
Monograph:
Mavranezouli, I, Brazier, JE, Rowen, D et al. (1 more author) (2011) Estimating a preference-based index from the Clinical Outcomes in Routine Evaluation - Outcome Measure (CORE-OM): valuation of CORE-6D. Discussion Paper. (Unpublished)

HEDS Discussion Paper 11/05

HEDS Discussion Paper 11/05


ScHARR
School of Health and
Related Research

Title: Estimating a Preference-Based Index from the Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM): valuation of CORE-6D

Running title: Preference-based index for common mental disorders

Authors: Ifigeneia Mavranezouli, MD, MSc, John E Brazier, PhD, Donna Rowen, PhD, Michael Barkham, PhD

IM, JEB, DR: Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, Sheffield, United Kingdom

IM: National Collaborating Centre for Mental Health, Centre for Outcomes Research and Effectiveness, Research Department of Clinical, Educational & Health Psychology, University College London, London, United Kingdom

MB: Centre for Psychological Services Research, University of Sheffield, Sheffield, United Kingdom

Corresponding author:

Ifigeneia Mavranezouli

National Collaborating Centre for Mental Health, Centre for Outcomes Research and Effectiveness, Research Department of Clinical, Educational & Health Psychology, University College London

Gower Street, London WC1E 6BT, United Kingdom

Tel: +44 (0)207 679 1964

Fax: +44 (0)207 91 68 511

e-mail: i.mavranezouli@ucl.ac.uk

Source of financial support: The study was funded by the MRC-NIHR Methodology Research Programme, project number 06/97/04. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

Keywords: condition-specific, CORE-6D, CORE-OM, health state valuation, mental health, preference-based index, time-trade-off.

Wordcount: 4,370

Abstract:

Background: The Clinical Outcomes in Routine Evaluation - Outcome Measure (CORE-OM) is used to evaluate the effectiveness of psychological therapies in people with common mental disorders. The objective of this study was to estimate a preference-based index for this population using CORE-6D, a health state classification system derived from CORE-OM consisting of a 5-item emotional component and a physical item, and to demonstrate a novel method for generating states that are not orthogonal.

Methods: Rasch analysis was used to identify 11 plausible 'emotional' health states from CORE-6D (rather than conventional statistical design that would generate implausible states). By combining these with the 3 response levels of the physical item of CORE-6D, 33 plausible health states can be described, of which 18 were selected for valuation. An interview valuation survey of 220 members of public in South Yorkshire, UK, was undertaken using the time-trade-off method to value the 18 health states; regression analysis was subsequently used to predict values for all possible states described by CORE-6D.

Results: A number of multivariate regression models were built to predict values for the 33 plausible health states of CORE-6D, using the Rasch logit value of the emotional health state and the response level of the physical item as independent variables. A cubic model with high predictive value (adjusted R^2 0.990) was finally selected, which can be used to predict utility values for all 927 states described by CORE-6D.

Conclusion: The CORE-6D preference-based index will enable the assessment of cost-effectiveness of interventions for people with common mental disorders using existing and prospective CORE-OM datasets. The new method for generating states may be useful for other instruments with highly correlated dimensions.

INTRODUCTION

Quality Adjusted Life Years (QALYs) are increasingly used as the measure of benefit in economic evaluations of health care technologies and programmes worldwide. Several preference-based measures (PBMs) have been developed aiming at the estimation of utility values that can be used for calculation of QALYs. Among the most widely used are the EuroQol-5D (EQ-5D),(1) the SF-6D,(2) and the HUI-3.(3) All three measures are generic and can therefore be used for the assessment of interventions and programmes targeted at different disease areas and patient populations.

However, generic PBMs may be less appropriate or sensitive in some medical conditions.(4;5) Especially in the area of mental health, there are concerns that generic PBMs may lack sensitivity in capturing important elements of health-related quality of life (HRQoL), due to their focus on physical aspects of health (for example, 4 out of 5 items of EQ-5D capture physical aspects of HRQoL). This has led to proposals for the development of a PBM specific to mental health, that will be suitable for use across a wide range of mental health conditions.(6-8) Currently, no such measure is available. A report examining the feasibility of incorporating patient-rated measures in mental health into a productivity measure for use in the UK identified the Clinical Outcomes in Routine Evaluation - Outcome Measure (CORE-OM) as a good candidate for this purpose.(9)

CORE-OM is a patient-based instrument that is widely used in the UK to evaluate the effectiveness of psychological therapies in people with common mental disorders.(10;11) It consists of 34 items, each with 5 levels of response (ranging from ‘not at all’ to ‘most or all the time’), tapping 4 conceptual domains: ‘subjective well-being’, ‘problems’, ‘functioning’ and ‘risk’. The validity, reliability and acceptability of CORE-OM has been demonstrated across a wide range of practice settings.(12;13) Based on these characteristics and given

the arguments favouring the development of a PBM specific to mental health, CORE-OM was selected as the basis for constructing a PBM for use in common mental disorders.

Derivation of a PBM from CORE-OM requires a three-step process: first, the development of a health state descriptive system; second, a valuation survey, in which respondents attach utility values in selected health states derived from the descriptive system; and third, modeling of the utility values leading to an algorithm that links all possible health states to utility values. Previous work has reported on the first stage of the construction of CORE-6D, a health state descriptive system derived from CORE-OM.(14) The primary objective of this paper is to report on the later stages covering the development of an algorithm linking all health states described by CORE-6D with appropriate utility values, using the results of a valuation survey on CORE-6D health states and further modeling. A secondary objective is to examine an alternative method for generating health states when dimensions are highly correlated using the results of Rasch analysis.

