Monograph:
Mulhern, B, Rowen, D, Brazier, J et al. (6 more authors) (2010) Developing a health state classification system from NEWQOL for epilepsy using classical psychometric techniques and Rasch analysis: A technical report. Discussion Paper. (Unpublished)

HEDS Discussion Paper 10/16


ScHARR
School of Health and
Related Research

Developing a health state classification system from NEWQOL for epilepsy using classical psychometric techniques and Rasch analysis: A Technical report

Brendan Mulhern (MRes)[1], Donna Rowen (PhD)[1], John Brazier (PhD)[1], Ann Jacoby (PhD)[2], Tony Marson (FRCP)[3], Dee Snape (MA)[2], Dyfrig Hughes(PhD)[4] Nick Latimer(MSc)[1] & Gus A. Baker (FBPsS)[3]

1 Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, UK

2 School of Population Community and Behavioural Sciences, University of Liverpool, UK

3 School of Clinical Sciences, University of Liverpool, UK

4 Centre for Economics and Policy in Health, University of Bangor, UK

Acknowledgements: This study was funded by Epilepsy UK

Corresponding author:

Brendan Mulhern, School of Health and Related Research, University of Sheffield, Regent Court, 30 Regent St, Sheffield, S1 4DA, United Kingdom, e-mail: b.mulhern@sheffield.ac.uk

Tel no: +44 (0)114 222 0794

Fax no: +44 (0)114 272 4095

Key words: Quality Adjusted Life Years, Health Related Quality of Life, Rasch Analysis, Preference-Based Measures of health, Health states, Epilepsy

Abstract

Aims: Resource allocation amongst competing health care interventions is informed by evidence of both clinical- and cost-effectiveness. Cost-utility analysis is increasingly used to assess cost effectiveness through the use of Quality Adjusted Life Years (QALYs). This requires health state values. Generic measures of health related quality of life (HRQL) are usually used to produce these values, but there are concerns about their relevance and sensitivity in epilepsy. This study develops a health state classification system for epilepsy from the NEWQOL battery, a validated questionnaire measuring QoL in epilepsy. The classification system will be amenable to valuation for calculating QALYs.

Methods: Factor and other psychometric analyses were undertaken to investigate the factor structure of the battery, and assess the validity and responsiveness of the items. These analyses were used alongside Rasch analysis to select the dimensions included in the classification system, and the items used to represent each domain. Analysis was carried out on a trial dataset of patients with epilepsy (n=1611). Rasch and factor analysis were performed on one half of the sample and validated on the remaining half. Dimensions and items were selected that performed well across all analyses.

Results: The battery was found to demonstrate reliability and validity but responsiveness across time periods for many of the items was low. A six dimension classification system was developed: worry about seizures, depression, memory, cognition, stigmatism and control, each with four response levels.

Conclusions: It is feasible to develop a health state classification system from a battery of instruments using a combination of classical psychometric, factor and Rasch analysis. This is the first condition-specific health state classification developed for epilepsy and the next stage will produce preference weights to enable the measure to be used in cost-utility analysis.

Introduction

As with healthcare in general, resources available for epilepsy are limited and need to be allocated efficiently. The allocation process is typically informed by economic evaluations of competing health technologies. Methods to evaluate the cost effectiveness of emerging interventions include the assessment of cost utility through the generation of Quality Adjusted Life Years (QALYs). QALYs provide a single measure of an individual's preference for a particular health state by combining an assessment of both quantity and quality of life (Torrance, 1986). In their guidelines for conducting economic evaluations, the National Institute for Health and Clinical Excellence (NICE) in England recommend the use of QALYs to measure the benefits of health interventions (National Institute for Health and Clinical Excellence, 2008).

QALYs are commonly generated using generic preference-based measures (PBM) of health such as the EQ-5D (Brooks, 1996; Dolan, 1997), or the SF-6D (Brazier, 2002). The EQ-5D is the measure for economic evaluations preferred by NICE and contains five dimensions (mobility, self care, usual activities, pain/discomfort, anxiety/depression), each with three health state levels. This means that it is possible to generate a total of 243 (3^5) health states from the instrument. To elicit preferences for each EQ-5D health state and to generate a preference-based single index score, a selection of the health states were valued using the standard preference elicitation technique Time Trade Off (TTO) developed by Torrance et al. (1972). The EQ-5D index scores range from -0.594 to 1, and are anchored at 0 for dead and 1 for full health (where minus scores are states that have been valued as worse than dead).

