67.50-90.00 | 0.754 | 0.204 | 0.675-0.900 | 0.80 | |
| 111453 | 99 | 66.77 | 21.89 | 75 | 55.00-81.00 | 0.666 | 0.227 | 0.526-0.850 | 0.75 | |
| 214411 | 76 | 65.45 | 21.76 | 70 | 50.00-80.00 | 0.652 | 0.222 | 0.500-0.830 | 0.70 | |
| 623133 | 76 | 54.26 | 21.85 | 58 | 40.00-70.00 | 0.540 | 0.224 | 0.400-0.700 | 0.58 | |
| 424421 | 74 | 54.07 | 22.93 | 59 | 40.00-73.00 | 0.532 | 0.241 | 0.350-0.730 | 0.55 | |

[4] EQ-5D VAS values quoted in 'The Measurement and Valuation of Health; First report on the main survey', The MVH Group, May 1994. Values used are the VAS tariff of means: Whole population – 10 year duration.

| *Health state description* | *Number of valuations* | *Mean raw VAS* | *Standard deviation for raw VAS* | *Median for raw VAS* | *Interquartile range for raw VAS* | *Mean adjusted VAS* | *Standard deviation for adjusted VAS* | *Interquartile range for adjusted VAS* | *Median for adjusted VAS* | *VAS published value set utility value (where available[4]* |
|---|---|---|---|---|---|---|---|---|---|---|
| 311655 | 99 | 50.00 | 22.98 | 49 | 32.00-70.00 | 0.490 | 0.237 | 0.300-0.700 | 0.48 | |
| 545622 | 74 | 49.52 | 21.51 | 50 | 31.50-66.00 | 0.484 | 0.234 | 0.300-0.650 | 0.50 | |
| 422655 | 76 | 42.8 | 22.34 | 40 | 25.00-65.00 | 0.411 | 0.267 | 0.239-0.650 | 0.40 | |
| 624343 | 74 | 42.77 | 19.21 | 40 | 30.00-60.00 | 0.411 | 0.195 | 0.250-0.575 | 0.40 | |
| 535645 | 99 | 35.45 | 22.20 | 35 | 15.00-50.00 | 0.340 | 0.242 | 0.150-0.500 | 0.33 | |
| 645655 (worst state) | 498 | 28.72 | 20.96 | 25 | 10.00-40.00 | 0.266 | 0.250 | 0.100-0.400 | 0.25 | |
| | | | | | | | | | | |
| *AQL-5D* | | | | | | | | | | |
| 13321 | 72 | 71.83 | 20.62 | 79 | 60.00-90.00 | 0.717 | 0.205 | 0.590-0.876 | 0.79 | |
| 21223 | 98 | 70.28 | 20.24 | 75 | 60.00-85.00 | 0.701 | 0.205 | 0.598-0.850 | 0.76 | |
| 53411 | 72 | 65.36 | 21.26 | 70 | 50.00-85.00 | 0.647 | 0.212 | 0.500-0.830 | 0.70 | |
| 32441 | 76 | 61.70 | 19.94 | 65 | 50.00-78.25 | 0.390 | 2.008 | 0.500-0.783 | 0.63 | |
| 12543 | 100 | 61.52 | 19.68 | 60 | 50.00-75.00 | 0.608 | 0.203 | 0.500-0.750 | 0.60 | |
| 45143 | 76 | 55.88 | 18.91 | 56 | 45.00-70.00 | 0.551 | 0.200 | 0.444-0.700 | 0.55 | |
| 23534 | 76 | 55.83 | 20.12 | 60 | 45.00-70.00 | 0.550 | 0.205 | 0.400-0.700 | 0.60 | |
| 52314 | 100 | 52.62 | 21.82 | 53.5 | 36.75-69.25 | 0.515 | 0.226 | 0.350-0.689 | 0.52 | |
| 15355 | 76 | 51.18 | 20.65 | 50 | 36.75-63.75 | 0.261 | 2.179 | 0.361-0.600 | 0.50 | |
| 34254 | 72 | 46.38 | 22.99 | 45 | 30.00-65.00 | 0.424 | 0.324 | 0.261-0.637 | 0.40 | |
| 55424 | 98 | 43.99 | 20.19 | 45 | 30.00-59.25 | 0.432 | 0.210 | 0.300-0.586 | 0.43 | |
| 34554 | 72 | 43.00 | 22.82 | 40 | 25.00-65.00 | 0.403 | 0.274 | 0.250-0.600 | 0.40 | |
| 55555 (worst state) | 494 | 34.47 | 21.22 | 30 | 20.00-49.00 | 0.289 | 0.816 | 0.158-0.480 | 0.30 | |
| | | | | | | | | | | |
| *HUI2* | | | | | | | | | | |
| 112222 | 75 | 70.97 | 20.49 | 75 | 60.00-85.00 | 0.706 | 0.210 | 0.600-0.856 | 0.75 | |
| 121132 | 76 | 66.86 | 18.44 | 70 | 55.00-80.00 | 0.661 | 0.194 | 0.550-0.800 | 0.69 | |
| 112123 | 99 | 64.82 | 23.09 | 70 | 50.00-85.00 | 0.639 | 0.235 | 0.500-0.849 | 0.69 | |
| 331131 | 98 | 53.88 | 24.77 | 50 | 40.00-75.00 | 0.499 | 0.375 | 0.327-0.713 | 0.50 | |
| 323331 | 76 | 50.46 | 21.58 | 50 | 40.00-64.00 | 0.490 | 0.224 | 0.376-0.650 | 0.50 | |
| 314431 | 75 | 45.78 | 24.34 | 45 | 25.00-63.50 | 0.442 | 0.267 | 0.246-0.622 | 0.45 | |
| 234111 | 74 | 43.92 | 18.09 | 44 | 34.00-54.50 | 0.428 | 0.192 | 0.308-0.533 | 0.40 | |
| 344222 | 76 | 42.36 | 23.50 | 42.5 | 25.00-55.75 | 0.407 | 0.244 | 0.250-0.550 | 0.40 | |