METHODS

The CORE-6D health descriptive system

CORE-6D is a 6-item health descriptive system consisting of a 5-item unidimensional 'emotional' component and a physical item.(14) Each item has 3 response levels: 'never', 'only occasionally or sometimes' and 'often, most or all the time'. The system describes 3^6 = 729 unique health states. The emotional component of CORE-6D was derived from CORE-OM using predominantly Rasch analysis(15) (supported by classical psychometric testing) to analyse a dataset containing information on 400 people with common mental disorders attending primary care services in the UK. The emotional component of CORE-6D comprises a unidimensional scale, which, combined with a physical item, creates a 2-dimensional scale, tapping emotional and physical symptoms in people with common mental disorders. The CORE-6D health state descriptive system is shown on Table 1.

Rasch analysis

Rasch analysis is a statistical approach for examining people's abilities, such as knowledge, skills and perceptions ('latent traits'); it is based on the principles of the Rasch model,(15) according to which the outcome of an encounter between a person and an item is exclusively governed by the 'ability' of the person and the 'difficulty' of the item. The Rasch model indicates an 'ideal' relationship between an observed response and the 'amount' of the latent trait measured by an item(16) and demonstrates what the expected responses to items should be, if interval scale measurement is to be achieved.(17) Rasch analysis orders persons according to their 'ability' or 'severity' (according to their 'amount' of the latent trait), and ranks questionnaire items according to their difficulty.(18) Subsequently, Rasch analysis assigns persons to different difficulty points ('locations') along the latent variable (Rasch model logit scale) generating groups of respondents of different ability/severity.(18) The Rasch model is underpinned by the principle of unidimensionality, which requires all items fitting into the Rasch model to express the same underlying latent trait.(17) Rasch analysis has been successfully used as a tool in the development and refinement of patient reported outcome measures, and more recently in the development of various condition-specific PBMs.(19-22)

Selection of health states for the valuation survey

The emotional component of CORE-6D can describe $3^5 = 243$ health states. However, this component has been shown to be unidimensional,(14) and so its items are not independent from each other, resulting in some item response combinations being implausible; e.g. "I make plans to end my life often, most or all the time" *and* "I never feel terribly alone and isolated". Use of conventional statistical approaches for generating health states (such as orthogonal arrays) is not appropriate in this case, because it is likely to generate infeasible health states due to the high correlation between items. We have applied a novel method for

generating health states, the ‘Rasch vignette approach’, in order to identify plausible health states amenable to valuation.(14) This approach relies on the inspection of the item threshold map for the unidimensional emotional component, an output of Rasch analysis, which depicts the most likely item response combinations expected for each location across the latent trait. Such combinations represent frequently observed, plausible ‘emotional’ health states experienced by the study population across the continuum of symptoms of the condition examined. To obtain the full CORE-6D state, emotional health states selected using the Rasch vignette approach, made up of 5 items, need to be subsequently combined with different response levels of the physical item for use in the valuation survey.

Inspection of the Rasch item threshold map of the emotional component of CORE-6D in Figure 1 helped identify the most likely item response combinations across the continuum of the emotional symptom severity. Items have been ordered from the easiest to the most difficult one, as indicated by their average location in the Rasch model. Shaded areas 0 (black), 1 (dark grey) and 2 (light grey) correspond to the 3 response levels, that is, ‘never’, ‘only occasionally or sometimes’, and ‘often, most or all the time’ respectively, with the exception of the positively worded item, the response levels of which are reversed. The map allows prediction of the most likely responses at various levels of symptom severity. For example, a person whose symptom severity corresponds to location +1 on the Rasch model logit scale is expected to most likely respond 22210. These item combinations represent frequently observed, plausible health states experienced by people with common mental disorders.

As illustrated in Table 2, 11 plausible emotional health states (response combinations) were identified; these cover 37.1% of response combinations obtained by the study sample (after excluding cases with one or more responses missing). These 11 emotional health states, combined with 3 response levels of the physical item of CORE-6D, produce a 2-dimensional

set of 11 x 3 = 33 plausible health states. Emotional health state 10 (22221) was not represented in the study sample (0 out of 400 cases as shown in Table 2) and was therefore excluded from further consideration. The remaining 10 emotional health states combined with the physical item at response level zero (never troubled by physical problems) were selected for valuation. To assess the impact of physical functioning on utility values, 4 of the emotional health states (including best state 00000, worst state 22222 and two intermediate states) were combined with different response levels (levels 1 and 2) of the physical item, so as to cover the full severity range captured by CORE-6D. Intermediate emotional states 3 (11000) and 7 (22110) were chosen for this purpose, based on their relative frequency in the study sample (shown in Table 2) and their location coverage (range) on the item threshold map (shown in Figure 1). In total, 18 plausible CORE-6D health states were selected for the valuation survey, plus 4 emotional health states with no reference to the physical item. Responses to the states describing only the emotional component will be analysed in a separate piece of work.

Three card blocks each containing 8 cards were used at valuation. Each card described one health state, consisting of one of the emotional health states from Table 2 combined with one of the response levels of the physical item. One of the card blocks consisted of 4 cards describing emotional health states only, without reference to the physical item, and of 4 cards describing the same emotional health states plus the physical item at response level zero. State 222220 was included in all 3 card blocks.