Epilepsy is a common neurological disorder that affects approximately 456,000 people of all ages in the United Kingdom (NHS, 2008). The condition is characterised by repeated seizures and the majority of recommended treatments are pharmacological. Research has focused on the positive and negative influences of experiencing epilepsy on health related quality of life (HRQL), and these include psychological comorbidities (Ettinger et al., 2004; Loring et al., 2004; Zeber et al., 2007),

stigma (Suurmeijer, 2001) and frequency or freedom from seizure (Birbeck et al., 2002; Jacoby et al., 1992). Reviews of this area have been published by Jacoby et al. (2008; 2009). However there is still more work required to define further the psychosocial and economic consequences of epilepsy.

Generic PBMs are generally used in the economic evaluation of interventions for epilepsy. However there is debate around their validity, and therefore the extent to which an accurate assessment of epilepsy specific cost utility can be made. For example, Stavem et al. (2001) found that some of the EQ-5D dimensions discriminated well by seizure status, but were less valid in patients who have used antiepileptic drugs and those with neurologic comorbidities. Selai et al (2000) found that the EQ-5D was not capturing all of the HRQL issues of relevance to patients with chronic intractable epilepsy, and the measure did not display responsiveness within this group.

Due to the uncertainty around using generic PBMs, there has been recent interest in developing condition-specific PBMs from non-preference-based condition-specific measures of Health Related Quality of Life (HRQL) (Brazier et al, 2008; Yang et al; 2008; Young et al, 2009). Standard condition-specific HRQL measures do not generate single index utility scores and therefore cannot be used to calculate QALYs. However they provide a strong basis for the first stage of the development of condition-specific PBMs which is the generation of a reduced health state classification system from the parent measure that is amenable to subsequent valuation. In recent work, a combination of classical psychometric and Rasch analysis (Rasch, 1960) has been used to develop condition-specific health state classification systems for overactive bladder syndrome (Young et al., 2009), asthma (Young et al., 2010), flushing symptoms (Young et al. 2010) and dementia (Mulhern et al., 2010). Classical psychometric techniques and Rasch analysis are used to investigate the factor structure of the instrument and the analyses are combined with clinical input to subsequently select an item that represents each factor. Combining these techniques is an accepted method of developing HRQL instruments (Tennant et al, 2004).

This technical report describes the first stage of the development of an epilepsy specific PBM. This involves the application of psychometric and Rasch analysis to develop a condition-specific health state classification system from an existing measure of HRQL in epilepsy. The parent instrument used is the NEWQOL (Abetz et al., 2000) which has been validated for use with patients with newly diagnosed epilepsy and has been used as a measure of HRQL in a large scale randomised controlled trial (RCT; Marson et al., 2005; 2007). The reduced health state classification system will subsequently be valued by both patients and the general population using a standard preference elicitation technique. The general population weights will provide a tool to inform the economic evaluation of epilepsy interventions.

Method

Sample

The sample used in this study consists of 1611 respondents with newly diagnosed epilepsy. The data was collected as part of the SANAD study (Marson et al. 2005; 2007), an RCT of immediate and deferred antiepileptic drug treatment carried out in UK outpatient clinics. The baseline data was used for this analysis. Of the overall sample, 54% were male and the mean age was 39 (range 16 to 86). Furthermore, 70% reported that their general health was “good” or “better” and 40% had experienced 10 or more seizures. Classical psychometric analyses (including factor analysis) were carried out on the full sample. The optimum number for Rasch analysis is 500 (Linacre, 1999), so therefore the sample used here was randomly split into two halves and the analysis was carried out on the first random half of the data (initial analysis) and validated on the second random half (validation analysis).

