MONASH University
Business and Economics

Centre for Health Economics

Research Paper 2007 (17)

The Assessment of Quality of Life (AQoL) II Instrument

Overview and Creation of the Utility Scoring Algorithm

Professor Jeff Richardson
Foundation Director, Centre for Health Economics, Monash University

Dr Stuart Peacock
Senior Scientist, Centre for Health Economics in Cancer, British Columbia Cancer Agency
Honorary Senior Research Fellow, Centre for Health Economics, Monash University

Mr Angelo Iezzi
Research Fellow, Centre for Health Economics, Monash University

Mr Neil Atherton Day
Principal Research Fellow, Centre for Program Evaluation, The University of Melbourne

Assoc Prof Graeme Hawthorne
Principal Research Fellow, Department of Psychiatry, The University of Melbourne

May, 2007

ISSN 1833 – 1173
ISBN 1 921187 16 6

TABLE OF CONTENTS

List of Tables

List of Figures

List of Boxes

ABSTRACT

MAU instruments seek to measure the 'utility' of health states in a way suitable for use in economic evaluation studies and, in particular, cost utility analysis (CUA). The Assessment of Quality of Life, Mark 2 (AQoL 2) project was undertaken specifically to increase the sensitivity of measurement in the region of full health, where most other instruments, including the earlier AQoL 1 instrument are relatively insensitive.

In sum, the AQoL 2 instrument estimates utility using a three stage procedure. Items are (i) weighted and combined using a multiplicative model to obtain dimension scores; (ii) these are similarly weighted and combined to obtain an initial AQoL score; (iii) this is then transformed econometrically to produce the final estimate of a health state utility.

As with AQoL 1 the research program also sought to experiment with new methods for achieving this. AQoL 1 was the first instrument to use a multi level descriptive system with five dimensions of health separately modelled and then combined. After experimentation it incorporated a new way of modelling the utility of health states worse than death.

AQoL 2 adopted this same multi level structure It was developed in 2 stages. The first used a series of confirmatory factor analysis using Lisrel, to construct dimension models. The second was a confirmatory factor (SEM) analysis of the overall AQoL which combined all of the dimensions. Utility scores were modelled in three stages. Time trade-off (TTO) importance weights were first combined into dimensions and to the dimensions into a single score using multiplicative (non stochastic) models (as with AQoL 1). However these were subsequently adjusted in a third stage econometric 'correction' based upon independently collected multi attribute – TTO – scores.

Section 2 of the paper below summarises the work published in Working Paper 144 'Overview of the Assessment of Quality of Life (AQoL) Mark 2 Project (Richardson et al 2004b). This summarises the AQoL program, outlines the methods used to obtain the AQoL descriptive system and the methods used to obtain the stage 1 and stage 2 utility algorithms based upon this system. Section 3 presents the results from these two stages of the modelling.

Section 4 of the paper outlines the methods data and results of the econometric analysis conducted to obtain a 'stage 3' adjustment or correction to improve the explanatory power of the model. It presents standard and new tests of the correction. One such task is the use of the corrected models to predict utility in an independently collected data set. Results using MA-TTO data collected for the VisQoL instrument are reported in Section 5. Model selection is discussed in Section 6.

Computation of final utilities is complex. Algorithms for the dimension and final utilities are provided in Appendix 3 . Syntax for the STATA statistical software package to carry out these calculations has been placed on the CHE website:
http://www.buseco.monash.edu.au/centres/che/

The Assessment of Quality of Life (AQoL) II Instrument: Overview and Creation of the Utility Scoring Algorithm

1 Background: The Use of Multi Attribute Utility (MAU) Instruments

Economic evaluation of health and health care related activities must quantity the *importance* of the quality of life of the outcome. This is done through the use of the Quality Adjusted Life Year (QALY) as a unit of outcome. As QALYs are defined as the product of life years and an index of the quality of life, this latter quantity requires measurement. This task is carried out by measuring the *strength of preference* for a health state relative to full health and death. The strength of preference in this context is referred to as the 'utility' of the health state.

Economic Evaluation and the Quality of Life

The quantification of 'utility' in cost utility analysis (CUA) requires two broad tasks. First, the health state under investigation must be described; secondly, a scaling technique such as the time trade-off (TTO) or standard gamble (SG) must be used to attach a numerical value to the health state such that this value measures the strength of a person's preference (utility value) for the health state.

Two broad approaches to this two stage procedure have normally been used[1], namely holistic (or 'composite') and multi-attribute utility (MAU) measurement (Torrance 1986). With the first of these, a scenario or vignette is constructed which describes the health state (Step 1). The entire scenario is then 'scaled' (Step 2), ie a survey is conducted specifically to elicit 'utility' values for the scenario.

With the second approach a generic 'descriptive system' or 'descriptive instrument' is created which is capable of describing a wide range of health states and utility weights are attached to every possible state. This is normally achieved by measuring a limited number of health states and using these to calibrate a model which is then used to infer the utility values of every other health state in the 'descriptive system'.[2] The model may either be derived by econometric analysis of the observed utilities (as with the EuroQoL/EQ5D (Williams 1995)) or through the use of decision analytic techniques to fit the simple additive model (as used in the Quality of Wellbeing Instrument (QWB) (Kaplan et al 1996) and 15D (Sintonen and Pekurinen 1993)) or a multiplicative model (the Health Utilities Index (HUI1 and 2) (Feeny, Torrance et al 1996). The fully scaled MAU instrument may then be used to estimate the utility of health states.

Both approaches have strengths and weaknesses. Holistic measurement permits a description which is tailored to a particular health state. Unique aspects of the health state, its content, its consequences, the process of health delivery, risk and prognosis may all be included in the vignette. Validation of health state specific vignettes, however, is seldom, if ever, carried out. By contrast, the generic descriptive system of the MAU approach may be unable to capture many of the nuances of the health state and be incapable of capturing the importance of the process or

---

1 In principle, these two steps can be collapsed by asking patients directly the value of the health state that they are currently experiencing. In practice this approach has seldom been used.

2 In principle every health state may be individually measured. In practice, the number of health states in the 'descriptive system' is normally so large that this is infeasible. The only example of this approach is the original Rosser Kind Index which is now seldom used because of its limited sensitivity.

context of the health state or intervention. However, this approach should, in principle, be based upon a descriptive system, the reliability and validity of which can be investigated using standard procedures.[3] After construction, the use of an MAU instrument is cheap and easy and allows the rapid estimation of utilities in the context of a longitudinal trial. This means that it is feasible to construct a time profile of each of the dimensions of health included in the instrument. Because of these respective strengths and weaknesses both techniques have a role in CUA.

To date, only a handful of generic instruments have attempted to measure utility; *viz,* the UK Rosser-Kind Index (Rosser 1993), the US QWB (Kaplan, Ganiats et al 1996), the Canadian HUI instruments (Feeny, Torrance et al 1996), the Finnish 15D (Sintonen and Pekurinen 1993) and the European Euroqol (Kind 1996). Extant descriptive systems (questionnaires) have been used to create utility from non utility instruments (Brazier et al 1998). While each of these instruments has particular strengths, to our knowledge none were constructed using normal psychometric principles to ensure construct validity and structural independence.

Consider, for example, this second issue. MAU theory postulates there should be no 'redundancy' amongst items in a descriptive system. That is, a single attribute should not be described in more than one way (von Winterfeldt and Edwards 1986). If redundancy occurs then the (dis)utility of the attribute will be double counted. A sufficient (but not necessary) condition for non-redundancy is that the different scales within the instrument are orthogonal.[4] However, the requirement of non redundancy appears to be in conflict with the need for 'sensitivity' and several instruments have reduced redundancy by the adoption of very simple descriptive systems; but this simplicity has been achieved at the expense of sensitivity. Other problems also exist. Some instruments have unsatisfactory models for inferring utility values; others have adopted scaling techniques which probably do not measure utility (Richardson 1994). Consequently there is a challenge to develop a generic instrument which overcomes these weaknesses.

---

3 HRQoL is a psychological concept. It cannot be directly measured, but is uniquely individual. Although instruments can be developed from other measurement – traditions such as clinometrics, economics or decision-making – this property of HRQoL suggests that the application of psychometrics is particularly appropriate during construction of the descriptive system.

4 It is not strictly necessary as scales may be 'environmentally correlated', which does not necessarily indicate double counting. Von Winterfeldt & Edwards illustrate this in the case of a manufacturing plant, the management of which is concerned with the cost of production and distribution (von Winterfeldt and Edwards 1986). These costs will correlate because each correlates with the scale of production. Despite this, there is no redundancy and each attribute is independently important. Even with this example, however, careful construction of the instrument can eliminate the correlation. There is no necessary reason why scale of production, *unit* production costs and *unit* distribution costs will correlate (if there are no economies of scale).

2 AQoL 1 and AQoL 2

The AQoL project was designed to assist with meeting this challenge. Specifically it sought to create an instrument where the descriptive system was:

- derived from psychometric procedures for instrument construction and which therefore has construct validity;

- sensitive to as much of the full universe of HRQoL as is practical;

- based upon structurally independent dimensions of health;
- based upon a description of ‘handicap’—problems in a social context—as distinct from a ‘within the skin’ descriptive system;
- maximised sensitivity while minimising the effects of redundancy upon final utility scores.
- Used the arguably most acceptable metric for measuring utility, the ITO IRichardson 1994, 2002)

It was the attempt to achieve this final objective – sensitivity without inflated disutilities – which initially led to the adoption of a hierarchical descriptive system. This permits redundancy – double counting of disutilities – within dimensions to achieve instrument sensitivity, but structural independence between dimensions. As discussed below the structure was not fully successful in achieving its initial objective and it was for this reason that a three part modelling of utilities was adopted for AQoL 2, where the third stage is an econometric correction to eliminate the effects of redundancy. Nevertheless the hierarchical modelling is independently valuable as it provides disaggregated descriptive information by the different dimensions of health included in the model.