Valuation survey

A valuation survey using face-to-face interviews was carried out in South Yorkshire, UK, aiming at determining public preferences for a number of health states derived from CORE-6D. Selected health states were valued using TTO. Interviews were conducted by trained and experienced interviewers from the Centre for Health and Social Care Research at

Sheffield Hallam University. Respondents were selected using sampling from streets in both urban and rural areas with a mix of socio-economic characteristics in the North of England using the AFD Names and Numbers version 3.1.25 database (AFD Software Limited, Ramsey, UK). Households in these areas received letters informing them that interviewers would be in their area and interviewers then visited houses. All willing participants were then interviewed in the respondent's own home. Addresses were visited up to four times on different days and times of day before an address was considered a non-responder. No financial reward was offered for participation in the survey. Ethical approval for the valuation survey was received by the ScHARR Research Ethics Committee at the University of Sheffield.

Respondents were asked firstly to self-complete EQ-5D and CORE-6D for their own health, so as to become familiarised with the idea of describing states, as well as with the items and response levels of CORE-6D. Subsequently, each respondent undertook warm-up ranking and TTO tasks and TTO valuations of eight health states. All respondents first ranked and valued 4 states and subsequently ranked and valued the remaining 4 states in the card block. In the card block that contained states without reference to the physical item, emotional states were ranked and valued first, followed by ranking and valuation of states that included the physical item at response level zero; in the other two card blocks, the 4 states valued first were chosen at random. The Measurement and Valuation of Health (MVH) group version of TTO was used including the visual prop designed by the MVH group (University of York) to allow comparisons of the survey findings with the EQ-5D tariff.(23) Because of the nature of some item responses (e.g. I make plans to end my life), respondents were informed in the cover letter and information sheet that the interview was about common mental and physical health problems. In the information sheet and in a thank you note left at the end of the interview all respondents were strongly recommended that they seek appropriate professional support either from their GP or from a professional

agency such as the Samaritans (contact details provided) if the interview raised personal issues for them. Respondents were also asked a number of background questions covering health, demographic and socio-economic characteristics and how difficult they found the valuation tasks.

Modelling utility values for all CORE-6D health states using Rasch analysis

The standard approach for modelling utility values has been by creating dummy variables for each level of every dimension of an instrument(2;23) and regressing these onto the health state values (obtained using TTO or standard gamble). However, this approach was not appropriate here, since the highly correlated items of the emotional component of CORE-6D were expected to produce significant, multiple interaction effects, and consideration of all possible interactions across different response levels of different items would require complex regression models as well as valuation of a large number of health states in order to predict utility values for all health states of the instrument. This can be avoided using an alternative method described by Young et al. that uses the relationship between the Rasch model logit value and the respective TTO utility value of a health state of a unidimensional measure to predict TTO utility values for all potential states of the measure.(24)

Nevertheless, this new method alone was not adequate for the estimation of utility values for CORE-6D; this is because CORE-6D is a 2-dimensional scale, consisting of a unidimensional emotional component and a physical item. In order to predict utility values for all health states described by CORE-6D taking into account the effect of the physical item, we adopted a hybrid approach: we used as a basis the methodology described by Young et al.,(24) appropriate for the prediction of utility values in the case of unidimensional measures such as the emotional component of CORE-6D, and also created dummy variables to represent the different severity levels of the physical item, which is a standard approach used for multidimensional measures.(2;23) Consequently, a series of regression analyses

were undertaken at the mean (health state) level to explore the relationship between the TTO value for each health state considered in valuation and

a. the respective Rasch model logit value corresponding to the emotional component of the health state, as indentified in previously undertaken Rasch analysis
b. the response level (0, 1 or 2) of the physical item of the health state, modeled in the form of 2 dummy dichotomous variables, one for response level 1 and one for response level 2.

A number of regression models were fitted, including simple linear, quadratic and cubic relationship. Model fit was compared using the coefficient of determination (i.e. the adjusted R-Squared) and the root mean squared error (RMSE) at the state level. The model with the best fit was selected in order to predict mean TTO values for all health states described by CORE-6D based on their respective Rasch model logit value and the response level of the physical item.

RESULTS

Valuation survey – respondents' characteristics

The valuation survey was conducted on 225 respondents, a response rate of 45.7% for respondents answering their door at the time of interview. The study achieved a completion rate of 99.7% for all 18 health states included in the TTO valuations considered in this study (4 missing TTO values). Characteristics of all respondents included in the analysis are presented in Table 3, which allows comparison of the study sample to the general population in South Yorkshire and England. The study sample had a higher average age, a higher proportion of females, home owners and retired individuals, and a lower proportion of employed/self-employed individuals. A large proportion of respondents reported that they found the rank and TTO tasks difficult (27.6% and 31.1% respectively). However,

interviewers reported it was doubtful whether the respondent understood the rank and TTO tasks in just 5.8% and 4.9% of the interviews, respectively.
(25)

Utility values obtained from the valuation survey

The TTO values obtained from the valuation survey are reported in Table 4 and Table 5. Table 4 provides descriptive statistics for the health state values obtained for each health state. It can be seen that the mean TTO values range from 0.96 (best state 000000) to 0.10 (worst state 222222). Table 5, which shows responses by card block, demonstrates the changes in obtained TTO values with increasing severity of physical and emotional symptoms: moving to states with more severe physical symptoms (i.e. increasing the response level of the physical item), while keeping the emotional health state unchanged, results in a decrease in the average TTO value; similarly, moving to states with more severe emotional symptoms (i.e. moving from emotional state 00000 to emotional state 22222), while keeping the response level of the physical item intact, also results in a decrease in the average TTO value. There is only one inconsistency to this pattern, observed in states 100000 and 110000; in this case the mean TTO value increased by a small and non-significant amount (from 0.87 to 0.88, respectively) despite of the increase in emotional symptom severity. This inconsistency can be explained by the fact that these health states were included in different card blocs and hence were valued by different respondents.

