Health Economists' Study Group Meeting, June 2010, University College Cork, Ireland

The development of a condition-specific preference-based measure for common mental disorders from the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM) using Rasch analysis

Ifigeneia Mavranezouli[1,2], John Brazier[1], Donna Rowen[1], Michael Barkham[3]

1. Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, Sheffield, UK
2. National Collaborating Centre for Mental Health, Centre for Outcomes Research and Effectiveness, Research Department of Clinical, Educational & Health Psychology, University College London, London, UK
3. Centre for Psychological Services Research, University of Sheffield, Sheffield, UK

Corresponding author:
Ifigeneia Mavranezouli
National Collaborating Centre for Mental Health
Centre for Outcomes Research and Effectiveness (CORE)
Research Department of Clinical, Educational & Health Psychology
University College London
Gower Street, London WC1E 6BT
Tel: +44 (0)207 679 1964
Fax: +44 (0)207 91 68 511
e-mail: i.mavranezouli@ucl.ac.uk

Acknowledgements: Ifigeneia Mavranezouli undertook this work as a PhD student in the School of Health and Related Research at the University of Sheffield. This study was funded by MRC-NIHR Methodology Research Programme, project number 06/97/04. We are grateful to the CORE System Trustees for giving approval for the development of the CORE-6D.

INTRODUCTION

Cost-utility analysis is a form of economic evaluation increasingly used for the assessment of health technologies and programmes worldwide. The most common outcome measure in cost-utility analysis is the Quality Adjusted Life Year (QALY). The number of QALYs in a health state is estimated by multiplying the time spent in this particular health state by the Health Related Quality of Life (HRQoL) experienced in this health state, which is measured using people's preferences reflected in a health state *utility value*. Preference-based measures (PBMs) are instruments that allow valuation of the HRQoL experienced in each of the health states characterising a clinical condition, disease area or patient population, resulting in estimation of utility values that can be used for calculation of QALYs.

Several PBMs have been successfully developed for this purpose. Among the most widely used are the EuroQol-5D (EQ-5D),[1] the SF-6D,[2] and the HUI-3.[3] These measures consist of a health state descriptive system and an algorithm that converts each health state into a utility value. For example EQ-5D consists of 5 items with 3 levels of response each, thus describing $3^5 = 243$ distinct health states. The health state descriptive system of EQ-5D has been valued by a representative sample of the general population in the UK (as well as in a number of other countries) using Visual Analogue Scale[4] and Time Trade-Off techniques.[5] Subsequent econometric modelling has led to the development of an algorithm that links each of the 243 health states with a utility value,[6] thus allowing the use of EQ-5D for generation of QALYs in cost-utility analysis.

All three measures cited above are generic and can therefore be used for the assessment of interventions and programmes targeted at different disease areas and patient populations. Various bodies worldwide, such as the National Institute for Health and Clinical Excellence (NICE) in England and Wales[7] and the US Panel on Cost-Effectiveness in Health and Medicine, a non-federal expert panel convened by the US Public Health Service,[8] support the use of generic PBMs in cost-utility analysis, because it allows comparability and consistency in the allocation of healthcare resources across a spectrum of technologies, disease areas and patient populations.

However, generic PBMs may be less appropriate or sensitive for capturing the HRQoL in some medical conditions.[9,10] Especially in the area of mental health, there are concerns that generic PBMs may lack sensitivity in capturing important elements of 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 condition-specific PBM that will be 'generic' for mental health, that is, it will be suitable for use across a wide range of mental health conditions.[11-13] Currently there is no such generic mental health PBM 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.[14]

CORE-OM is a 34-item instrument developed to evaluate the effectiveness of psychological therapies in people with common mental disorders.[15] 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 dimensional structure of CORE-OM is presented in Table 1. The validity, reliability and acceptability of CORE-OM as an effectiveness measure has been demonstrated across a wide range of practice settings.[16,17] Moreover, it is a widely used patient-based tool for measuring mental health outcomes in the British National Health Service.[14,18] Based on these characteristics and given the arguments favouring the development of a generic mental health PBM, CORE-OM was selected as the basis for constructing a PBM specific to 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, further 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.[19] The objective of this paper is to report on the latter 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.