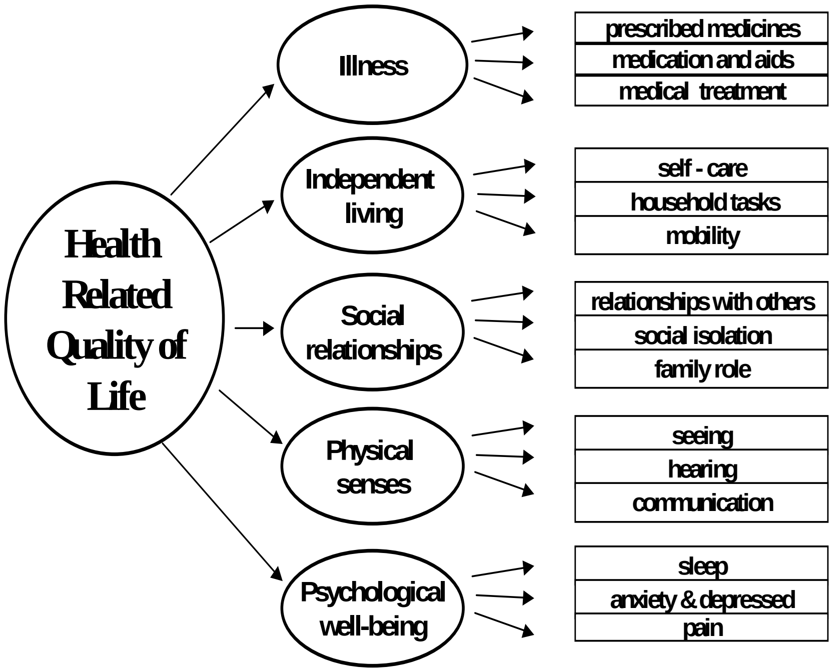
The procedures adopted in this part of the project resulted in an instrument which is unique in two respects: *viz,*

(i) use of an hierarchical descriptive structure in which structural independence is achieved between dimensions but not within dimensions; and

(ii) a descriptive system which can claim to have construct validity, which increases confidence in the validity of the health state descriptions.

The structure of AQoL 1 is described in Figure 1 and the instrument construction is described in Hawthorne, Richardson and Day (1997)

Figure 1 Structure of the AQoL 1 Instrument



A large five instrument validation study was carried out to compare AQoL with the HUI 3, EQ5D, 15D and SF36 (first Brazier weights). Results are described in Hawthorne et al (2001). The major conclusion from this study was that the very low correlation between instruments raises serious doubts about their validity. AQoL 1 performed as well as other instruments in terms of its predictive power and showed greater upper level sensitivity, ie in the range of full health. However, along with the Health Utilities Index, the magnitude of disutilities appeared very high. This is not necessarily a problem for ranking the quality of life of different health states. It does, however, cast doubt upon the validity of absolute QALY numbers.

AQoL 2

AQoL 2 sought to improve AQoL 1 in several respects. These were:

(i) to increase in the sensitivity of the descriptive system in the region of full health and a descriptive system which permitted the evaluation of health promotional activities;

(ii) to create two scaling systems based upon the time trade-off (as with AQoL 1) and the person trade-off (PTO) scaling methodologies. The appropriate choice of scaling instrument has not been determined in the literature;

(iii) to rework of the utility scores employing techniques to test and, if necessary, eliminate one possible source of bias in previous methodologies (including AQoL 1), viz a 'focusing effect' (Ubel and Loewenstein, 2001) (Richardson 1994, 2002);

(iv) to test and use 'deliberative weights' which permit and encourage the contemplation of the health states over a significant time period;

(v) to use a flexible three stage modelling methodology to combine disaggregated dimension scores into a single index of utility which minimises the effects of redundancy and preference dependency.

The different approaches to this problem in AQoL 1 and AQoL 2 were a consequence of the experience with AQoL 1. As noted, in the five instrument comparative study it was found that AQoL 1 and the Canadian Health Utilities Index (HUI 3) systematically produced lower scores than in other instruments. These were also the two instruments which had used a multiplicative model for combining the disaggregated scores. This approach to modelling has the great advantage of being able to combine a very large number of items and item responses. This contrasts with an econometric model (EQ5D) where the necessary sample size of respondents needed to produce valid TTO scores sharply rises with the number of possible health states. However the results above suggest that AQoL 1 did not successfully eliminate redundancy and that the simple multiplicative model with less than perfect item orthogonality may produce higher disutility scores when more opportunities arise for detecting disutility in a health state, ie when additional items are included in the instrument.

AQoL 1 sought to overcome this problem by seeking sensitivity within dimensions, ie relaxing the assumption of orthogonality – but minimising the effect of this. This was done by first achieving orthogonality between dimensions and, secondly, independently measuring the maximum disutility of each dimension and constraining dimension scores to be less than or equal to this dimension disutility. That is, redundancy within AQoL 1 dimensions cannot result in a disutility greater than the independently measured dimension disutility. However, this strategy may have been compromised in AQoL 1. In order to allow sensitivity within dimensions a significant level of redundancy was permitted. This will have the effect of depressing dimension scores. The *within dimension* redundancy is not prevented by the orthogonality *between* dimensions. Combining depressed dimensions will produce an overall depressed estimate of global disutility.

An additional factor is likely to depress the scores obtained from a large multi attribute model. To obtain the importance weights for items their worst level is typically compared with death or the dimension/instrument 'all worst' health state, where every item is at its worst value (AQoL used death as the lower boundary for dimension measurement). As described by Ubel and Loewenstein (2001) there may be a 'focussing effect' which results in an exaggerated utility score for single items as respondents to the scaling survey focus upon the disutility of the item and forget other positive elements of the health state. AQoL 2 tackled this source of error through its presentation and use of interview props (Richardson et al 2004).

To eliminate these residual effects AQoL 2 adopted a three stage modelling strategy. In the first two parts the multiplicative modelling of dimension and overall DU scores, the requirement of orthogonality was again dropped and the criterion for instrument selection became its psychometric properties, ie the extent to which an item contributes to the dimension description (a requirement which implicitly limits – but does not seek to explicitly eliminate – double counting).

Elimination of the effects of non orthogonality is achieved in the third stage of the modelling when scores from the multiplicative modelling, are used to econometrically explain TTO values obtained from the independently collected multi attribute health states constructed from the AQoL 2 descriptive system. The final AQoL 2 algorithm uses the best fitting equation.

This procedure ensures that the predicted algorithm produces scores comparable with the scores of the multi attribute health states used in stage 3 modelling. The effects of a generalised redundancy – low level double counting – across all of the dimensions can be eliminated with a simple model with only the second stage AQoL score as an explanatory variable. Use of a power function ensures that the transformation passes through the point of best and lowest disutility ie (0,0) and the point of maximum disutility (1,1). The procedure can take account of dimension specific redundancy by the inclusion of dimension dummy variables. Importantly, it can also take some account of preference dependency where the score for one dimension varies with the level (and therefore score) of a second dimension. This problem is mitigated by the inclusion of interaction terms for combinations of dimensions.

Stages of the Analyses

In broad terms AQoL 1 and AQoL 2 have a very similar methodology. This involves the following steps:

(i) Selection of a 'theory' relating to the appropriate form of the construct 'health'. The theory determines the contents of the 'descriptive system'.

(ii) Creation of an item bank, ie a large number of alternative items which describe the construct;

(iii) The creation of a descriptive system using items from the item bank and an analytical process to determine the best items for the description of the health construct;

(iv) The construction of a 'scaling system'; viz, a formula to attach a utility score to every health state described by the descriptive system.

Both AQoL 1 and AQoL 2 sought to create a descriptive system based upon handicap; that is, a description of health states in terms of their effect upon individuals in a social context and the impact upon individuals functioning in a social context.

AQoL 2 sought to describe health states in a way that would be relevant for the evaluation of health promotional interventions. With both instruments a databank was constructed from the existing literature and from a series of interviews with health care providers and patients .

The final selection of items for inclusion in the AQoL instruments were based upon a two part analysis. In the first, items were subject to logical and linguistic analysis to determine the appropriateness of content and grammar. A reduced number of items were then included in a survey in which respondents completed every item. These were then subject to exploratory then confirmatory factor analysis for AQoL 1 and structural equation modelling procedures in AQoL 2. These techniques group items according to the pattern of responses and evaluate the 'explanatory power' of different groupings with respect to the empirically obtained patterns of responses.

Finally, the AQoL 2 project was designed to permit respondent deliberation about health states. Key parameters were collected on each of two occasions separated by at least 2 weeks. In the interim one half of respondents were asked to complete a number of tasks designed to evoke deliberation about reducing their length of life in exchange for a better quality of life. Subsequent comparison of results are reported in Peacock et al (2004).

3 Utility Results for Stage 1 and Stage 2 Modelling

Two postal surveys and two sets of interviews (with the same respondent) were conducted. The first postal survey was for the completion of items in the item bank. The two sets of interviews were carried out to obtain time trade-off, person trade-off and self TTO data.

The second postal survey was used in preference to additional time trade-off interviews because of the cost of individual interviews and the burden already placed upon interview respondents. The strategy required the subsequent transformation of rating scale into time trade-off scores. (See Richardson et al. 2004b).

The number of respondents and response rates are shown in Table 1.