| *Health state description* | *Number of valuations* | *Mean raw VAS* | *Standard deviation for raw VAS* | *Median for raw VAS* | *Interquartile range for raw VAS* | *Mean adjusted VAS* | *Standard deviation for adjusted VAS* | *Interquartile range for adjusted VAS* | *Median for adjusted VAS* | *VAS published value set utility value (where available[4]* |
|---|---|---|---|---|---|---|---|---|---|---|
| 133444 | 98 | 39.82 | 26.32 | 33.5 | 20.00-60.00 | 0.350 | 0.355 | 0.150-0.600 | 0.30 | |
| 144325 | 75 | 38.46 | 27.32 | 30 | 15.00-65.00 | 0.369 | 0.296 | 0.150-0.621 | 0.30 | |
| 125425 | 77 | 34.70 | 24.05 | 30 | 15.00-49.50 | 0.323 | 0.269 | 0.119-0.495 | 0.30 | |
| 445234 | 98 | 19.87 | 17.26 | 15 | 5.00-30.00 | 0.155 | 0.223 | 0.037-0.275 | 0.15 | |
| 455445 (worst state) | 501 | 11.79 | 16.21 | 5 | 0.00-15.00 | 0.077 | 0.229 | 0.000-0.150 | 0.05 | |
| | | | | | | | | | | |
| *ICECAP* | | | | | | | | | | |
| 12321 | 100 | 87.11 | 11.29 | 90 | 85.00-95.00 | 0.872 | 0.114 | 0.850-0.950 | 0.90 | |
| 21131 | 73 | 86.33 | 14.15 | 90 | 85.00-95.00 | 0.857 | 0.149 | 0.817-0.950 | 0.90 | |
| 31212 | 72 | 80.47 | 14.99 | 85 | 71.25-90.00 | 0.805 | 0.151 | 0.709-0.900 | 0.85 | |
| 22242 | 72 | 74.07 | 20.57 | 80 | 65.00-90.00 | 0.735 | 0.217 | 0.650-0.900 | 0.80 | |
| 23324 | 100 | 68.71 | 20.87 | 75 | 50.00-85.75 | 0.682 | 0.217 | 0.513-0.879 | 0.75 | |
| 33333 | 72 | 58.75 | 23.16 | 65 | 40.00-78.00 | 0.567 | 0.290 | 0.400-0.780 | 0.62 | |
| 14344 | 100 | 57.65 | 25.13 | 58 | 35.00-80.00 | 0.569 | 0.261 | 0.350-0.800 | 0.58 | |
| 43111 | 72 | 55.07 | 21.50 | 55 | 40.00-72.00 | 0.538 | 0.227 | 0.400-0.706 | 0.55 | |
| 44143 | 73 | 44.38 | 22.51 | 45 | 30.00-60.00 | 0.398 | 0.275 | 0.211-0.553 | 0.44 | |
| 43443 | 72 | 40.42 | 20.05 | 39.5 | 25.50-60.00 | 0.386 | 0.217 | 0.250-0.600 | 0.36 | |
| 43334 | 72 | 37.25 | 23.44 | 30 | 20.00-50.00 | 0.352 | 0.245 | 0.190-0.500 | 0.30 | |
| 42444 | 99 | 36.82 | 23.32 | 30 | 17.00-55.00 | 0.343 | 0.258 | 0.150-0.550 | 0.30 | |
| 44444 (worst state) | 490 | 26.07 | 22.66 | 20 | 10.00-40.00 | 0.226 | 0.275 | 0.052-0.400 | 0.16 | |
| | | | | | | | | | | |
| *OPUS* | | | | | | | | | | |
| 1111 | 74 | 89.80 | 11.09 | 95 | 85.75-95.00 | 0.899 | 0.111 | 0.891-0.952 | 0.94 | |
| 2121 | 74 | 66.97 | 21.84 | 72 | 53.75-85.00 | 0.663 | 0.213 | 0.538-0.823 | 0.71 | |
| 3121 | 76 | 65.13 | 19.13 | 70 | 50.00-80.00 | 0.641 | 0.199 | 0.500-0.800 | 0.65 | |
| 2212 | 99 | 60.27 | 21.85 | 60 | 45.00-80.00 | 0.596 | 0.224 | 0.450-0.798 | 0.60 | |
| 3132 | 76 | 53.76 | 19.02 | 55 | 35.00-70.00 | 0.519 | 0.194 | 0.350-0.690 | 0.50 | |
| 2123 | 74 | 48.74 | 24.24 | 49.5 | 30.00-71.25 | 0.451 | 0.300 | 0.285-0.692 | 0.46 | |
| 1322 | 99 | 48.70 | 25.51 | 50 | 25.00-70.00 | 0.289 | 1.915 | 0.242-0.700 | 0.50 | |
| 1233 | 99 | 46.16 | 23.79 | 45 | 25.00-65.00 | 0.451 | 0.244 | 0.250-0.650 | 0.45 | |
| 3221 | 99 | 45.62 | 23.51 | 45 | 27.00-65.00 | 0.349 | 0.979 | 0.250-0.643 | 0.40 | |