The development of CORE-6D

The methodology used to derive CORE-6D from CORE-OM has been published elsewhere.[19] Here we summarise the results of that piece of work. CORE-6D is a 6-item health descriptive system consisting of a 5-item unidimensional 'emotional' component and a physical item. Each item has 3 response levels: 'never', 'only occasionally or sometimes' and

'often, most or all the time'. The descriptive system describes $3^6 = 729$ unique health states. The emotional component of CORE-6D has been derived from CORE-OM using predominantly Rasch analysis[20] (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 2.

Principles of Rasch analysis

Rasch analysis is a statistical approach for examining people's properties, such as knowledge, skills and perceptions ('latent traits'); it is based on the principles of the Rasch model,[20] 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[21] and demonstrates what the expected responses to items should be, if interval scale measurement is to be achieved.[22] Rasch analysis orders persons according to their ability (or 'severity', depending on the latent trait), and ranks questionnaire items according to their difficulty.[23] 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.[23] 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.[22] Rasch analysis has been successfully used as a tool for the development of various condition-specific PBMs.[24-27]

METHODS

Selection of health states for the valuation survey

The emotional component of CORE-6D can describe $3^5 = 243$ health states. However, as this constitutes a unidimensional scale, 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 am able to do most things I need to often, most or all the time". 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.[19] 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 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 helped identify the most likely item response combinations across the continuum of the emotional symptom severity. Figure 1 provides the item threshold map of the emotional component of CORE-6D. 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 3, 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 3) and was therefore excluded from further consideration. All 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 3) 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, as described below, which will be analysed elsewhere.

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. The version of TTO used here asks people to imagine they will be in a particular health state for 10 years, and then asks them to consider trading these years for a number of shorter periods in full health (x years). At the point where respondents are unable to choose, the value of the state is given as x/10.

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. 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 eight TTO valuations of health states. 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.[6] 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.

Three card blocks each containing 8 cards were used at valuation. Each card described one health state, based on a selection of emotional health states combined with different 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 (i.e. never troubled by physical problems). Responses to this card block were used to inform another piece of work currently under way as part of the same MRC-NIHR funded project,[28] the objective of which is to explore whether (non)reference to an extra aspect of health (in this case physical health) affects public's preferences (see companion HESG paper).[29] Due to this study design 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. State 222220 was included in all 3 card blocks.

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

The relationship between the Rasch model logit values and the respective TTO utility values of health states described by unidimensional measures has already been established by Young et al. and used at the development of PBMs from existing unidimensional scales.[30] However, this method is 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 used as a basis the methodology described by Young et al[30] for unidimensional measures, but also created

dummy variables to represent the different severity levels of the physical item, which is a standard approach used for multidimensional measures.[2,6]

Consequently, a series of regression analyses were undertaken 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 square of the correlation coefficient, R_c^2) 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 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 4, which allows comparison of the study sample to the general population in South Yorkshire and England. The study sample had a higher average age and a higher proportion of females, home owners and retired individuals and a lower proportion of employed/self-employed individuals. A significant proportion of respondents reported that they found the rank and TTO tasks difficult (27.6% and 31.1% respectively). The interviewers regarded it was doubtful whether the respondent understood the rank and TTO tasks in 5.8% and 4.9% of the interviews, respectively.

Utility values obtained from the valuation survey

The TTO values obtained from the valuation survey are reported in Table 5 and Table 6. Table 5 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 6 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 (from 0.87 to 0.88, respectively) despite of the increase in emotional symptom severity.

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 a TTO value 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 relationship) and 2 dummy variables accounting for the response level of the physical item.