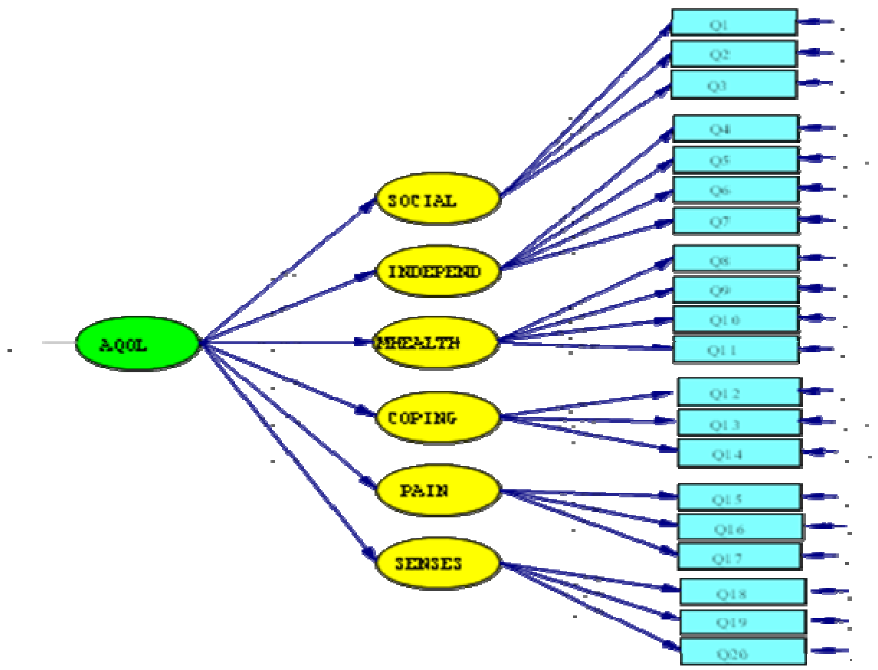
Table 1 Data collection for AQoL 2

Purpose	respondents n	response rate %
Postal survey 1: completion of items in item bank	316	44
Interview 1:= dimension worsts, multi attribute health states (TTO scores)	411	47
Interview 2 Multi attribute health states (cont) Person trade-off, self TTO	411	47
Postal Survey 2: Item responses, item worst scores (Rating Scale)	163	40

AQoL 2 Descriptive System

Construction of the descriptive system is described in Richardson et al 2004b. The AQoL descriptive system is reproduced as Appendix 1 and its structure is depicted in Figure 2 (below).

Figure 2 Structure of AQoL 2



R
MS
EA = 0.054, CFI = 0.97

The instrument consists of 6 dimensions and 20 items. Each of these had between 4 and 6 response levels. The model was built in two stages using confirmatory factor analysis (a particular application of the general SEM technique which focuses upon locating latent variables). First each dimensions was developed analysing the items that were anticipated to be related to that particular domain (e.g. social, independent living etc.) The construction questionnaire had many items intended to measure each dimension, so this investigation explored correlated errors, cross-loading items and looked for any structure within the dimension that might indicate a subscale. The end point of this stage of the analysis was selection of a small number of items that comprehensively but parsimomiously measured the dimension under development. The second stage of the analysis related to overall quality of life; the AQoL. **Figure** 2 shows the final result of this analysis which provides a model that closely fits the data from which it is derived. The CFI of 0.97, that is, is considerably higher than the accepted value of 0.90 that indicates a satisfactory fit. The RMSEA of 0.054 is well below 0.08, generally accepted as the minimum level for satisfactory fit. This is an exceptionally good result which underpins the validity of the model as a representation of the structure underlying the data from our construction sample. (Brown M and Cudeck R 1993,Yu 2002)[5]

Table 2 compares this structure with the structure of AQoL 1. AQoL 2 subsumes AQoL 1 both in terms of the range of health states covered by each item and by the elements of health embodied

[5] Yu, 2002, investigates goodness of fit indices where data deviate substantially from normality and recommends CFI > .95 and RMSEA around .5 top .6 as providing acceptable Type 1 (5) and Type II errors

in the instrument. In particular AQoL 2 gives additional items relating to coping and pain. It does not include items relating to illness as did the initial AQoL instrument but which were subsequently removed during the validation process.

The descriptive system defines 6.58 x 10^{13} (or about 13,000) health states for every human being on the planet). The overwhelming majority of these are irrelevant as they consist of implausible combinations of health states. Nevertheless the number indicates that the instrument achieves an unprecedented level of detail with respect to the dimensions of health which it encompasses.

Table 2 Comparison of AQoL 1 and AQoL 2

Dimension	AQoL 1*		AQoL 2	
	Items	Combinations	Items	Combinations
1 Independent living	3 (444)	64	4 (5665)	900
2 Social relationships	3 (444)	64	3 (544)	80
3 Physical senses	3 (444)	64	3 (664)	144
4 Psychological wellbeing	3 (444)	64	4 (5555)	625
5 Pain	-		3 (445)	80
6 Coping	-		3 (555)	125

* Following a number of validating studies the illness dimension was removed from the final scoring algorithm

AQoL 2 Scaling System ('Tariffs')

The first stage of the scaling of AQoL 2 consisted of the use of a multiplicative model. This is similar to equation 1 below.

$$U(AQoL) = U_1 * U_2 * U_3 * U_4 * U_5 * U_6 \ldots (1)$$

The actual model is somewhat more flexible. It is calculated using disutilities rather than utilities and these are adjusted for the relative importance of each of the dimensions (1)-(6). This results in equation 2 in which w_i are the dimension (or item) weights and k is the overall scaling constant. This is obtained by solving equation 3. It is similar to the requirement in an additive model that the dimension weights sum to unity.

$$DU = \frac{1}{k}\left[\prod_{i=1}^{n}\left[1 + kw_i DU_i\left(x_{ij}\right)\right] - 1\right] \quad \ldots (2)$$

$$k = \prod_{i=1}^{n}(1 + kw_i) - 1 \quad \ldots (3)$$

The relationship between utility and disutility is given in equation 4.

$$U^* = 1 - DU^* \quad \ldots (4)$$

The model was applied at two levels; first, to combine items into dimensions and, secondly, to combine dimensions into the overall AQoL score. Table 3 reports item response utilities on a (1-0) scale where the endpoints of the scale are the item best (DU = 0.00) and worst (DU = 1.00) response levels. Key parameters were calculated using data from the second of the 2 interviews at which they were elicited. DU scores were somewhat lower than obtained in the initial interview indicating that deliberation and the passage of time had, as postulated, resulted in some demunction of the perceived disutility of health states.

In Table 4 the individual item importance weights, w_i, are calculated as the item worst score (on a scale from dimension best (0.00) to dimension worst 1.00) health state). These are multiplied by the dimension scaling factor (k_d) which is derived from the item weights and from equation 3 above. The overall or net item weight is used to construct the dimension formulae shown in Box 1.

Similar results are shown for the overall AQoL in Table 5. In this w_d represents the dimension importance weights measured on an AQoL best (DU = 0.00) to death scale (DU = 1.00). The AQoL scaling constant, k, is derived from these six weights. The product of the weights and the scaling constant give the effective weights W_d. It is used to derive the overall AQoL formula in Box 1.

Table 3 Item Disutilities (TTO Scores) for use in Dimension Models

	Dimension 1	Dimension 2	Dimension 3	Dimension 4	Dimension 5	Dimension 6
Response Level	**Mean**	**Mean**	**Mean**	**Mean**	**Mean**	**Mean**
	Item 1	**Item 5**	**Item 8**	**Item 12**	**Item 15**	**Item 18**
1	0.00	0.00	0.00	0.00	0.00	0.00
2	0.07	0.07	0.13	0.06	0.13	0.03
3	0.44	0.46	0.39	0.34	0.64	0.22
4	0.82	0.84	0.84	0.72	1.00	0.62
5	1.00	1.00	1.00	1.00		0.84
6						1.00
	Item 2	**Item 6**	**Item 9**	**Item 13**	**Item 16**	**Item 19**
1	0.00	0.00	0.00	0.00	0.00	0.00
2	0.03	0.19	0.14	0.06	0.20	0.02
3	0.24	0.76	0.39	0.38	0.76	0.21
4	0.47	1.00	0.83	0.77	1.00	0.59
5	0.84		1.00	1.00		0.83
6	1.00					1.00
	Item 3	**Item 7**	**Item 10**	**Item 14**	**Item 17**	**Item 20**
1	0.00	0.00	0.00	0.00	0.00	0.00
2	0.04	0.20	0.10	0.06	0.07	0.19
3	0.25	0.65	0.33	0.42	0.34	0.70
4	0.57	1.00	0.78	0.83	0.75	1.00
5	0.83		1.00	1.00	1.00	
6	1.00					
	Item 4		**Item 11**			
1	0.00		0.00			
2	0.04		0.06			
3	0.30		0.37			
4	0.80		0.84			
5	1.00		1.00			

Notes: Item best and worst disutilities are set equal to 0.00 and 1.00 respectively. Item best and worst responses are set as endpoints for rating scale evaluations.

Table 4 Item Weights for use in Dimension Models

Dimension \ Item	(-) k_d w_i	= wt_i	Dimension \ Item	(-) k_i w_i	= wt_i
Ind Living			Coping		
1	(0.978) * (.39)	= 0.38	1	(0.930) * (.42)	= 0.39
2	(0.978) * (.59)	= 0.59	2	(0.930) * (.64)	= 0.60
3	(0.978) * (.63)	= 0.63	3	(0.930) * (.77)	= 0.72
4	(0.978) * (.79)	= 0.79			
Social & Family			Pain		
1	(0.923) * (.64)	= 0.64	1	(0.902) * (.63)	= 0. 57
2	(0.923) * (.70)	= 0.70	2	(0.902) * (.77)	= 0.69
3	(0.923) * (.51)	= 0.51	3	(0.902) * (.65)	= 0.59
Mental Health			Sensory		
1	(0.983) * (.64)	= 0.64	1	(0.851) * (.58)	= 0.49
2	(0.983) * (.59)	= 0.59	2	(0.851) * (.46)	= 0.39
3	(0.983) * (.65)	= 0.65	3	(0.851) * (.60)	= 0.51
4	(0.983) * (.71)	= 0.71			

Table 5 Dimension Weights for use in AQoL Model

Dimension	(-) k_d	x w_d	=	$wt_d^{(i)}$	*Alg*
1 Ind. Living	0.965	x (.47)	=	0.454	0.472
2 Social	0.965	x (.45)	=	0.434	0.448
3 Mental Health	0.965	x (.48)	=	0.455	0.479
4 Coping	0.965	x (.35)	=	0.338	0.345
5 Pain	0.965	x (.59)	=	0.569	0.592
6 Senses	0.965	x (.64)	=	0.618	0.637
AQoL	W/k 1.102/.954	=	=	1.15	

Key
k_d = Dimension scaling constant
w_d = Dimension weight = dimension all worst
wt_d = Final dimension weight
(i) = Final colum corrected for rounded decimal places, in col 1, 2