Modelling utility values of CORE-6D health states using respective Rasch model logit values and the response level of the physical item

Rasch model logit values for each emotional health state were rescaled and anchored at 0.96 (best emotional state 00000) and 0.23 (worse emotional state 22222), which were the observed mean TTO values obtained from the valuation survey. In order to predict TTO values for all health states described by CORE-6D, a number of mean (health state) level

regression models across all 18 states were explored using as independent variables the Rasch model rescaled logit value (assuming simple linear, quadratic and cubic relationships) and 2 dummy variables accounting for the response level of the physical item.

The following model specifications were tested:

Model 1 – simple linear relationship: $y = \alpha + \beta_1 R + \gamma_1 P_1 + \gamma_2 P_2$

Model 2 – quadratic relationship: $y = \alpha + \beta_2 R^2 + \gamma_1 P_1 + \gamma_2 P_2$

Model 3 – cubic relationship: $y = \alpha + \beta_3 R^3 + \gamma_1 P_1 + \gamma_2 P_2$

Model 4 – quadratic relationship: $y = \alpha + \beta_1 R + \beta_2 R^2 + \gamma_1 P_1 + \gamma_2 P_2$

Model 5 – cubic relationship: $y = \alpha + \beta_1 R + \beta_3 R^3 + \gamma_1 P_1 + \gamma_2 P_2$

Model 6 – cubic relationship: $y = \alpha + \beta_2 R^2 + \beta_3 R^3 + \gamma_1 P_1 + \gamma_2 P_2$

Model 7 – cubic relationship: $y = \alpha + \beta_1 R + \beta_2 R^2 + \beta_3 R^3 + \gamma_1 P_1 + \gamma_2 P_2$

where y is the mean predicted TTO value, R is the Rasch model rescaled logit value, P_1 is a dummy variable for response level 1 of the physical item (I have been troubled by aches, pains, physical problems only occasionally or sometimes), P_2 is a dummy variable for response level 2 of the physical item (I have been troubled by aches, pains, physical problems often, most or all the time), α is the constant, and β_i and γ_i are regression coefficients.

The regression coefficients and goodness of fit statistics for all 7 models are shown in Table 6. The adjusted R-Squared statistics varied from 0.773 (model 3) to 0.990 (model 7). Dummy variable P_1 was non-significant in any of the models. In model 7 the level of significance was only slightly above 0.05 (0.069). Based on having the lowest RMSE statistics of 0.0275, the largest model that contained linear, quadratic and cubic terms for the logit value and both physical dummies (model 7) was selected for the prediction of TTO values for all health states described by CORE-6D. It also had the best fit in terms of adjusted R-Squared.

Figure 2 allows the comparison between actual mean TTO values obtained from the valuation survey for the selected CORE-6D health states, and predicted TTO values for all potential health states described by CORE-6D, derived from the regression model 7. The x axis of the graph represents Rasch rescaled logit values that cover the full severity range of all potential emotional health states described by CORE-6D. There are three lines on the graph, one for each level of the physical item. The 3 lines have an s-shape reflecting the cubic relationship between the Rasch logit scale and the TTO health state value.

An SPSS syntax file that allows calculation of CORE-6D utility values from CORE-OM data is available from the corresponding author on request.

DISCUSSION

This paper describes the development of a PBM from the health state descriptive system CORE-6D, which, in turn, has been derived from CORE-OM, an outcome measure for common mental disorders widely used in clinical practice in the UK. The development of the CORE-6D PBM involved a 3-stage process, using predominantly Rasch analysis: this was first used to derive the unidimensional emotional component of CORE-6D from CORE-OM, and to identify 11 plausible emotional health states amenable to valuation.(14) These 11 emotional health states, combined with the 3 response levels of the physical item of CORE-6D, produced 33 plausible health states, 18 of which were selected for valuation. Following the valuation survey, a number of regression models were tested to explore the best option in order to predict utility values for all health states described by CORE-6D, based on the respective Rasch model rescaled logit value of the emotional state and the response level of the physical item.

The novel methodology developed in this study for the generation of health states for valuation and the subsequent prediction of utility values for all CORE-6D states was dictated by the 2-dimensional structure of CORE-6D. Generation of health states from the emotional component of CORE-6D was achieved by inspection of the Rasch item threshold map, which indicated the most frequent, and more importantly, plausible emotional health states observed in the study population.(14) Standard statistical approaches for generating health states such as orthogonal arrays would not be appropriate in this case, as these would likely result in the selection of implausible health states, due to the unidimensionality of the emotional component and the high correlation across its items. Plausible emotional health states were then combined with the 3 response levels of the physical item of CORE-6D in order to develop 2-dimensional health states amenable to valuation.

Following the valuation survey, the study then built on the approach described by Young et al.(24) that uses regression analysis to estimate the relationship between the Rasch model logit values and observed TTO values, in order to generate utility values for all states described by a unidimensional measure. This approach was considered appropriate given the high correlation between the 5 items of the unidimensional emotional component of CORE-6D. Use of the standard approach for modeling utility values by creating dummy variables for each level of every item of the measure(2;23) would have required far more states to be valued. In contrast, using the Rasch logit values offers a more efficient solution. Our study successfully adapted the approach reported by Young et al. by incorporating dummy variables in regression analysis to account for the different severity levels of the physical item of CORE-6D (according to the standard approach used to model utility values for multidimensional measures(2;23)). Our mixed approach enabled us to predict utility values for all potential health states described by CORE-6D. This approach can therefore be used in order to estimate utility values for multidimensional measures that encompass one or more unidimensional components.