The following models were tested:

Model 1 – simple linear relationship: $y = \text{constant} + b_1R + c_1P_1 + c_2P_2$

Model 2 – quadratic relationship: $y = \text{constant} + b_2R^2 + c_1P_1 + c_2P_2$

Model 3 – cubic relationship: $y = \text{constant} + b_3R^3 + c_1P_1 + c_2P_2$

Model 4 – quadratic relationship: $y = \text{constant} + b_1R + b_2R^2 + c_1P_1 + c_2P_2$

Model 5 – cubic relationship: $y = \text{constant} + b_1R + b_3R^3 + c_1P_1 + c_2P_2$

Model 6 – cubic relationship: $y = \text{constant} + b_2R^2 + b_3R^3 + c_1P_1 + c2P_2$

Model 7 – cubic relationship: $y = \text{constant} + b_1R + b_2R^2 + b_3R^3 + c_1P_1 + c_2P_2$

[*] Rescaling was achieved using the formula: $z_i = max_{new} + r \cdot (min_{Rasch} \cdot x_i)$, where:
z_i = Rasch model rescaled logit value of emotional state i;
x_i = Rasch model original logit value of emotional state i
max_{new} = maximum value of the new scale
min_{Rasch} = minimum value of the Rasch original logit scale
r = range of the new scale / range of the Rasch original logit scale

where y is mean predicted TTO value, R is 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), and b_i and c_i are regression coefficients.

The regression coefficients and goodness of fit statistics for all 7 models are shown in Table 7. The R_c^2 statistics varied from 0.813 (model 3) to 0.993 (model 7). Dummy variable P_1 was non-significant in all models. In model 7 the level of significance was slightly above 0.05 (0.069). Based on the RMSE statistics, model 7 was selected to be used for the prediction of TTO values for all health states described by CORE-6D. Figure 2 allows the comparison between actual TTO values obtained from the valuation survey for the selected CORE-6D health states, and predicted TTO values for all health states described by CORE-6D, derived from the regression model 7.

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 effectiveness 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.[19] 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 for 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.[19] 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.[30] 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 (such as the emotional component of CORE-6D). Our study successfully adapted this approach by incorporating dummy variables in regression analysis, to account for the different severity levels of the physical item of CORE-6D (a standard approach used to model utility values for multidimensional measures).[2,6] Our mixed 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 R_c^2 = 0.993 and RMSE = 0.0275, which can be used to predict utility values for all 729 health states described by CORE-6D. The new PBM 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.[11,31] One of the limitations 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 to mental disorders other than depression and anxiety, and therefore cannot be used as a 'generic' mental health PBM. Nevertheless, 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).[32]

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.

The appropriateness and sensitivity of the new measure is going to be assessed as a next step of this study, with CORE-6D being compared with the generic EQ-5D using a dataset of patients with common mental disorders that have received self-help psychological therapy, which contains data on both CORE-OM and EQ-5D.[33] 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.

REFERENCES

1. Brooks R. EuroQol: the current state of play. Health Policy 1996; 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; 21(2):271-292.
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; 40(2):113-128.
4. Gudex C, Dolan P, Kind P, Williams A. Health state valuations from the general public using the visual analogue scale. Qual Life Res 1996; 5(6):521-531.
5. Dolan P, Gudex C, Kind P, Williams A. The time trade-off method: results from a general population study. Health Econ 1996; 5(2):141-154.
6. Dolan P. Modeling valuations for EuroQol health states. Med Care 1997; 35(11):1095-1108.
7. National Institute for Health and Clinical Excellence. Guide to the Methods of Technology Appraisal. London: National Institute for Health and Clinical Excellence, 2008.
8. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-effectiveness in Health and Medicine. JAMA 1996; 276(15):1253-1258.
9. 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.
10. Brazier J, Tsuchiya A. Preference-based condition-specific measures of health: what happens to cross programme comparability? Health Econ 2010; 19(2):125-129.
11. Brazier J. Measuring and valuing mental health for use in economic evaluation. J Health Serv Res Policy 2008; 13 Suppl 3:70-75.
12. Chisholm D, Healey A, Knapp M. QALYs and mental health care. Soc Psychiatry Psychiatr Epidemiol 1997; 32(2):68-75.
13. Knapp M, Mangalore R. "The trouble with QALYs..." Epidemiol Psichiatr Soc 2007; 16(4):289-293.
14. Jacobs R. Investigating Patient Outcome Measures in Mental Health. CHE Research Paper Number 48. York: Centre for Health Economics, University of York, 2009.
15. 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; 69(2):184-196.