Box 1 Multiplicative Disutility Equations

Dimensions	
General Formula	$DU_d = \frac{1}{k}\left[1 - \prod_{i=1}^{n}\left(1 - kw_i DU_i\right)\right]; k_d > 0$
Independent Living	$DU_1 = 1.02\left[1 - (1 - 0.38du_1)(1. - 0.58du_2)(1 - 0.62du_3)(1 - 0.77du_4)\right]$
Social and Family	$DU_2 = 1.08\left[1 - (1 - 0.59du_5)(1. - 0.65du_6)(1 - 0.47du_7)\right]$
Mental Health	$DU_3 = 1.02\left[1 - (1 - 0.63du_8)(1. - 0.58du_9)(1 - 0.64du_{10})(1 - 0.70du_{11})\right]$
Coping	$DU_4 = 1.08\left[1 - (1 - 0.39du_{12})(1. - 0.60du_{13})(1 - 0.72du_{14})\right]$
Pain	$DU_5 = 1.08\left[1 - (1 - 0.57du_{15})(1 - 0.39du_{16})(1 - 0.59du_{17})\right]$
Senses	$DU_6 = 1.18\left[1 - (1 - 0.49du_{18})(1 - 0.39du_{19})(1 - 0.51du_{20})\right]$
General Formula	$DU_{AQoL} = \frac{W}{k}\left[1 - \prod_d \left(\left(1 - kw_d DUx_i\right)\right)\right]; k > 0$
$DU_{AQoL} = 1.150\left[1 - (1 - 0.454DU_1)(1 - 0.448DU_2)(1 - 0.479DU_3)(1 - 0.345DU_4)(1 - 0.592DU_5)(1 - 0.637DU_6)\right]$	

Key: W = the conversion factor between the 0-1 (death, full health) model

Each of the 6 dimension and the AQoL disutilities may be transformed into utility scores using the equation $U_i = 1 - DU$ where U_i and DU_i are utility are disutility respectively.

Scaling the Multiplicative Model: An Example

Assigning a utility score to a health state involves the following steps.

(i) Completing the AQoL questionnaire and determining the 20 response levels which define the health state.

(ii) Reading the 20 item disutility scores, du_i, which correspond with the response levels from Table 3. These ‘disutilities’ are measured on a (0-1) scale with the item best (0.00) and worst (1.00) defining the endpoints.

(iii) Entering the item disutility scores, du_i, into the corresponding equation in Box 1. Calculate the six dimension disutility scores DU_d. These disutilities are measured on a (0-1) scale where the endpoints are the dimension best and dimension ‘all worst’ (all items at their worst level).

(iv) Entering the six dimension DU_d scores into the final AQoL equation in Box 1. The score obtained is the predicted disutility for the health state from the stage 1 multiplicative model.

(v) Converting disutilities into utilities using the equation $U = 1 - DU$.

The five steps are illustrated for a randomly chosen health state in Box 2.

Box 2 Calculating a utility score: A numerical example

1 Complete the AQoL questionnaire to obtain 20 response levels; 1 per item

Example: Response levels are:
D 1(1,1,2,1); D 2(2,2,3); D 3 (3,2,3,1); D 4(1,1,1); D 5(2,1,1); D 6(2,1,2)

2 Read the 20 disutility scores from Table 3

In the example:

$$D1(0,0.04,0) : D2(.07,.19,.65) : D3(.39,.14,.33,.00) D4(0,0,0) : D5(.13,.0,.0) D6(.03,00,.19)$$

3 Enter the 20 disutility scores into the equations in Box 1

$$DU_1 = 1.02[1-(1-38*0)(1-.58*0)(1-.62*.04)(1-.77*0)] \quad = 0.03$$
$$DU_2 = 1.08[1-(1-.59*.07)(1-.65*.19)(1-.47*.0) \quad = 0.17$$
$$DU_3 = 1.02[1-(1-.63*.39)(1-.66*.14)(1-.64*.33)(1-.7*0) \quad = 0.40$$
$$DU_4 = 1.08[1-(1-.39*0)(1-.60*0)(1-.72*.0) \quad = 0.00$$
$$DU_5 = 1.08[1-(1-.69*.13)(1-.57*0)(1-.57*.0) \quad = 0.10$$
$$DU_6 = 1.18[1-(.1-.4*0.03)(1-.39*0)(1-.51*.19) \quad = 0.12$$

4 Enter the DU_i scores into the AQoL formula Box 1

$$DU_{AQol} = 1.15[1-(1-.472*.03)(1-.448*.17)(1-.479*.4)$$
$$(1-.345*0.0)(1-.592*.1)(1-.637*.12)] = .42$$

5 Convert disutility to utilities from the equation $U = 1 - DU_i$
Dimension Utilities = 0.97; 0.83; 0.6; 1.00; 0.9; 0.88
Global U = 0.58;

4 Stage Three Econometric Correction: Methods, Data and Results

The purpose of the stage 3 econometric ‘correction’ was to compensate for errors introduced by the multiplicative model which might arise because of structural or preference dependency or because of the possibility of a focusing effect in responses obtained in the scaling interview. The third stage modelling is based upon the assumption that the TTO score for a holistic (multi attribute) health sate (MA) TTO represents the ‘gold standard’; - the procedure for combining attributes (items and dimensions) should result in the same estimate of utility as obtained by the evaluation of the multi attribute health state defined by the same attributes.

Methods

If it is no longer assumed that the multiplicative model is without bias then there is no clear guideline concerning the relationship between the stage 2 AQoL and the gold standard.

It cannot be assumed, for example, that the same functional relationship between (MA) TTO and AQoL scores would exist at high levels and at low levels of disutility. Rather, the functional relationship could adjust the stage 2 AQoL score at least one of four forms: (i) uniform adjustment of the overall score/error in all dimensions; (ii) adjustment of a single dimension; (iii) correction for redundancy and preference dependence between two or more dimensions; (iv) structural change in the relationship at different levels of disutility. This implies the need for what might be described as ‘loose cannon’ modelling – the empirical exploration of alternative relationships. Only two general constraints exist. First, the functional relationship must produce utility (TTO) scores between 1.00 and 0.00 when AQoL scores have the value 1.00 and 0.00 respectively. These scores correspond with best health and death respectively. Secondly, a positive AQoL increment should result in a positive increase in utility.

To achieve this degree of flexibility the relationship in equation 5 was employed.

$$TTO = AQoL^x$$

$$x = \alpha_0 + \sum_j \alpha_i D_i + \sum_i \sum_j B_{ij} D_i D_j + \sum_{i=1}^{4} shift_i \qquad \text{... (5)}$$

Where
- α_0 = constant
- D_i = dimension score for dimension i
- $D_i D_j$ = dimension D_i times Dimension D_j
- $shift_i$ = dummy variables indicating that the TTO has a disutility score in excess of 0.2 0.4 0.6 0.8

If the multiplicative model had no bias then the coefficient α_0 would equal unity and other coefficients would be insignificant. If α is not equal to 1.00, a uniform exponential adjustment is made to all DU values. D_i coefficient adjust for any net bias in individual dimensions; D_i D_j coefficients adjust for interaction between dimensions and the shift coefficients included eliminate any residual over or under estimation of the true TTO values at particular levels of utility. The power function relationship ensures that the predicted score passes through the points (0.00, 0.00) and (1.00, 1.00). As the modelling is initially conducted in disutility space these points represent best health and death respectively in the results below. As the exponent may be negative it is possible for a disutility score to be in excess of 1.00, indicating a health state worse than death.

Data and Econometric Modelling

Three sets of multi attribute health state data were available. As part of the TTO interviews 411 respondents were asked to evaluate 3 or 4 multi attribute health states selected from 18 holistic ('e'-type) health states which were, in turn, constructed from the AQoL descriptive system. These were constructed to include interaction between all combinations of dimensions. From the 365 useable interviews 1042 multi attribute health state valuations were obtained ('MA.TTO's'). This relatively small sample size was dictated by the research budget. Secondly, the MA.TTO scores were 'deconstructed' to form 'pseudo TTO' health states which describe less severe symptoms and thereby increase the sensitivity of the results to health state values close to full health.[6] Thirdly, a single 'D type health state' value was elicited form each interview respondent. This was the MA (TTO) score for the AQoL all worst health state. All MA (TTO) states were measured on a full health (0.00)-death (1.00) scale

The different combinations of these data were employed with the various combinations of the model variables in equation 5. This resulted in 40 basic models: 4 data sets each with 10 combinations of variables which are possible from the variable sets in equation 5

Stage 3 regressions were estimated using a random effects (RE) model[7] and initially evaluated using the conventional Wald statistic. Additional criteria to distinguish between models are outlined below. 'Short list' models were used to predict MAU (TTO) scores independently calculated for the validation of an MA instrument for measuring visual acuity, the VisQoL

Results

Table 6 reports the results for three sets of equations estimated with three sets of data. Set A includes only the stage 2 estimates of AQoL as an explanatory variable. Set B also includes dimension dummy variables D_i and dimension interaction terms, D_iD_j. Set C adds the shift dummy variables defined in equation 5. The three equations in each set correspond with the use of observed multi attribute TTO data only; the inclusion of pseudo MA data (P) and the addition of instrument all worst scores, D.

Several conclusions may be drawn from Table 6. First, results confirm the expectation that utilities observed in the level 2 multiplicative model inflate DU scores. All models predict disutilities less than predicted by the multiplicative model. Secondly, this effect was not uniform across the dimensions. Third, interaction terms were significant in regressions which included them. Finally, the choice of data set had a significant effect upon results, with the larger data bases producing better fitting models.

Table 7 reports the correlation between the TTO scores for MA health scenarios, predicted by the 9 models, and the correlation with the stage 1 AQoL score. Unsurprisingly these are very high and do not provide a basis for discriminating between models.

Selection Criteria: In addition to the Wald Chi 2 statistics, a number of procedures were undertaken to test the goodness of fit of the models. These were (i) the 'internal predictive

---

6 For example, with 3 dimensions $E(U_1\ U_2\ U_3)$ PE_1 and PE_2 are created as $(U_1\ 0\ U_3)$ and $(0\ U_2\ 0)$. Scores were assigned by pro rata allocation of the DU of state E between the two states in proportion to the two MA scores for PE_1 and PE_2 derived from the AQoL 2.0 (multiplicative) model.