The results of regression analysis indicated the solution of a cubic model with RMSE = 0.0275 and adjusted R-Squared = 0.990, which can be used to predict utility values for all 729 health states described by CORE-6D. These results compare very favourably with other studies, where the RMSE was typically above 0.05 and the adjusted R-Squared below 0.6.(2;23;26-28) These results suggest that using the Rasch logit value rather than individual dummy variables is a more efficient way for predicting utility values for states not included in the valuation survey.

One limitation of the new measure is that it is only suitable for common mental disorders, such as depression and anxiety. CORE-OM has not been designed for use in other mental disorders such as schizophrenia, bipolar disorder, personality disorders etc. Consequently, CORE-6D cannot be used for the estimation of QALYs at the evaluation of interventions targeted at mental disorders other than depression and anxiety, and therefore cannot be used as a 'generic' mental health PBM. Nonetheless, common mental disorders constitute the most prevalent group of disorders in the UK, experienced by 16.2% of people aged 16-64 years in England (for comparison, psychotic disorders are experienced by 0.4% of the same population).(29)

Another limitation of CORE-6D is that it largely focuses on emotional symptoms, as it includes 5 emotional items and only one physical item. The composition of CORE-6D reflects the structure of CORE-OM (from which CORE-6D was derived), which is a measure primarily designed for the monitoring of emotional, rather than physical, symptoms. Inclusion of one physical item in CORE-6D allows a rather crude representation of physical symptoms, which, nevertheless, enables the assessment and valuation of both emotional and physical dimensions of HRQoL in people with common mental disorders.

Compared with generic PBMs, condition-specific PBMs, such as CORE-6D, are expected to be more relevant and sensitive to the condition they have been designed for; on the other hand, they are characterised by a number of limitations, such as their inability to capture side-effects of treatment and comorbidities, and the distortions created by focusing effects.(5) Use of condition-specific PBMs raises concerns regarding their comparability to generic measures in the wider resource allocation context, although it has been argued that comparability across different PBMs can be improved if utility values are obtained using the same valuation technique, on a scale with common anchors (full health and death), and elicited from the same population.(30) The role of generic and condition-specific PBMs has been (and still is) an important subject of debate.(31-34) Nevertheless, the new PBM described in this paper can be used for the estimation of QALYs in cost-utility analyses of healthcare technologies for people with common mental disorders, where the use of generic PBMs has been reported to be problematic.(6;35)

The appropriateness and sensitivity of CORE-6D is going to be assessed as a next step of this study, with the new measure being compared with generic PBMs such as EQ-5D and SF-6D in populations of people with common mental disorders. Given that CORE-OM is an instrument routinely used for the clinical monitoring of people with common mental disorders in the UK, the preference-based CORE-6D is expected to contribute to the wider assessment of the cost-effectiveness of interventions for common mental disorders using existing and prospective CORE-OM datasets.

Acknowledgments: We are grateful to the CORE System Trustees for giving approval for the development of the CORE-6D.

REFERENCES

(1) Brooks R. EuroQol: the current state of play. Health Policy 1996 Jul;37(1):53-72.

(2) Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ 2002 Mar;21(2):271-92.

(3) Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, et al. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Med Care 2002 Feb;40(2):113-28.

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

(5) Brazier J, Tsuchiya A. Preference-based condition-specific measures of health: what happens to cross programme comparability? Health Econ 2010 Feb;19(2):125-9.

(6) Brazier J. Measuring and valuing mental health for use in economic evaluation. J Health Serv Res Policy 2008 Oct;13 Suppl 3:70-5.

(7) Chisholm D, Healey A, Knapp M. QALYs and mental health care. Soc Psychiatry Psychiatr Epidemiol 1997 Feb;32(2):68-75.

(8) Knapp M, Mangalore R. "The trouble with QALYs...". Epidemiol Psichiatr Soc 2007 Oct;16(4):289-93.

(9) Jacobs R. Investigating Patient Outcome Measures in Mental Health. CHE Research Paper Number 48. York: Centre for Health Economics, University of York; 2009.

(10) Barkham M, Margison F, Leach C, Lucock M, Mellor-Clark J, Evans C, et al. Service profiling and outcomes benchmarking using the CORE-OM: toward practice-based

evidence in the psychological therapies. Clinical Outcomes in Routine Evaluation-Outcome Measures. J Consult Clin Psychol 2001 Apr;69(2):184-96.

(11) Barkham M, Stiles WB, Connell J, Twigg E, Leach C, Lucock M, et al. Effects of psychological therapies in randomized trials and practice-based studies. Br J Clin Psychol 2008 Nov;47(Pt 4):397-415.

(12) Barkham M, Gilbert N, Connell J, Marshall C, Twigg E. Suitability and utility of the CORE-OM and CORE-A for assessing severity of presenting problems in psychological therapy services based in primary and secondary care settings. Br J Psychiatry 2005 Mar;186:239-46.

(13) Evans C, Connell J, Barkham M, Margison F, McGrath G, Mellor-Clark J, et al. Towards a standardised brief outcome measure: psychometric properties and utility of the CORE-OM. Br J Psychiatry 2002 Jan;180:51-60.

(14) Mavranezouli I, Brazier JE, Young TA, Barkham M. Using Rasch analysis to form plausible health states amenable to valuation: the development of CORE-6D from a measure of common mental health problems (CORE-OM). Qual Life Res 2011 Apr;20(3):321-33.

(15) Rasch G. Probabilistic models for some intelligence and attainment tests. Chicago: University of Chicago Press; 1960.

(16) Tesio L, Simone A, Bernardinello M. Rehabilitation and outcome measurement: where is Rasch analysis-going? Eura Medicophys 2007 Sep;43(3):417-26.