| *Health state description* | *Number of valuations* | *Mean raw VAS* | *Standard deviation for raw VAS* | *Median for raw VAS* | *Interquartile range for raw VAS* | *Mean adjusted VAS* | *Standard deviation for adjusted VAS* | *Interquartile range for adjusted VAS* | *Median for adjusted VAS* | *VAS published value set utility value (where available[4]* |
|---|---|---|---|---|---|---|---|---|---|---|
| 2331 | 74 | 44.42 | 20.94 | 42 | 30.00-60.00 | 0.409 | 0.253 | 0.239-0.600 | 0.41 | |
| 3313 | 74 | 40.45 | 21.89 | 40 | 24.50-55.00 | 0.386 | 0.224 | 0.219-0.550 | 0.40 | |
| 1333 | 74 | 36.43 | 24.77 | 30 | 18.75-54.25 | 0.318 | 0.329 | 0.150-0.550 | 0.27 | |
| 3333 (worst state) | 496 | 25.55 | 20.84 | 20 | 10.00-40.00 | 0.186 | 0.870 | 0.071-0.350 | 0.20 | |
| Best state | 1488 | 99.40 | 3.01 | 100.00 | 100.00-100.00 | 0.999 | 0.010 | 1.000-1.000 | 1.00 | |
| Dead | 1488 | 3.08 | 8.75 | 0.00 | 0.00-1.00 | 0.000 | 0.001 | 0.000-0.000 | 0.00 | |