7 Since respondents were asked to answer between 3 and 6 E type questions an RE model was used to correctly account for clustering in the data.

power’; - a quasi diagnostic test to test the explanatory power of stage 3 estimates but, more importantly, to determine if these estimates provide an unbiased estimate of the TTO scores from which they were constructed; (ii) analysis of absolute errors; (iii) comparison of predicted and actual change between health states; (iv) external validation – comparison of actual and predicted scores from an independently collected dataset.

Table 6 Stage 3 regression results: MA-TTO on AQoL stage 2 predicted values

	Set A			Set B			Set C		
	M1	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**	**M8**	**M9**
	E (1)	**EP (1)**	**EPD (1)**	**E (1)**	**EP (1)**	**EPD (1)**	**E (1)**	**EP (1)**	**EPD (1)**
n	1042	4136	4501	1042	4136	4501	1042	4136	4501
Wald Chi 2(1)	655	2156	4368	908	2764	5087	1451	6917	8677
AQoL (α)	1.84	1.15	1.51	2.07*	1.14	1.46	1.82	1.46	1.45
D1						0.52*			
D2					0.24*	0.43	-2.54		
D3							-1.97		
D5				-1.5			-1.29*		
D6								0.92	0.70
D_1D_2				-13.4	-3.89*	-3.97	-21.49	-2.84	-4.69
D_1D_5				6.9					
D_1D_6				19.1		8.47			
D_2D_4								2.57*	
D_2D_5								-4.88*	
DD									-1.42**
34 D_3D_4						2.57*			-2.23*
36 D_3D_6				-27.5	-13.58	-4.02*	-10.52*	-18.83	
L1									
L2							0.99	0.42	0.42
L3							2.53	1.16	1.10
L4							6.42	2.89	2.68
									5.30

Key
(1) All coefficients significant at 0.000 level unless designated
* significant to 0.005 level
** significant at 0.05 level

Stage 4 Internal Predictive Power

To conduct this test the predicted TTO score form the stage 3 models reported above were used as the (only) independent variable in a linear regression to explain MA (TTO) scores. Since RE modelling should result in an unbiased estimate of the dependent variable, OLS modelling should be used to test for bias. For each model the constant term was initially suppressed. An unbiased model would result in b = 1.00 in equation 6 below:

In the second regression a constant term was included to test the (more demanding) null hypothesis: that a = 0.00; b = 1.00 in equation (7) below

$$\text{MA (TTO)} = b.M_i + e \qquad ...(6)$$

$$\text{MA (TTO)} = a + bM_i + e \qquad ...(7)$$

Where M_i = stage 3 utility predicted by model M_i

The two regressions test whether or not models provide an unbiased estimate of average and marginal TTO scores respectively. The second of these is particularly relevant for the estimation of QoL changes effected by a health program.

Tables 8.1-8.4 present results for model 1 using individual (E + D) data – Table 8.1; and for model 2 using individual (E + D) data mean (E + D) data and mean (E) data – Tables 8.2, 8.3 and 8.4 respectively.

Coefficients for 'b' in Table 8.1 indicate a close correspondence between average predicted and average observed data except for model 4 and 8. While a linear regression constrained to pass through the origin may have a slope close to unity, this need not occur when a constant is included. Results in Table 8.2 reflect this. Slope coefficients are less than unity and suggest that stage 3coefficients may underestimate the full effect of a change in model 1 scores upon predicted TTO. Models 4 and 7 perform particularly badly on this test with model 9 having the least bias.

When mean data were used for this test (Tables 8.3, 8.4) model 7 and model 8 predicted low R^2 coefficients; models 1 to model 3 had relatively small b coefficients in the range 0.69-0.73. However R^2 coefficients in the range 0.90-0.95 indicate high explanatory power. Once again M9 provided the best estimate, explaining 90 percent of variation and with a b coefficient within 6 percentage points of the gold standard value of 1.00.

Results of a further test are reported in Table 9. For each of the five best fitting models the difference between various combinations of observed TTO scores were subtracted from the difference predicted by model scores, ie that is, observed (TTO_i - TTO_j) was subtracted from predicted ($AQoL_i$ - $AQoL_j$). The reported result is therefore the absolute error from single model scores as compared with gold standard (directly elicited) TTO scores when the difference in utilities between two health states is estimated. Table 9 reports the difference for 18 pairs of health states[8] for which this test could be conducted (signs are ignored).

[8] Because different combinations of health states were administered to different respondents there were relatively few health states where a comparison of this form could be made.

Results do not clearly indicate the relatively goodness of fit of the five models. M9 resulted in the smallest number of relatively large errors with M6 and M3 following in that order. However the chief conclusion to be drawn from the table is that the pattern of errors created by the different models differs and some of the real incremental changes are very different from estimates. This suggests that the database may have underrepresented these health states in the interview protocol or that co-linearity may have resulted in selected error. It is difficult to draw stronger conclusions from this analysis and, particularly how AQoL compares with other instruments with respect to this task, as the test is, to our knowledge unique in the literature.

VisQoL Validation Study, Model Selection

While AQoL is general in its scope, like all such instruments it is more or less suited to particular health states. As noted earlier comparisons of existing MAU instruments indicate very different health state utility scores for the same health state. This is primarily attributable to differences in the descriptive instruments. For this reason it is desirable to 'validate' an instrument in each context in which it is used or, better, to modify and selectively extend or recalibrate an instrument if it is to be extensively used in that particular context. At the time of writing (April 2007) three extensions to the AQoL have been completed or are underway. The target population for these instruments are, first, obese children in Fiji, Tonga, New Zealand and Australia; secondly, psychiatric patients (PsyQoL) and third the visually impaired (VisQoL). From this latter study TTO values were obtained from visually impaired patients and from a representative cross section of the remainder of the adult, English speaking community. Health states included a number relating specifically to impaired vision to the full AQoL 2 and to the AQoL 1 MA health states. In total 752 MA utility scores were available for analysis. These were used in the 'validation' study reported below. Details of the construction of the descriptive system are given in Misajon et al. (2005).

Three tests of the AQoL were carried out with the VisQoL data. The first two done were the regression of average VisQoL utility scores upon the scores predicted from the different AQoL models.

$$\text{VisQoL (TTO)} = a + b \text{ AQoL} \qquad \text{... (8)}$$

As previously, an unbiased model would result in coefficients of a = 0.00; and b = 1.00.

Results are reported in Table 10 below. Unreported R^2 values were also small. This result is to be expected. Most of the variation is between individuals with the same health state and there is only a single predicted AQoL score to 'explain' this variation; that is, most of the variation cannot be explained because of the construction of the test. The use of mean data overcomes this problem. The 'between' health state variation may, potentially, be explained by differences in the health state scores predicted by the AQoL models. This is the relevant test since CUA studies require mean, not individual, scores. Table 10 row 2 indicates very high explanatory power of all models with between 63 and 79 percent of variation explained.

Coefficients from the individual database suggest significant bias with models 1-7 implying that variation in observed VisQoL TTO scores was only about 40-46 percent of the variation in AQoL model scores. With this criterion models 7-9 give better prediction. In row 2 this pattern is largely repeated. Coefficients for model 1 to model 6 indicate that VisQoL TTO scores varied by only 52-59 percent as much as predicted AQoL model scores over the range of observations. However model 7 to model 9 provide estimates of b which are very close to unity and therefore include little bias.

5 Model Selection and Discussion

Results from the previous five tables are summarised in Table 11. Relatively good results are boxed; relatively poor results are shaded. It is clear from Table 11 that model 9 outperforms other models using the majority of the tests conducted and it is, therefore, the preferred model.

For this latter reason the CHE Monash website provides the algorithm for both model 1 and model 9. These are also reproduced in Appendix 2. The former is included as it entails least adjustment and is estimated only from MA (TTO) scores directly elicited.

6 Conclusion

The AQoL 2 set out to provide a more sensitive instrument than AQoL 1 and to identify and overcome methodological problems arising from this task. The resulting model strongly suggests that the structure of preferences for health states does not lend itself to simple modelling. This arises, fundamentally, from the complexity of the underlying construct. Results here suggest both structural and preference dependency. The methods developed in this paper should, in principle, minimise both of these sources of error. Results with the VisQoL data provide confidence in this conclusion. The results reported in this paper include the most rigorous tests for MAU models reported to date. It is difficult, however, to put the outcome of these tests in broad perspective as there are no comparable test results in the literature. In absolute terms however they suggest that AQoL 2 is a good instrument.

Table 7 Correlation Matrix

	AQoL	**TTO**	**M1**	**M2**	**M3**	**M4**	**M5**
AQoL [(1)]	1.00						
Mean **(2)**	0.89	1.00					
M1	0.98	0.92	1.00				
M2	0.99	0.90	0.99	1.00			
M3	0.99	0.91	0.99	0.99	1.00		
M4	0.94	0.91	0.96	0.95	0.96	1.00	
M5	0.99	0.90	0.98	0.99	0.98	0.94	1.00
M6	0.98	0.93	0.99	0.98	0.99	0.98	0.97
M7							
M8	0.93	0.88	0.95	0.94	0.95	0.96	0.93
M9	0.95	0.88	0.97	0.96	0.97	0.95	0.93
	M6	M7e	M8	M9			
M19re	1.00						
M60re	0.93	1.00					
M69re	0.96	0.95	1.00				
M71re	0.96	0.97	0.98	1.00			

Key

(1) AQoL multiplicative model
(2) Mean TTO data
(3) DU estimates using RE regression models

Table 8 Stage 4 OLS regressions:

Observed (MA) TTO on predicted TTO, Models 1-9

8.1	**MA(TTO) = $bM_i + e$**				**Individual E + D Data, n = 1403**				
	M1	M2	M3	M4	M5	M6	M7	M8	M9
Orig No.	1	7	9	11	17	19	60	69	71
b	0.84	0.86	0.86	0.79	0.87	0.86	1.07	1.32	1.07
R^2	(0.88)	(0.88)	(0.88)	(0.50)	(0.71)	(0.86)	(0.38)	(0.49)	(0.86)
8.2	**MA(TTO) = $a + bM_i + e$**				**Individual E + D Data, n = 1407**				
	M1	M2	M3	M4	M5	M6	M7	M8	M9
a	0.08	0.02	0.02	0.34	0.17	0.04	0.40	0.33	0.11
b	0.76	0.84	0.84	0.42	0.68	0.80	0.40	0.64	0.86
R^2	(0.75)	(0.74)	(0.74)	(0.18)	(0.41)	(0.68)	(0.07)	(0.12)	(0.72)
8.3	**MA(TTO) = $a + bM_i + e$**				**Mean E + D Data, n = 74**				
	M1	M2	M3	M4	M5	M6	M7	M8	M9
a	0.06	0.04	0.04	0.07	0.03	0.03	0.04	0.02	0.03
b	0.74	0.73	0.73	0.74	0.81	0.75	1.13	1.27	1.05
R^2	0.95	0.93	0.93	0.62	0.84	0.92	0.40	0.52	0.90
8.4	**MA(TTO) = $a + bM_i + e$**				**Mean E Data, n = 19**				
	M1	M2	M3	M4	M5	M6	M7	M8	M9
a	0.10	0.05	0.05	0.15	0.02	0.04	0.17	0.13	0.10
b	0.69	0.73	0.73	0.62	0.84	0.76	0.88	1.01	0.96
R^2	0.93	0.90	0.90	0.40	0.72	0.88	0.19	0.22	0.92

The R^2 test statistic for regressions omitting the constant cannot be interpreted in the normal way (as the percentage of variation 'explained'). Nevertheless they reflect a satisfactory fit except for models 7 and 8.