16. 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; 186:239-246.
17. 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; 180:51-60.
18. 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; 47(Pt 4):397-415.
19. Mavranezouli I, Brazier JE, Young TA, Barkham M. Using Rasch analysis to form plausible health states amenable to valuation: the development of CORE-5D utility from CORE-OM in order to elicit preferences for common mental health problems. Health Economics and Decision Science Discussion Paper 09/09. Sheffield: ScHARR, University of Sheffield, 2009.
20. Rasch G. Probabilistic models for some intelligence and attainment tests. Chicago: University of Chicago Press, 1960.
21. Tesio L, Simone A, Bernardinello M. Rehabilitation and outcome measurement: where is Rasch analysis-going? Eura Medicophys 2007; 43(3):417-426.
22. 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; 57(8):1358-1362.
23. Bond TG, Fox CM. Applying the Rasch Model: Fundamental Measurement in the Human Sciences. Mahwah, New Jersey: Lawrence Erlbaum Associates Inc, 2006.
24. McTaggart-Cowan H, Brazier J, Tsuchiya A. Combining Rasch And Cluster Analysis: A Novel Method For Developing Rheumatoid Arthritis States For Use In Valuation Studies. Health Economics and Decision Science Discussion paper No. 08/15. Sheffield: School of Health and Related Research, 2009.
25. 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; 62(8):845-856.

26. 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; 18(2):253-265.
27. Young T, Yang Y, Brazier J, Tsuchiya A. The use of Rasch analysis as a tool in the construction of a preference based measure: the case of AQLQ. Health Economics and Decision Science Discussion Paper No. 07/01. Sheffield: School of Health and Related Research, 2007.
28. NIHR Methodology Panel project. Developing and testing methods for deriving preference-based measures of health from condition specific measures (and other patient based measures of outcome) - the CoSMeQ (Condition-Specific Methodology for estimating QALYs) study. Project number 06/97/04. Chief Investigator: Brazier J.
29. Brazier J, Rowen D, Tsuchiya A, et al. What a pain: adding a generic dimension to a condition-specific preference-based measure. Paper presented at HESG meeting, University College Cork, June 2010.
30. 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; DOI 10.1007/s11136-010-9646-0.
31. 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; 17(3):485-495.
32. 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.
33. Richards A, Barkham M, Cahill J, Richards D, Williams C, Heywood P. PHASE: a randomised, controlled trial of supervised self-help cognitive behavioural therapy in primary care. Br J Gen Pract 2003; 53(495):764-770.
34. Kind P, hardman G, Macran S. UK population norms for EQ-5D. York, Centre for Health Economics Discussion Paper Series, University of York, 1999.

Table 1. The dimensional structure of CORE-OM

Domain	Item N^{o}	Item
Subjective Well Being	4	I have felt ok about myself
	14	I have felt like crying
	17	I have felt overwhelmed by my problems
	31	I have felt optimistic about my future
Symptoms – anxiety	2	I have felt tense, anxious or nervous
	11	Tension/anxiety have prevented me doing important things
	15	I have felt panic or terror
	20	My problems have been impossible to put to one side
Symptoms – depression	5	I have felt totally lacking in energy and enthusiasm
	23	I have felt despairing or hopeless
	27	I have felt unhappy
	30	I have thought I am to blame for my problems & difficulties
Symptoms – physical	8	I have been troubled by aches, pains, physical problems
	18	I have had difficulty of getting to sleep or staying asleep
Symptoms – trauma	13	I have been disturbed by unwanted thoughts and feelings
	28	Unwanted images or memories have been distressing me
Functioning – general	7	I have felt able to cope when things go wrong
	12	I have been happy with the things I've done
	21	I have been able to do most things I needed to
	32	I have achieved the things I wanted to
Functioning – close relationships	1	I have felt terribly alone and isolated
	3	I have felt I have sb to turn to for support when needed
	19	I have felt warmth or affection for someone
	26	I have thought I have no friends
Functioning – social relationships	10	Talking to people has felt too much for me
	25	I have felt riticized by other people
	29	I have been irritable when with other people
	33	I have felt humiliated or shamed by other people
Risk/harm to self	9	I have thought of hurting myself
	16	I made plans to end my life
	24	I have thought it would be better if I were dead
	34	I have hurt myself physically or taken risks with my health
Risk/harm to others	6	I have been physically violent to others
	22	I have threatened or intimidated another person