(17) Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Rheum 2007 Dec 15;57(8):1358-62.

(18) Bond TG, Fox CM. Applying the Rasch Model: Fundamental Measurement in the Human Sciences. Mahwah, New Jersey: Lawrence Erlbaum Associates Inc; 2006.

(19) McTaggart-Cowan HM, Brazier JE, Tsuchiya A. Clustering Rasch results: a novel method for developing rheumatoid arthritis states for use in valuation studies. Value Health 2010 Sep;13(6):787-95.

(20) Sundaram M, Smith MJ, Revicki DA, Elswick B, Miller LA. Rasch analysis informed the development of a classification system for a diabetes-specific preference-based measure of health. J Clin Epidemiol 2009 Aug;62(8):845-56.

(21) Young T, Yang Y, Brazier JE, Tsuchiya A, Coyne K. The first stage of developing preference-based measures: constructing a health-state classification using Rasch analysis. Qual Life Res 2009 Mar;18(2):253-65.

(22) Young TA, Yang Y, Brazier JE, Tsuchiya A. The Use of Rasch Analysis in Reducing a Large Condition-Specific Instrument for Preference Valuation: The Case of Moving from AQLQ to AQL-5D. Med Decis Making 2011 Jan;31(1):195-210.

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

(24) Young TA, Rowen D, Norquist J, Brazier JE. Developing preference-based health measures: using Rasch analysis to generate health state values. Qual Life Res 2010 Aug;19(6):907-17.

(25) Kind P, hardman G, Macran S. UK population norms for EQ-5D. York: Centre for Health Economics Discussion Paper Series, University of York; 1999.

(26) Yang Y, Brazier JE, Tsuchiya A, Young TA. Estimating a preference-based index for a 5-dimensional health state classification for asthma derived from the asthma quality of life questionnaire. Med Decis Making 2011 Mar;31(2):281-91.

(27) Yang Y, Brazier J, Tsuchiya A, Coyne K. Estimating a preference-based single index from the Overactive Bladder Questionnaire. Value Health 2009 Jan;12(1):159-66.

(28) Brazier J, Czoski-Murray C, Roberts J, Brown M, Symonds T, Kelleher C. Estimation of a preference-based index from a condition-specific measure: the King's Health Questionnaire. Med Decis Making 2008 Jan;28(1):113-26.

(29) McManus S, Meltzer H, Brugha T, Bebbington P, Jenkins R. Adult psychiatric morbidity in England, 2007: Results of a household survey. Leeds: The NHS Information Centre for Health and Social Care; 2009.

(30) Brazier JE, Ratcliffe J, Salomon J, Tsuchiya A. Measuring and valuing health benefits for economic evaluation. Oxford/New York: Oxford University Press; 2007.

(31) Dowie J. Decision validity should determine whether a generic or condition-specific HRQOL measure is used in health care decisions. Health Econ 2002 Jan;11(1):1-8.

(32) Brazier J, Fitzpatrick R. Measures of health-related quality of life in an imperfect world: a comment on Dowie. Health Econ 2002 Jan;11(1):17-9.

(33) Feeny D. Commentary on Jack Dowie, "Decision validity should determine whether a generic or condition-specific HRQOL measure is used in health care decisions". Health Econ 2002 Jan;11(1):13-6.

(34) Guyatt G. Commentary on Jack Dowie, "Decision validity should determine whether a generic or condition-specific HRQOL measure is used in health care decisions". Health Econ 2002 Jan;11(1):9-12.

(35) Moock J, Kohlmann T. Comparing preference-based quality-of-life measures: results from rehabilitation patients with musculoskeletal, cardiovascular, or psychosomatic disorders. Qual Life Res 2008 Apr;17(3):485-95.

Table 1. The CORE-6D descriptive system

Emotional component

1	I never feel terribly alone and isolated	0
	I feel terribly alone and isolated only occasionally or sometimes	1
	I feel terribly alone and isolated often, most or all the time	2
2	I never feel panic or terror	0
	I feel panic or terror only occasionally or sometimes	1
	I feel panic or terror often, most or all the time	2
3	I never feel humiliated or shamed by other people	0
	I feel humiliated or shamed by other people only occasionally or sometimes	1
	I feel humiliated or shamed by other people often, most or all the time	2
4	I am able to do most things I need to often, most or all the time	0
	I am able to do most things I need to only occasionally or sometimes	1
	I am not able to do the things I need to	2
5	I never make plans to end my life	0
	I make plans to end my life only occasionally or sometimes	1
	I make plans to end my life often, most or all the time	2

Physical health item

6	I am never troubled by aches, pains or other physical problems	0
	I am troubled by aches, pains or other physical problems only occasionally or sometimes	1
	I am troubled by aches, pains or other physical problems often, most or all the time	2

Table 2. Health states of the emotional component of CORE-6D as identified by the item threshold map and frequency of each health state in the study sample.