Table 9: Errors in estimates of incremental changes

[TTOs TTOO(e_1)-TTO(e_j)]-[Model e_i – model e_j] Absolute values, ie ignore sign

variable	Difference in mean TTO scores	Model 1 OBS-Pr ed	Model 3 OBS-Pred	Model 6 OBS-Pred	Model 9 OBS-Pred
column	1	2	3	4	5
chg e1 to e2	0.10	0.00	0.01	0.01	0.01
chg e1 to e3	0.08	0.04	0.04	005	0.05
chg e2 to e3	-0.02	0.03	0.05	0.03	0.03
chg e4 to e5	0.09	0.23	0.25	0.29	0.11
chg e4 to e6	-0.13	0.19	0.16	0.10	0.06
chg e5 to e6	-0.23	0.42	0.42	0.39	0.18
chg e7 to e8	-0.03	0.03	0.03	0.00	0.02
chg e7 to e9	0.04	0.16	0.20	0.15	0.04
chg e8 to e9	0.07	0.13	0.17	0.15	0.01
chg e10 to e11	0.29	0.32	0.36	0.33	0.01
chg e10 to e12	0.15	0.26	0.24	0.16	0.04
chg e11 to e12	-0.14	0.06	0.11	0.17	0.03
chg e13 to e14	-0.09	0.04	0.04	0.03	0.02
chg e13 to e15	0.11	0.00	0.00	0.02	0.14
chg e14 to e15	0.21	0.04	0.04	0.01	0.11
chg e16 to e17	0.11	0.09	0.07	0.16	0.08
chg e16 to e18	0.12	0.00	0.01	0.04	0.16
chg e17 to e18	0.01	0.09	0.09	0.12	0.08
Average*	0.11	0.12	0.13	0.12	0.07
Error 0.00-0.049		8	7	7	9
0.05-0.099		3	3	1	4
0.1-0.39		44	7	10	5
0.4*		3	1	0	0
Total					

*Average of absolute value ignoring the sign

Table 10: Results from the VisQoL validation study

Coefficient	Row	Stage 4 –type regressions: VisQoL (TTO) = a + b M_i + e								
(data)		M1	M2	M3	M4	M5	M6	M7	M8	M9
b (individual) n=752		0.43	0.45	0.43	0.41	0.44	0.46	0.76	0.64	0.65
b (mean) n=	1	0.59	0.54	0.56	0.55	0.52	0.58	1.08	0.86	0.87
R^2 (mean)	2	0.79	0.78	0.79	0.78	0.79	0.78	0.63	0.73	0.74

Notes: 'a' coefficients, R^2 (individual data) not reported

Relatively good results

Relatively poor results

Table 11. Summary of selection criteria

	M1	M2	M3	M4	M5	M6	M7	M8	M9
Stage 3	Various databases								
Wald	655	2156	4368	908	2764	5687	1431	6917	8677
Stage 4a	n = 1407								
B	0.84	0.86	0.86	0.79	0.87	0.86	1.07	1.32	1.02
R^2	0.88	0.88	0.88	0.50	0.71	0.86	0.38	0.49	0.86
Stage 4b	n = 74								
B	0.74	0.73	0.72	0.74	0.81	0.75	1.13	127	1.05
R^2	0.95	0.93	0.93	0.62	0.84	0.92	0.40	0.52	0.90
Error, predicted change, n = 19									
	0.12		0.13			0.13			

Key:

Relatively good results

Relatively poor results

VisQoL see Table 10

References

Brazier J, Usherwood T, HHasrper R, Thomas K (1998) Deriving a preference based single indix from the UK SF36*,* Health Survey*, Journal Clinical Edipimiology, 1998, 1115-1128.*

Brown, M.W., **Cudeck**, R (1993) “Alternate ways of assessing model fit” in Bollen K.A., Long, J.S. (Eds), Testing Structural Equation Models, Sage, 1993.

Feeny D, Torrance G et al 1996, ‘Health utilities index’, *Quality of Life and Pharmacoeconomics in Clinical Trials*, B Spilker, Lippincott-Raven Publishers, Philadelphia.

Hawthorne G, Richardson J et al 1997, The Assessment of Quality of Life (AQoL) Instrument: Construction, Initial Validation and Utility Scaling’, Working Paper 76 Centre for Health Program Evaluation, Monash University. *[this one not mentioned in the text]*

Hawthorne G, Richardson J and Day NA 2001, ‘A comparison of the assessment of quality of life (AQoL) with four other generic utility instruments’, *Annals of Medicine*, vol 33, no 5, pp 358-370.

Kaplan R, Ganiats T et al 1996, ‘The quality of wellbeing scale’, *Medical Outcome Trust Bulletin*, 4:2-3.

Kind P 1996, ‘The EuroQoL instrument: an index of health related quality of life’, *Quality of Life and Pharmacoeconomics in Clinical Trials*, B Spilker, Lippincott-Raven Publishers, Philadelphia.

Richardson J 1994, ‘Cost utility analysis: what should be measured’, *Social Science and Medicine*, 39:(1)7-21.

Richardson J 2002, ‘Evaluating summary measures of population health’, Chapter 3.7 pp 147-160 in (eds) CJL Murray, JA Salomon, CD Mathers, AD Lopez, *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications,* World Health Organization, Geneva.

Richardson, J., Hawthorne, G., Day, N.A, Peacock, S., Iezzi, A. 2004a, The Assessment of Quality of Life (AQoL) II Instrument: Conceptualising the Assessment of Quality of Life Instrument Mark 2 (AQoL 2) Methodological Innovations and the Development of the AQoL Descriptive System, Working Paper 141, Health Economics Unit, Monash University, Melbourne

Peacock, S., Richardson, J., Hawthorne, G., Day, N.A., Iezzi, A. 2004, The Assessment of Quality of Life (AQoL) 2 Instrument: The Effect of Deliberation and Alternative Utility Weights in a Multi Attribute Utility Instrument, Working Paper 143, Health Economics Unit, Monash University, MelbourneRichardson, J., Peacock, S., Hawthorne, G., Day, N.A., Iezzi, A. 2004d, The Assessment of Quality of Life (AQoL) 2 Instrument: Overview of the Assessment of Quality of Life Mark 2 Project, Working Paper 144, Health Economics Unit, Monash University, Melbourne.

Rosser R 1993, ‘A health index and output measure’, *Quality of Life Assessment: Key Issues in the 1990s*, S Walker and R Rosser (eds), Kluwer Academic Publishers, Dordrecht.

Sintonen H and Pekurinen M 1993, 'A fifteen dimensional measure of health related quality of life (15D) and its applications', *Quality of Life Assessment*, S Walker and R Rosser (eds), Kluwer Academic Publishers, Dordrecht.

Torrance G 1986 'Measurement of health state utilities for economic appraisal: a review', *Journal of Health Economics*, 5:1-30.

Ubel P, , G Loewenstein, et al (2001). "Do nonpatients underestimate the quality of life associated with chronic health conditions because of a focusing illusion?" *Medical Decision Making* ***21****: 190-199.*

von Winterfeldt DV and Edwards W 1986, *Decision Analysis and Behavioural Research*, Cambridge University Press, Cambridge.

Williams A 1995, 'The measurement and valuation of health: final report on the modelling of valuation tariffs', MVH Group, Centre for Health Economics, University of York, York.

YU, C-Y. (2002) Evaluating Cutoff Criteria of Model Fit Indices for Latent Variable Models with Binary and Continuous Outcomes. University of California, Dissertation for PhD.

Appendix 1 AQoL 2 Questionnaire

Assessment of Quality of Life (AQoL) Mark 2.

<u>Dimension 1: Independent Living</u>

Q1 How much help do I need with household tasks (eg preparing food, cleaning the house or gardening):

- I can do all these tasks very quickly and efficiently without any help
- I can do these tasks relatively easily without help
- I can do these tasks only very slowly without help
- I cannot do most of these tasks unless I have help
- I can do none of these tasks by myself.

Q2 Thinking about how easy or difficult it is for me to get around by myself outside my house (eg shopping, visiting):

- getting around is enjoyable and easy
- I have no difficulty getting around outside my house
- a little difficulty
- moderate difficulty
- a lot of difficulty
- I cannot get around unless somebody is there to help me.

Q3 Thinking about how well I can walk:

- I find walking or running very easy
- I have no real difficulty with walking or running
- I find walking or running slightly difficult. I cannot run to catch a tram or train, I find walking uphill difficult
- walking is difficult for me. I walk short distances only, I have difficulty walking up stairs
- I have great difficulty walking. I cannot walk without a walking stick or frame, or someone to help me
- I am bedridden.