**Figure 1** Health states for the ranking tasks for blocs 1-6

BLOC 1 – A, B, C. *n=25*

| Task 1 | Task 2 | Task 3 |
|---|---|---|
| A1 | A1 | B1 |
| A2 | A4 | B4 |
| A3 | A5 | B5 |
| B1 | C1 | C1 |
| B2 | C2 | C4 |
| B3 | C3 | C5 |
| Best state | Best state | Best state |
| Dead | Dead | Dead |

BLOC 2 – D, E, F. *n=25*

| Task 1 | Task 2 | Task 3 |
|---|---|---|
| D1 | D1 | E1 |
| D2 | D4 | E4 |
| D3 | D5 | E5 |
| E1 | F1 | F1 |
| E2 | F2 | F4 |
| E3 | F3 | F5 |
| Best state | Best state | Best state |
| Dead | Dead | Dead |

BLOC 3 – E, F, A. *n=25*

| Task 1 | Task 2 | Task 3 |
|---|---|---|
| E1 | E1 | F1 |
| E6 | E8 | F8 |
| E7 | E9 | F9 |
| F1 | A1 | A1 |
| F6 | A6 | A8 |
| F7 | A7 | A9 |
| Best state | Best state | Best state |
| Dead | Dead | Dead |

BLOC 4 – B, C, D. *n=25*

| Task 1 | Task 2 | Task 3 |
|---|---|---|
| B1 | B1 | C1 |
| B6 | B8 | C12 |
| B7 | B9 | C13 |
| C1 | D1 | D1 |
| C10 | D10 | D12 |
| C11 | D11 | D13 |
| Best state | Best state | Best state |
| Dead | Dead | Dead |

BLOC 5 – D, A, E. *n=25*

| Task 1 | Task 2 | Task 3 |
|---|---|---|
| D1 | D1 | A1 |
| D6 | D8 | A12 |
| D7 | D9 | A13 |
| A1 | E1 | E1 |
| A10 | E6 | E8 |
| A11 | E7 | E9 |
| Best state | Best state | Best state |
| Dead | Dead | Dead |

BLOC 6 – F, B, C. *n=25*

| Task 1 | Task 2 | Task 3 |
|---|---|---|
| F1 | F1 | B1 |
| F6 | F8 | B12 |
| F7 | F9 | B13 |
| B1 | C1 | C1 |
| B10 | C10 | C12 |
| B11 | C11 | C13 |
| Best state | Best state | Best state |
| Dead | Dead | Dead |