[adapted from Mavranezouli et al., Quality of Life Research 2011; 20(3): 321-33]

Item	Health states										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
I feel terribly alone and isolated	N	S	S	S	S	O	O	O	O	O	O
I feel panic or terror	N	N	S	S	S	S	O	O	O	O	O
I feel humiliated or shamed by other people	N	N	N	S	S	S	S	O	O	O	O
I am able to do most things I need to	O	O	O	O	S	S	S	S	S	N	N
I make plans to end my life	N	N	N	N	N	N	N	N	S	S	O
Frequency of each health state in the study sample	5.3%	5.9%	6.2%	5.0%	5.6%	2.7%	2.7%	1.5%	1.5%	0.0%	0.6%

N = never; S = only occasionally or sometimes; O = often, most or all the time; the 4th item is positively worded and therefore response levels are reversed

Table 3. Characteristics of respondents in the valuation survey and comparison with population characteristics for South Yorkshire and England

	Respondents (n=225)	South Yorkshire*	England*
Mean age (s.d.)	48.86 (17.16)	-	-
Age distribution			
18-40	32.7%	41.2%	41.6%
41-65	48.0%	39.1%	39.1%
Over 65	19.3%	19.7%	19.3%
Female	58.7%	51.2%	51.3%
Married/Partner	69.8%	NA	-
Employed or self-employed	51.3%	56.1%	60.9%
Unemployed	3.1%	4.1%	3.4%
Long-term sick	5.4%	7.7%	5.3%
Full-time student	5.4%	7.5%	7.3%
Retired	22.3%	14.4%	13.5%
Own home outright or with a mortgage	81.0%	64.0%	68.7%
Renting property	20.0%	36.0%	31.3%
Secondary school is highest level of education	37.9%	NA	-
Average EQ-5D score (s.d.)	0.83(0.28)	NA	0.86(0.23)†
TTO completion rate	99.7%	-	-
Respondent found 1st rank valuation task difficult	27.6%	-	-
Respondent found 1st TTO valuation task difficult	31.1%	-	-
Interviewer doubted whether respondent understood 1st rank task	5.8%	-	-
Interviewer doubted whether respondent understood 1st TTO task	4.9%	-	-

* 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 surveyed persons aged 18 and above only. Age distribution is here reported as the percentage of all adults aged 18 and over.

† Interviews conducted in the Measurement and Valuation of Health (MVH) study.[25]

Table 4. TTO values by health state obtained in the valuation survey

CORE-6D health state	TTO value								
	N	Mean	SD	Minimum	Percentile 25	Median	Percentile 75	Maximum	Mode
000000	75	0.96	0.13	0.08	0.99	1.00	1.00	1.00	1.00
000001	75	0.93	0.14	0.33	0.93	1.00	1.00	1.00	1.00
000002	76	0.82	0.32	-0.93	0.78	0.93	1.00	1.00	1.00
100000	74	0.87	0.22	0.08	0.84	1.00	1.00	1.00	1.00
110000	75	0.88	0.25	-0.73	0.85	1.00	1.00	1.00	1.00
110001	76	0.86	0.27	-0.93	0.80	0.96	1.00	1.00	1.00
110002	75	0.74	0.31	-0.83	0.57	0.83	1.00	1.00	1.00
111000	74	0.79	0.29	-0.23	0.69	0.93	1.00	1.00	1.00
111100	74	0.76	0.33	-0.40	0.53	0.93	1.00	1.00	1.00
211100	75	0.66	0.35	-0.63	0.50	0.73	1.00	1.00	1.00
221100	76	0.57	0.44	-0.93	0.45	0.63	0.93	1.00	1.00
221101	74	0.49	0.47	-0.88	0.30	0.50	0.88	1.00	1.00
221102	74	0.40	0.49	-0.93	0.14	0.44	0.83	1.00	1.00
222100	74	0.47	0.43	-0.93	0.20	0.50	0.84	1.00	1.00
222110	75	0.38	0.45	-0.98	0.08	0.44	0.70	1.00	1.00
222220	225	0.23	0.52	-0.98	0.00	0.30	0.53	1.00	1.00
222221	74	0.21	0.50	-0.93	-0.08	0.23	0.50	1.00	1.00
222222	75	0.10	0.53	-0.93	-0.33	0.10	0.48	1.00	1.00

Table 5. Mean TTO values for each CORE-6D health state included in valuation survey by severity of emotional and physical symptoms (standard deviation in parenthesis). Each card bloc is highlighted in a different shade; all respondents valued state 222220, shaded in black.

CORE-6D	Response levels of physical item		
Emotional component	0	1	2
00000	0.96 (0.13)	0.93 (0.14)	0.82 (0.32)
10000	0.87 (0.22)		
11000	0.88 (0.25)	0.86 (0.27)	0.74 (0.31)
11100	0.79 (0.29)		
11110	0.76 (0.33)		
21110	0.66 (0.35)		
22110	0.57 (0.44)	0.49 (0.47)	0.40 (0.49)
22210	0.47 (0.43)		
22211	0.37 (0.45)		
22221			
22222	0.23 (0.52)	0.21 (0.50)	0.10 (0.53)

Table 6. Regression models for prediction of mean TTO values (y) from Rasch model rescaled logit values (R) after adding 2 dummy variables (P_1, P_2) to account for the response level of the physical item (response levels 1 and 2, respectively)