Q4 Thinking about washing myself, toileting, dressing, eating or looking after my appearance:

- these tasks are very easy for me
- I have no real difficulty in carrying out these tasks
- I find some of these tasks difficult, but I manage to do them on my own
- many of these tasks are difficult, and I need help to do them
- I cannot do these tasks by myself at all.

Dimension 2: Social and Family

Q5 My close and intimate relationships (including any sexual relationships) make me:

- very happy
- generally happy
- neither happy nor unhappy
- generally unhappy
- very unhappy

Q6 Thinking about my health and my relationship with my family:

- my role in the family is unaffected by my health
- there are some parts of my family role I cannot carry out
- there are many parts of my family role I cannot carry out
- I cannot carry out any part of my family role.

Q7 Thinking about my health and my role in my community (that is to say neighbourhood, sporting, work, church or cultural groups):

- my role in the community is unaffected by my health
- there are some parts of my community role I cannot carry out
- there are many parts of my community role I cannot carry out
- I cannot carry out any part of my community role.

Dimension 3: Mental Health

Q8 How often did I feel in despair over the last seven days?

- never
- occasionally
- sometimes
- often
- all the time.

Q9 And still thinking about the last seven days: how often did I feel worried:

- never
- occasionally
- sometimes
- often
- all the time.

Q10 How often do I feel sad?

- never
- rarely
- some of the time
- usually
- nearly all the time.

Q11 When I think about whether I am calm and tranquil or agitated:

. always calm and tranquil
. usually calm and tranquil
. sometimes calm and tranquil, sometimes agitated
. usually agitated
. always agitated.

<u>Dimension 4: Coping</u>

Q12 Thinking about how much energy I have to do the things I want to do, I am:

. always full of energy
. usually full of energy
. occasionally energetic
. usually tired and lacking energy
. always tired and lacking energy.

Q13 How often do I feel in control of my life?

. always
. mostly
. sometimes
. only occasionally
. never.

Q14 How much do I feel I can cope with life's problems?

. completely
. mostly
. partly
. very little
. not at all.

<u>Dimension 5: Pain</u>

Q15 Thinking about how often I experience serious pain. I experience it:

. very rarely
. less than once a week
. three to four times a week
. most of the time.

Q16 How much pain or discomfort do I experience:

- none at all
- I have moderate pain
- I suffer from severe pain
- I suffer unbearable pain.

Q17 How often does pain interfere with my usual activities?

- never
- rarely
- sometimes
- often
- always
-
- .

Dimension 6: Senses

Q18 Thinking about my vision (using my glasses or contact lenses if needed):

- I have excellent sight
- I see normally
- I have some difficulty focusing on things, or I do not see them sharply. E.g. small print, a newspaper or seeing objects in the distance.
- I have a lot of difficulty seeing things. My vision is blurred. I can see just enough to get by with.
- I only see general shapes. I need a guide to move around
- I am completely blind.

Q19 Thinking about my hearing (using my hearing aid if needed):

- I have excellent hearing
- I hear normally
- I have some difficulty hearing or I do not hear clearly. I have trouble hearing softly-spoken people or when there is background noise.
- I have difficulty hearing things clearly. Often I do not understand what is said. I usually do not take part in conversations because I cannot hear what is said.
- I hear very little indeed. I cannot fully understand loud voices speaking directly to me.
- I am completely deaf.

Q20 When I communicate with others, e.g. by talking, listening, writing or signing:

- I have no trouble speaking to them or understanding what they are saying
- I have some difficulty being understood by people who do not know me. I have no trouble understanding what others are saying to me.
- I am understood only by people who know me well. I have great trouble understanding what others are saying to me.
- I cannot adequately communicate with others.

AQoL Study Background Questions

Please tick ☑ one box per question.

21 Are you:

❑ male ❑ female

22 In what year were you born? 19____

23 Where were you born?

❑ Australia ❑ Other country? ☞ Which one? ____________

24 Is English your first language?

❑ yes ❑ no ☞ Specify: ____________

25 What is your highest level of education?

- ❑ primary schooling only
- ❑ secondary schooling completed
- ❑ secondary schooling not completed. ☞ How many years completed? ___
- ❑ trade qualification or TAFE: ☞ Specify course: _____________
- ❑ University or other tertiary study
- ❑ Other or not applicable: please describe: ______________________

26 Which best describes your work situation: (Tick as many boxes as apply)

- ❑ full-time: self-employed or employee
- ❑ part-time: self-employed or employee
- ❑ unemployed, seeking work
- ❑ working in the home / home duties
- ❑ retired
- ❑ student
- ❑ other: please describe: ____________________

If You Are Employed Or Self-Employed Or Seeking Work:

27 What is your occupation? ________________________________

28 What do you do in your job? ________________________________

29 Do you receive any Government pension or benefit?

- ❑ no
- ❑ yes ☞ Which pension(s) or benefit(s): ____________________

30 Are you:

- ❑ married or living with a partner
- ❑ single: never married
- ❑ single: widowed
- ❑ single: divorced or separated

31 How would you rate your current level of health, for someone of your age?

- ❑ excellent
- ❑ very good
- ❑ good
- ❑ fair
- ❑ poor
- ❑ very poor
- ❑ extremely poor

32 Mark one box on the scale to show how important or unimportant is religion or spirituality is in your life

❑	❑	❑	❑	❑	❑	❑
↓		↓		↓		↓
very important		important		unimportant		very unimportan t

33 Please mark one box to show your HOUSEHOLD income, either annually, monthly or weekly. Include income that comes to the household from all sources. You may estimate either before or after tax.

	yearly	**monthly**	**fortnightly**	**weekly**
❑	under $20,000	under $1,665	under $800	under $385
❑	$20,001-$30,000	$1,665-$2,500	$800-$1,155	$385-$575
❑	$30,001-$40,000	$2,501-$3,330	$1,156-$1,535	$576-$770
❑	$40,001-$50,000	$3,331-$4,165	$1,536-$1,925	$771-$960
❑	$50,001-$60,000	$4,166-$5,000	$1,926-$2,305	$961-$1,155
❑	$60,001-$80,000	$5,001-$6665	$2,306-$3,075	$1,156-$1,540

- [ ] more than $80,000 more than $6665 more than $3,075 more than $1,540

34 Please mark a box to show whether your answer is before or after tax.

- [ ] before tax
- [ ] after tax

Thank you! Please bring this questionnaire with you when you attend the group session/interview.

Appendix 2 AQoL II Utility Algorithm

The following AQoL II Utility Algortihm is written in syntax for the Stata/SE 8 statistical software package. The data file which the algorithm is applied to should contain a single record for each resondepent. Column 1 should typically contain a unique identifier for each respondent. This should be followed by 20 columns for AQoL II data: one column for each AQoL II item. The variable for the first AQoL II item should be named aq1, the second named aq2, etc. Data entries for each AQoL II item should indicate the item level the respondent picked for that item. For example, if the respondent picked the top level for item 1, the entry for aq1 would be "1". If the respondent picked the 3rd best level for item 1, the entry for aq1 would be "3". The algorithm then calculates utilities based on item response for each of the AQoL II items.

DISUTILITIES FOR ITEM RESPONSES (A TYPES)

```
**Variable definitions:

**aq1-aq20 are item responses from administration of AQOLII
**duaq1-duaq20 are item disutilites
**dud1-dud6 are dimenstion disutilities
**duaqol=aqol disutility (0-1) scale
**duaqolld=aqol disutility (Life-Death) scale

**Note if data from aq1-aq20 is missing (i.e. a respondent did not answer an item)
  the observation is dropped from the utility model

generate duaq1=0 if aq1==1
replace duaq1=0.073441 if aq1==2
replace duaq1=0.435044 if aq1==3
replace duaq1=0.819933 if aq1==4
replace duaq1=1 if aq1==5

generate duaq2=0 if aq2==1
replace duaq2=0.032946 if aq2==2
replace duaq2=0.240038 if aq2==3
replace duaq2=0.470953 if aq2==4
replace duaq2=0.839769 if aq2==5
replace duaq2=1 if aq2==6

generate duaq3=0 if aq3==1
replace duaq3=0.041418 if aq3==2
replace duaq3=0.250737 if aq3==3
replace duaq3=0.569589 if aq3==4
replace duaq3=0.826952 if aq3==5
```

```
replace duaq3=1 if aq3==6

generate duaq4=0 if aq4==1
replace duaq4=0.040249 if aq4==2
replace duaq4=0.297115 if aq4==3
replace duaq4=0.797217 if aq4==4
replace duaq4=1 if aq4==5

generate duaq5=0 if aq5==1
replace duaq5=0.074061 if aq5==2
replace duaq5=0.46053 if aq5==3
replace duaq5=0.840618 if aq5==4
replace duaq5=1 if aq5==5

generate duaq6=0 if aq6==1
replace duaq6=0.193057 if aq6==2
replace duaq6=0.758943 if aq6==3
replace duaq6=1 if aq6==4

generate duaq7=0 if aq7==1
replace duaq7=0.196852 if aq7==2
replace duaq7=0.648117 if aq7==3
replace duaq7=1 if aq7==4

generate duaq8=0 if aq8==1
replace duaq8=0.133418 if aq8==2
replace duaq8=0.392291 if aq8==3
replace duaq8=0.837871 if aq8==4
replace duaq8=1 if aq8==5

generate duaq9=0 if aq9==1
replace duaq9=0.141557 if aq9==2
replace duaq9=0.391622 if aq9==3
replace duaq9=0.824482 if aq9==4
replace duaq9=1 if aq9==5

generate duaq10=0 if aq10==1
replace duaq10=0.097358 if aq10==2
replace duaq10=0.329611 if aq10==3
replace duaq10=0.783667 if aq10==4
replace duaq10=1 if aq10==5

generate duaq11=0 if aq11==1
replace duaq11=0.06389 if aq11==2
replace duaq11=0.368499 if aq11==3
replace duaq11=0.837281 if aq11==4
replace duaq11=1 if aq11==5
```

```
generate duaq12=0 if aq12==1
replace duaq12=0.056137 if aq12==2
replace duaq12=0.337631 if aq12==3
replace duaq12=0.72245 if aq12==4
replace duaq12=1 if aq12==5

generate duaq13=0 if aq13==1
replace duaq13=0.055008 if aq13==2
replace duaq13=0.381755 if aq13==3
replace duaq13=0.77363 if aq13==4
replace duaq13=1 if aq13==5

generate duaq14=0 if aq14==1
replace duaq14=0.056503 if aq14==2
replace duaq14=0.42309 if aq14==3
replace duaq14=0.825994 if aq14==4
replace duaq14=1 if aq14==5

generate duaq15=0 if aq15==1
replace duaq15=0.133048 if aq15==2
replace duaq15=0.642428 if aq15==3
replace duaq15=1 if aq15==4

generate duaq16=0 if aq16==1
replace duaq16=0.200438 if aq16==2
replace duaq16=0.757555 if aq16==3
replace duaq16=1 if aq16==4

generate duaq17=0 if aq17==1
replace duaq17=0.071958 if aq17==2
replace duaq17=0.338367 if aq17==3
replace duaq17=0.751957 if aq17==4
replace duaq17=1 if aq17==5

generate duaq18=0 if aq18==1
replace duaq18=0.032737 if aq18==2
replace duaq18=0.22308 if aq18==3
replace duaq18=0.621633 if aq18==4
replace duaq18=0.842872 if aq18==5
replace duaq18=1 if aq18==6

generate duaq19=0 if aq19==1
replace duaq19=0.024276 if aq19==2
replace duaq19=0.204844 if aq19==3
replace duaq19=0.585908 if aq19==4
replace duaq19=0.825651 if aq19==5
```