Table 2. 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, physical problems	0
	I am troubled by aches, pains, physical problems only occasionally or sometimes	1
	I am troubled by aches, pains, physical problems often, most or all the time	2

Table 3. 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.

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

	Respondents (n=225)	South Yorkshire[1]	England[1]
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)[2]
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%	-	-

Notes: 1. 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.
2. Interviews conducted in the Measurement and Valuation of Health (MVH) study.[34]

Table 5. 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 6. Mean TTO values for each CORE-6D health state included in valuation survey by severity of emotional and physical symptoms (standard deviation in parenthesis)

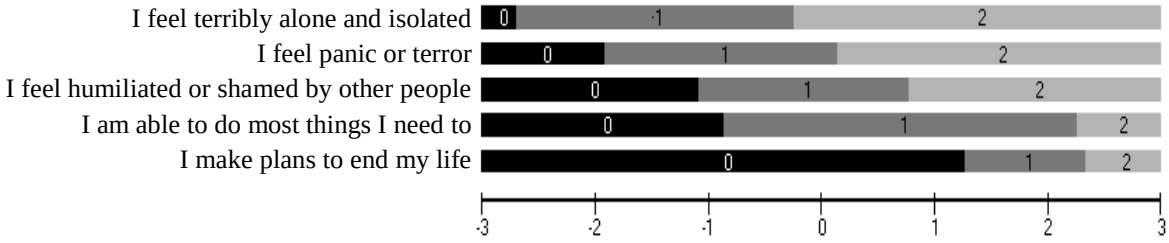
CORE-6D	Response levels of physical item		
Emotional component	0	1	2
00000	0.96 (0.13)	0.93 (0.14)	0.82 (0.32)
10000	0.87 (0.22)		
11000	0.88 (0.25)	0.86 (0.27)	0.74 (0.31)
11100	0.79 (0.29)		
11110	0.76 (0.33)		
21110	0.66 (0.35)		
22110	0.57 (0.44)	0.49 (0.47)	0.40 (0.49)
22210	0.47 (0.43)		
22211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

Table 7. Regression model results and goodness of fit statistics for predicting 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	Constant	b_1	b_2	b_3	c_1	c_2	R_c^2	RMSE
Model 1 $y = constant + b_1R + c_1P_1 + c_2P_2$	0.008 (0.036)^{ns}	1.057 (0.052)			-0.044 (0.032)^{ns}	-0.151 (0.032)	0.961	0.0533
Model 2 $y = constant + b_2R^2 + c_1P_1 + c_2P_2$	0.302 (0.043)		0.844 (0.075)		-0.070 (0.054)^{ns}	-0.177 (0.054)	0.906	0.0916
Model 3 $y = constant + b_3R^3 + c_1P_1 + c_2P_2$	0.416 (0.052)			0.779 (0.103)	-0.085 (0.077)^{ns}	-0.193 (0.077)	0.813	0.1292
Model 4 $y = constant + b_1R + b_2R^2 + c_1P_1 + c_2P_2$	-0.130 (0.073)^{ns}	1.585 (0.255)	-0.443 (0.211)		-0.029 (0.029)^{ns}	-0.137 (0.029)	0.976	0.0478
Model 5 $y = constant + b_1R + b_3R^3 + c_1P_1 + c_2P_2$	-0.108 (0.055)^{ns}	1.388 (0.138)		-0.282 (0.111)	-0.028 (0.027)^{ns}	-0.135 (0.027)	0.979	0.0452
Model 6 $y = constant + b_2R^2 + b_3R^3 + c_1P_1 + c_2P_2$	0.099 (0.026)		2.624 (0.186)	-1.758 (0.181)	-0.029 (0.020)^{ns}	-0.137 (0.020)	0.985	0.0331
Model 7 $y = constant + b_1R + b_2R^2 + b_3R^3 + c_1P_1 + c_2P_2$	0.366 (0.104)	-1.695 (0.645)	5.712 (1.184)	-3.446 (0.660)	-0.033 (0.017)^{ns}	-0.141 (0.017)	0.993	0.0275