Materials

The NEWQOL measure was developed by Abetz et al. (2000) to assess HRQL in recently diagnosed epilepsy, and a subset of the items are included in this study. NEWQOL consists of a range of validated measures developed both for general use across a range of conditions and specifically for epilepsy. The subset of NEWQOL measures and items (n=82) included in the initial analysis for this study assess mental health, cognition and neuropsychological problems, mastery/locus of control, stigma related to having epilepsy, the impact of epilepsy on a number of life areas, worry about seizures, social restrictions, and a full range of adverse events related to epilepsy. The measures included are the A-B Neuropsychological Assessment Schedule (ABNAS, Aldenkamp et al, 2002), Liverpool Adverse Events Profile (AEP, Baker et al., 1995), Hospital Anxiety and Depression Scale (HADS, Zigmond & Snaith, 1983), the mastery/locus of control scale (Pearlin & Schooler, 1974), stigma scale (Jacoby, 1994), the perceived impact of seizures scale (Jacoby et al., 1994), a social restriction item (Jacoby et al., 1992), and the seizure worry scale (Jacoby, 1994). Table 1 provides further details about each of the measures.

Analysis

The intention of the analysis was to derive a multidimensional patient reported health state classification system from the NEWQOL measure. The number of items included in the classification system is reduced to one per dimension whilst retaining as many of the epilepsy-specific HRQL concepts included in the original NEWQOL as possible. A series of steps guides the analysis process (see figure 1), which involves applying the classical psychometric methods and Rasch analysis outlined below to produce the final classification system. Input from a range of experts including epilepsy clinicians and the developers of a selection of the measures included in the NEWQOL battery were also involved in the selection of the classification system. This ensured that the item selected for each dimension was relevant to epilepsy and displayed good face validity. Item text and the associated response options that are selected for each dimension of the classification system

form the basis of the health state levels used to generate epilepsy specific health states that are subsequently valued. It is important to alter the text of the original item as little as possible so that responses can be clearly mapped onto the classification. Psychometric and factor analysis was carried out using SPSS version 16 (SPSS, 2007) and Rasch analysis using Rasch Unidimensional Measurement Models 2020 (Rasch Unidimensional Measurement Models, 1997-2004).

Stage 1: Establishing NEWQOL dimensionality:

The dimensionality of the NEWQOL and the domains to be included in the classification system were established using exploratory factor analysis and input from epilepsy clinicians. Factor analysis assesses the factor structure of instruments by examining the correlation between each item and a range of factor structures. These were defined both by using the standard criterion of eigenvalues > 1 and by forcing a range of solutions. Factor solutions with 4 to 12 factors were investigated. Items were removed from factors if they did not load ≥ 0.4 on any factor, or cross loaded within 0.2 across more than one factor (Ferguson & Cox, 1993).

Stage 2: Rasch analysis to eliminate items per dimension:

Rasch analysis is part of the item response theory (IRT) group of analysis techniques. Rasch converts responses to items into a continuous logit scale whereby the position of the individual is related to the severity of the underlying trait that the scale is measuring. In the development of health instruments the underlying trait is the aspect of HRQL that the item assesses. Item responses are assumed to be a function of the location of both the person and the item on the logit scale. Rasch analysis is applied to each of the dimensions. This is because the technique assumes unidimensionality so is therefore used to assess items that are measuring the same underlying

construct as suggested by the factor analysis. The following criteria guide the selection of items for each domain.

Item level ordering

The ordering of the item response categories is investigated. If items are disordered it demonstrates that responders cannot differentiate between item levels. The response levels of disordered items are merged and the Rasch model is refitted. If this occurs then the item is excluded from the selection process, although they are retained in the Rasch model.

Differential item functioning

Items are checked for Differential Item Functioning (DIF) which assesses whether responses to items differ dependent on participant characteristics. There are two types of DIF, uniform and non uniform. Uniform DIF occurs when responses between groups consistently differs across the logit scale (for example females consistently score more highly on a certain trait than males). Non-uniform DIF occurs when responses systematically differ at different severity levels. In this study DIF by both gender and age group (split as age 16-35 (47% of the sample), 36-55 (34%) and 56 or older (19%)) has been investigated. If any items display DIF they are split into component factors and the model is refitted. Items split for DIF are no longer considered for inclusion in the classification system.