replace duaq19=1 if aq19==6

generate duaq20=0 if aq20==1
replace duaq20=0.186826 if aq20==2
replace duaq20=0.694913 if aq20==3
replace duaq20=1 if aq20==4

****Modelling Quality Of Life**

**example for 3 item dimension:

**DUd1 = 1/kd1[(1+kd1*w1*DUx1j)*(1+kd1*w2*DUx2j)*(1+kd1*w3*DUx3j)-1]

**where DUd1 = DU for dimension 1
** kd1 = scaling constant for dimension 1
** w1 = item 1 item worst disutility (ave value)
** (using 0-1 dimension best-dimension worst scale)
** DUx1j = item 1 disutility for jth item level

**Ud1=1-DUd1

MODELLING DIMENSIONS

DIMENSION 1 - IND LIV

**DIMENSION SCALING CONSTANT kd1=-0.978
**IND LIV HAS 4 ITEMS

***ITEM WORST WEIGHTS (Wi) FROM duttob1 variables
**duttob1b w1=0.385412
**duttob1c w2=0.593819
**duttob1d w3=0.630323
**duttob1e w4=0.794888

**4 item formula
**dud1=(1/kd1)*[(1+(kd1*w1*duaq1))*(1+(kd1*w2*duaq2))*(1+(kd1*w3*duaq3))*(1+(kd1*w4*duaq4))-1]

generate dud1=(1/-0.978)*[(1+(-0.978*0.385412*duaq1))*(1+(-0.978*0.593819*duaq2))*(1+(-0.978*0.630323*duaq3))*(1+(-0.978*0.794888*duaq4))-1]

DIMENSION 2 - SOCIAL & FAMILY

**DIMENSION SCALING CONSTANT kd2 = -0.923
**SOC & FAM HAS 3 ITEMS

**ITEM WORST WEIGHTS (Wi) FROM duttob2 variables
**duttob2b w5=0.64303
**duttob2c w6=0.697742
**duttob2d w7=0.508658

**3 item formula
**dud2=(1/kd2)*[(1+(kd2*w5*duaq5))*(1+(kd2*w6*duaq6))*(1+(kd2*w7*duaq7))-1]

generate dud2=(1/-0.923)*[(1+(-0.923*0.64303*duaq5))*(1+(-0.923*0.697742*duaq6))*(1+(-0.923*0.508658*duaq7))-1]

DIMENSION 3 - MENTAL HEALTH

***DIMENSION SCALING CONSTANT kd3 = -0.983
***MENTAL HEALTH HAS 4 ITEMS

***ITEM WORST WEIGHTS (Wi) FROM duttob3 variables
**duttob3b w8=0.640377
**duttob3c w9=0.588422
**duttob3d w10=0.648748
**duttob3e w11=0.71122

**4 item formula
**dud3=(1/kd3)*[(1+(kd3*w8*duaq8))*(1+(kd3*w9*duaq9))*(1+(kd3*w10*duaq10))*(1+(kd3*w11*duaq11))-1]

generate dud3=(1/-0.983)*[(1+(-0.983*0.640377*duaq8))*(1+(-0.983*0.588422*duaq9))*(1+(-0.983*0.648748*duaq10))*(1+(-0.983*0.71122*duaq11))-1]

DIMENSION 4 - COPING

***DIMENSION SCALING CONSTANT kd4 = -0.930
***COPING HAS 3 ITEMS

***ITEM WORST WEIGHTS (Wi) FROM duttob4 variables
**duttob4b w12=0.415694
**duttob4c w13=0.636994
**duttob4d w14=0.773296

**3 item formula
**dud4=(1/kd4)*[(1+(kd4*w12*duaq12))*(1+(kd4*w13*duaq13))*(1+(kd4*w14*duaq14))-1]

generate dud4=(1/-0.930)*[(1+(-0.930*0.415694*duaq12))*(1+(-0.930*0.636994*duaq13))*(1+(-0.930*0.773296*duaq14))-1]

DIMENSION 5 - PAIN

***DIMENSION SCALING CONSTANT kd5 = -0.962
***PAIN HAS 3 ITEMS

***ITEM WORST WEIGHTS (Wi) FROM duttob5 variables
**duttob5c w15=0.631833
**duttob5b w16=0.767573
**duttob5d w17=0.652241

**3 item formula
**dud5=(1/kd5)*[(1+(kd5*w15*duaq15))*(1+(kd5*w16*duaq16))*(1+(kd5*w17*duaq17))-1]

generate dud5=(1/-0.962)*[(1+(-0.962*0.631833*duaq15))*(1+(-0.962*0.767573*duaq16))*(1+(-0.962*0.652241*duaq17))-1]

DIMENSION 6 - SENSORY PERCEPTION

***DIMENSION SCALING CONSTANT kd6 = -0.851
***SENSPER HAS 3 ITEMS

***ITEM WORST WEIGHTS (Wi) FROM duttob6 variables
**duttob6b w18=0.580696
**duttob6c w19=0.463022
**duttob6d w20=0.604613

**3 item formula
**dud6=(1/kd6)*[(1+(kd6*w18*duaq18))*(1+(kd6*w19*duaq19))*(1+(kd6*w20*duaq20))-1]

generate dud6=(1/-0.851)*[(1+(-0.851*0.580696*duaq18))*(1+(-0.851*0.463022*duaq19))*(1+(-0.851*0.604613*duaq20))-1]

MODELLING THE AQOL ON A 0-1 DISUTILITY SCALE

***DIMENSION SCALING CONSTANT kA= -0.965

**6 item formula
**DUAQOL = (1/kA)*[(1+(kA*wd1*DUx1j))*(1+(kA*wd2*DUx2j))*(1+(kA*wd3*DUx3j))*(1+(kA*wd4*DUx4j))*(1+(kA*wd5*DUx5j))*(1+(kA*wd6*DUx6j))-1]

**where DUAQOL = DU for AQOL (0-1 Scale)

** kA = scaling constant for AQOL
** wd1 = dimension 1 dimension worst disutility (ave value)
** DU1j = dimension 1 disutility from Dimension Disutility modelling above

***DIMENSION WORST WEIGHTS (wdi) FROM dut2c1-dut2c6 variables (rescaled by dividing by AQOL D type divide by 1.132181)
**dut2c1 wd1=0.4724105
**dut2c2 wd2=0.4477805
**dut2c3 wd3=0.4788146
**dut2c4 wd4=0.3454342
**dut2c5 wd5=0.5920923
**dut2c6 wd6=0.637341

***scalar for wdi = 1/1.132181 = 0.883251

**6 item formula
**DUAQOL = (1/kA)*[(1+(kA*wd1*0.8*dud1))*(1+(kA*wd2*0.8*dud2))*(1+(kA*wd3*0.8*dud3))*(1+(kA*wd4*0.8*dud4))*(1+(kA*wd5*0.8*dud5))*(1+(kA*wd6*0.8*dud6))-1]

generate duaqol=(1/-0.965)*[(1+(-0.965*0.4724105*0.883251*dud1))*(1+(-0.965*0.4477805*0.883251*dud2))*(1+(-0.965*0.4788146*0.883251*dud3))*(1+(-0.965*0.3454342*0.883251*dud4))*(1+(-0.965*0.5920923*0.883251*dud5))*(1+(-0.965*0.637341*0.883251*dud6))-1]

MODELLING THE AQOL ON A LIFE-DEATH DISUTILITY SCALE

**AQOL ALL WORST DISUTILITY=SCALAR
**dut2daq=1.132181

generate duaqolld=1.132181*duaqol

STAGE II MULTIPLICATIVE MODEL - SECOND STAGE CORRECTION

**dummy variables for ranges of duaqolld

gen aqoldv1=0
gen aqoldv2=0
gen aqoldv3=0
gen aqoldv4=0
gen aqoldv5=0

replace aqoldv1=1 if duaqolld>=0 & duaqolld<0.25

```
replace aqoldv2=1 if duaqolld>=0.25 & duaqolld<0.5
replace aqoldv3=1 if duaqolld>=0.5 & duaqolld<0.75
replace aqoldv4=1 if duaqolld>=0.75 & duaqolld<1.0
replace aqoldv5=1 if duaqolld>=1.0
```

Model 1

```
generate m1re=duaqolld^(1.8407651)
```

Model 9

```
generate
   m9re=duaqolld^[(1.4544379)+(0.6357759*dud6*.70142711)+(0.470309*dud1*0.4468181
   *dud2*-4.6857753)+(0.4468181*dud2*0.6357759*dud6*-
   1.4205317)+(0.4779371*dud3*0.3459682*dud4*-
   2.2346052)+(aqoldv2*.42313558)+(aqoldv3*1.1013539)+(aqoldv4*2.6770203)+(aqoldv5*
   5.3075813)]
```