Model	α	β_1	β_2	β_3	γ_1	γ_2	Adjusted R-Squared	RMSE
Model 1 $y = \alpha + \beta_1 R + \gamma_1 P_1 + \gamma_2 P_2$	0.008 (0.833)	1.057 (0.000)			-0.044 (0.189)	-0.151 (0.000)	0.961	0.0533
Model 2 $y = \alpha + \beta_2 R^2 + \gamma_1 P_1 + \gamma_2 P_2$	0.302 (0.000)		0.844 (0.000)		-0.070 (0.219)	-0.177 (0.006)	0.886	0.0916
Model 3 $y = \alpha + \beta_3 R^3 + \gamma_1 P_1 + \gamma_2 P_2$	0.416 (0.000)			0.779 (0.000)	-0.085 (0.284)	-0.193 (0.025)	0.773	0.1292
Model 4 $y = \alpha + \beta_1 R + \beta_2 R^2 + \gamma_1 P_1 + \gamma_2 P_2$	-0.130 (0.100)	1.585 (0.000)	-0.443 (0.056)		-0.029 (0.329)	-0.137 (0.000)	0.969	0.0478
Model 5 $y = \alpha + \beta_1 R + \beta_3 R^3 + \gamma_1 P_1 + \gamma_2 P_2$	-0.108 (0.072)	1.388 (0.000)		-0.282 (0.025)	-0.028 (0.329)	-0.135 (0.000)	0.972	0.0452
Model 6 $y = \alpha + \beta_2 R^2 + \beta_3 R^3 + \gamma_1 P_1 + \gamma_2 P_2$	0.099 (0.002)		2.624 (0.000)	-1.758 (0.000)	-0.029 (0.170)	-0.137 (0.000)	0.985	0.0331
Model 7 $y = \alpha + \beta_1 R + \beta_2 R^2 + \beta_3 R^3 + \gamma_1 P_1 + \gamma_2 P_2$	0.366 (0.004)	-1.695 (0.022)	5.712 (0.000)	-3.446 (0.000)	-0.033 (0.069)	-0.141 (0.000)	0.990	0.0275

RMSE = root mean squared error; *p* values in parenthesis

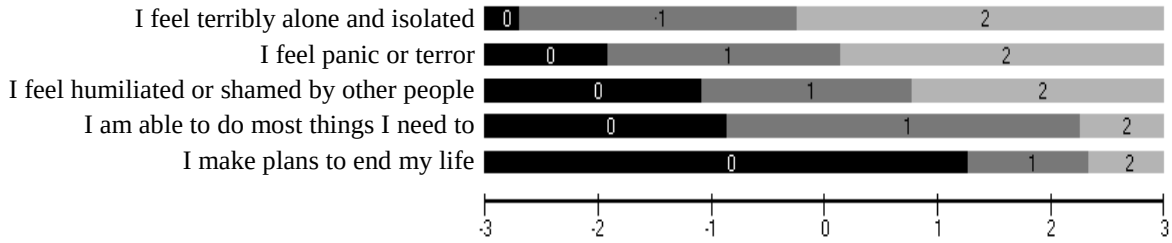
Figure legends:

Figure 1. Rasch item threshold map of the emotional component of CORE-6D

Figure 2. Mean observed (from the valuation survey) and modelled (based on regression model 7) TTO values by Rasch rescaled logit value.

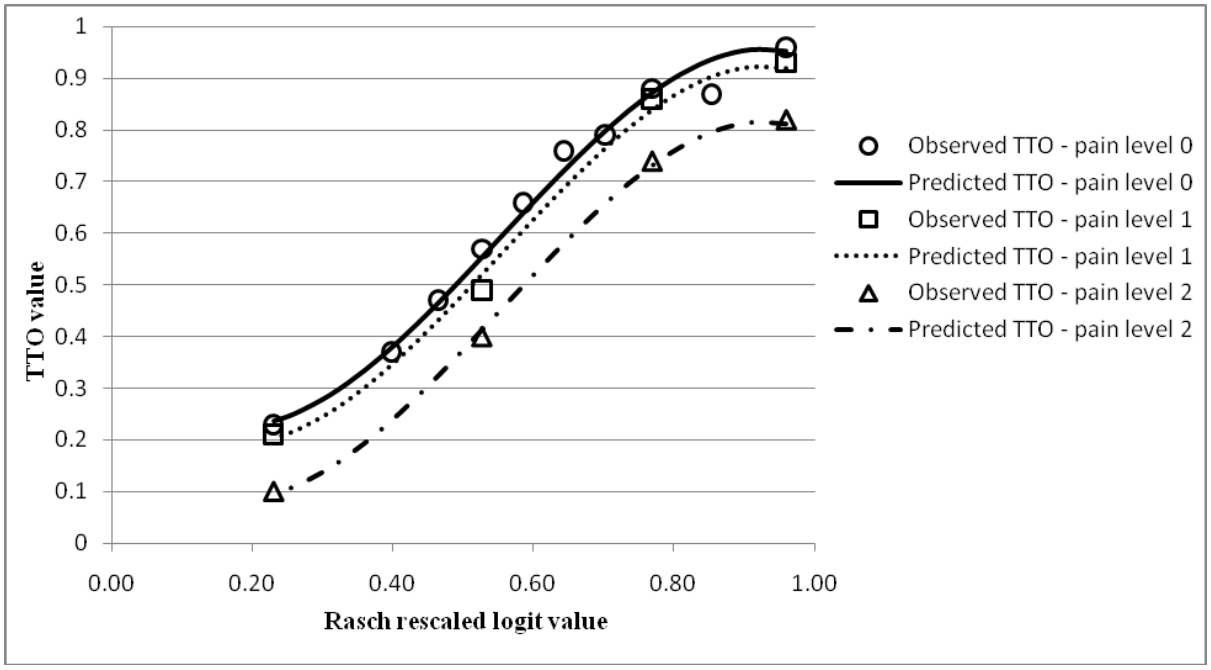
Figure 1.

[adapted from Mavranezouli et al., Quality of Life Research 2011; 20(3): 321-33]



0 = never; 1 = only occasionally or sometimes; 2 = often, most or all the time; note that the fourth item is positively worded and therefore response levels are reversed

Figure 2.



TTO = time trade-off

Note: Modelled TTO values are predicted using the Rasch rescaled logit value of the emotional health state and the response level of the physical item 'I am troubled by aches, pains, physical problems' (level 0 = never; level 1 = only occasionally or sometimes; level 2 = often, most or all the time)