Goodness of fit

Goodness of fit is investigated with the aim of removing items that do not fit the overall dimension level Rasch model and so are not consistent with the unidimensional scale. The objective is to ensure that the items included in the final model, and therefore available for selection, all fit the Rasch model and provide overall goodness of fit. This is done by studying item-trait interactions, and item and person fit residuals.

Item trait interactions:

Item trait interactions assess the fit of items to the model for responders (dependent on the position of the respondent lie on the latent scale). The overall difference between the observed and expected response is measured using the chi-squared ($\chi 2$) test statistic which is non-significant (i.e. > 0.01) for a model providing good fit. The lowest fitting items are removed sequentially until the remaining items fit the model and the overall fit statistic is non significant. Items that are removed from the model will not be considered for the final classification system.

Item/person fit residuals:

Item fit residuals assess the amount of divergence between the expected and observed responses for each item included in the model. Person fit residuals assess the difference for individual respondents. The mean fit residual should be approximately 0 and the standard deviation around 1, and residuals > 2.5 or < -2.5 are considered high and indicate a large divergence from what is expected for that item or individual. Items and persons significantly outside this range are removed and the model is refitted.

Stage 3: Psychometric and Rasch analysis to select one item per dimension

After applying the tests included at stage 2, most dimensions have more than one item that could be included in the classification system. Stage three involves selecting the most appropriate item from each dimension. A selection of classical psychometric and Rasch statistics guides this process. The psychometric tests carried out in this study included missing data, floor/ceiling effect and responsiveness analysis. The main Rasch criteria used are the item spread at logit 0 (i.e. the spread of response at the average item difficulty) and the item range. High spread and item range indicate that the item cover a large range of condition severity. Good item range incorporates values above

and below 0 (i.e. more severe and less severe cases respectively). Item goodness of fit statistics and clinical input are also used to guide the selection process.

Stage 4: Item level reduction

Analysis of the performance of each item level is carried out to investigate whether the number of response levels included for each item can be reduced. This is because it is essential to ensure that information relating to item levels is not redundant. It is also possible that respondents will have difficulties distinguishing between response levels, and therefore the valuation of the health states will be more complex than if the item levels are collapsed. The distribution of responses across each category is investigated, as is the ordering of the levels on the Rasch logit scale.

Results:

Stage 1: Establishing NEWQOL dimensionality:

Factor analysis:

The five factor solution explaining 53.3% of the variance in the model provided the best fit and included factors defined as cognition and memory, mental health, control, stigma and impact of epilepsy. The items included in each factor are displayed in table 2.

Selection of dimensions for the classification system:

A number of alterations and additions were made to the initial factor structure to ensure that the classification system included as many facets of epilepsy related HRQL as possible. It was also important to ensure that the items that were not amenable to the generation of health states were removed. This process was carried out by an expert panel who assessed each of the five factors as well as the items that were not included in the five factor model. Changes relating to four factors

are described below. Table 3 displays the dimensions that were used to develop the classification system:

Worry about seizures:

Following clinical input regarding the importance of worry relating to past and potential future seizures on epilepsy patients HRQL, it was decided to include an extra dimension investigating worry and anxiety about seizures. This dimension included two items investigating worry about past and future seizures.

Memory and cognition:

The memory and cognition factor investigates two different facets of HRQL. As no item covers both concepts, selecting one item for this factor would mean removing an area of HRQL from the classification system that may be important in a neuropsychological condition such as epilepsy. Following this it was decided to split the factor into memory and cognition sub-dimensions and perform the Rasch analysis on each. Most of the items in this factor were taken from the ABNAS. Therefore the items included in the sub-dimensions was informed by the original factor analysis of this measure (Aldenkamp and Baker, 1997) which found a six dimension structure, with factors defined as fatigue, slowing, memory, concentration, motor coordination and language. Items from the concentration, slowing and language factors were included in the cognition sub-dimension. Items from the memory factor were included in the memory sub-dimension. Items from the fatigue and motor coordination factors and non relevant somatic items from the AEP related to unsteadiness and dizziness were removed from the analysis at this stage.