**Figure 2** Comparison of adjusted VAS value and published value set

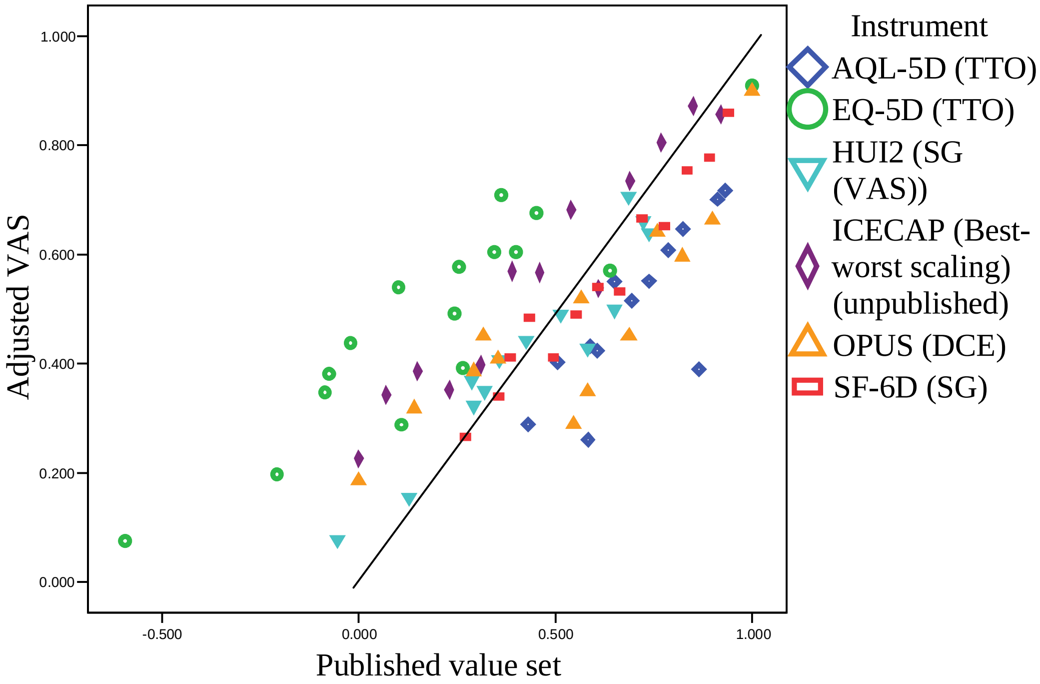
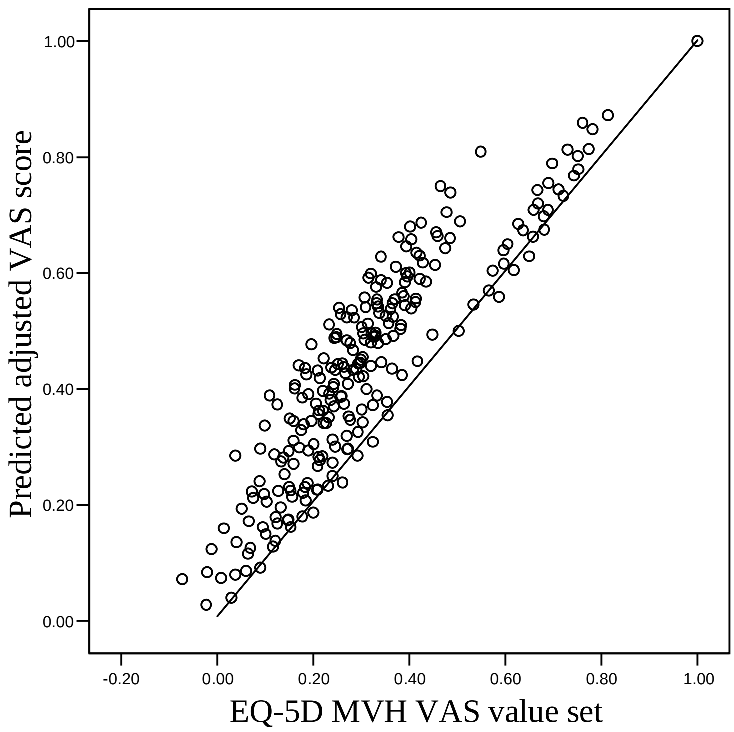



**Figure 3** Comparison of EQ-5D VAS value set for all EQ-5D states and predicted adjusted VAS score using maximum likelihood random effects regression[5]




[5] EQ-5D VAS values quoted in 'The Measurement and Valuation of Health; First report on the main survey', The MVH Group, May 1994. Values used are the VAS tariff of means: Whole population – 10 year duration.