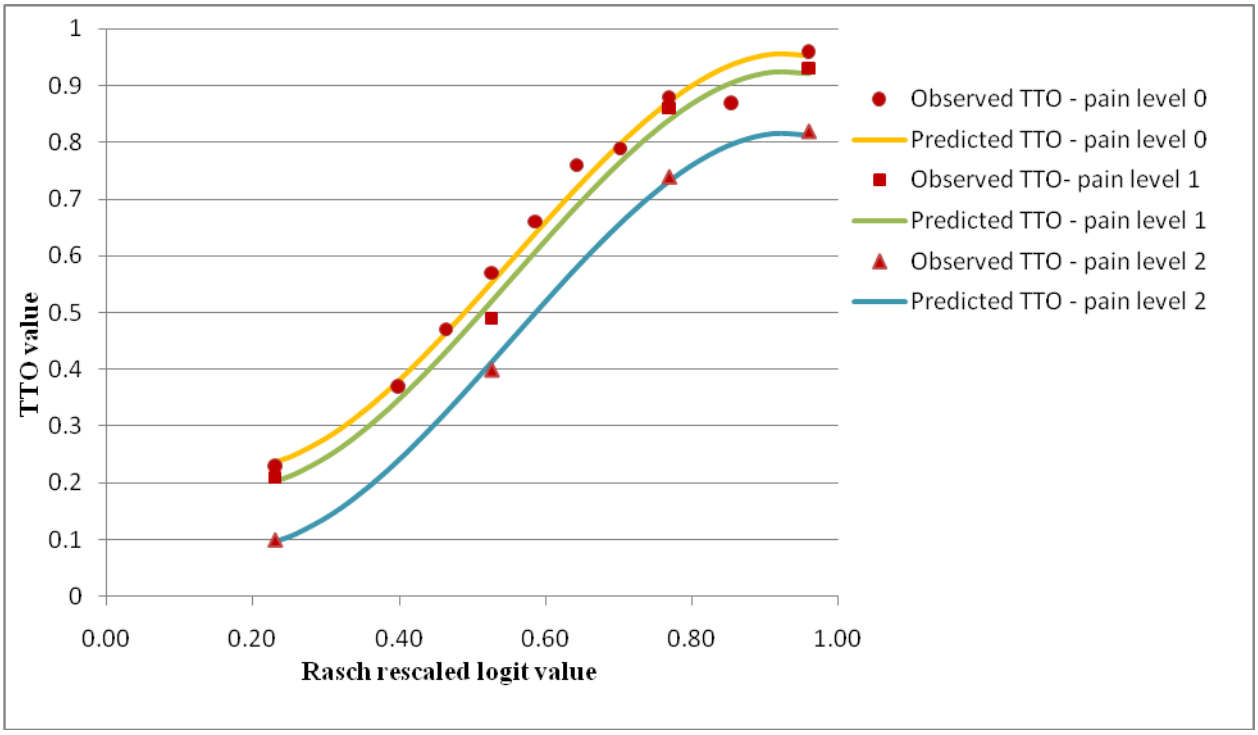
R_c = correlation coefficient; RMSE = root mean squared error

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



0 = never; 1 = only occasionally or sometimes; 2 = often, most or all the time; note that 4th item is positively worded and therefore response levels are reversed

Figure 2. Mean observed TTO values (from the valuation survey) and predicted TTO values (based on regression model 7). Predicted TTO values depend on the Rasch model rescaled logit value of each 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)



TTO = time trade-off