Mental health:

The mental health factor includes items relating to concepts that could broadly be defined as depression and anxiety. Therefore it was decided to split the items into depression and anxiety sub-dimensions. The majority of the items in this factor are taken from the HADS, and therefore the sub-

dimension split was informed by the original development of this instrument which includes depression and anxiety subscales (Zigmond and Snaith, 1983). Research suggests that anxiety in epilepsy is associated with increased seizure activity (Jacoby et al., 1996) and poor seizure control (Mensah et al., 2007). Therefore as the worry about seizures dimension focuses more on epilepsy specific anxiety rather than the more general focus of the anxiety sub-dimension, it was decided to use the worry about seizures dimension to investigate anxiety. The anxiety sub-dimension was excluded from further consideration.

Impact of epilepsy

The wording of the impact factor items means that it is not possible to generate health state levels. This is because the items are transition questions that ask how epilepsy has affected 'for better or worse' an area of the patient's life (for example their relationship with their friends). Therefore this dimension was not included in the health state classification system.

Rasch analysis and item selection by dimension

The item selection process for each of the six NEWQOL factors is described below. Tables 4 and 5 display in detail psychometric and Rasch analysis results for each of the items included in each dimension.

Cognition:

Stage 2: Item elimination

The cognition sub-dimension includes 10 items from the ABNAS scale. Across both the initial and validation analyses, items b ('My mind does not work as fast as it should') and h ('My thinking has slowed down') display DIF by age and item j ('I have difficulty concentrating on the things that I am doing') does not fit the Rasch model. These items were therefore no longer considered for selection

to the health state classification system. Three further items were also excluded from selection due to both DIF by age and misfit to the Rasch model on either the initial or validation analysis. These are item d ('I have difficulties in following a book or film'), item f ('I have problems finding the correct word') and item v ('I sometimes stutter or am unable to find the correct words').

Stage 3: Item selection

Across the analyses, four items remained for selection. Of these, item l ('I have problems understanding what I read') displayed low spread at logit 0 so was not considered further. Items p ('I can't concentrate for more than a short period of time'), u ('I get distracted easily') and w ('I feel I react too slowly to things that are said to me') all display similar severity coverage and coverage at logit 0 across both analyses. Of these, p was selected as the item assesses a cognition related issue that is more general, and therefore more prevalent, than the concepts covered by the other remaining items.

Memory:

Stage 2: Item elimination

The memory component of the overall memory and cognition factor includes 5 items. Across both the initial and validation analysis, ABNAS item c ('I have difficulties remembering the names of people') displays uniform DIF by age and ABNAS o ('I forget things people have said to me') does not fit the model. Both items were removed from consideration for the health state classification. On the initial analysis, ABNAS t ('I get confused and forget what I was doing') displayed DIF by age and on the validation analysis ABNAS i ('I forget all kinds of things, for example an appointment') did not fit the Rasch model.

Stage 3: Item selection

The only remaining item from across both analyses was AEP 18 ('memory problems') and this item was selected for the health state classification system.

Depression:

Stage 2: Item elimination

Six items are included in the depression component of the mental health factor. Across both the initial and validation analysis, HADS items 2 ('I can still enjoy the things I used to enjoy') and 14 ('I can enjoy a good book or radio or TV program') were disordered and HADS 12 ('I look forward with enjoyment to things') displayed DIF by age. HADS item 6('I feel cheerful') displayed DIF by age on the validation analysis. These four items were not considered for the health state classification system.

Stage 3: Item selection:

Two items remain for selection to the classification system. These are AEP 17 ('Depression') and HADS 4 ('I can laugh and see the funny side of things'). Although HADS item 4 displays better range and spread than AEP item 17 it was decided to use AEP item 17 as the item assesses the overall factor concept directly. *Control:*

Stage 2: Item elimination

The control dimension includes 5 items, and the response categories for all of the items are ordered on the logit scale. Control items b ('I sometimes feel that I am pushed around in my life') and g ('There is little I can do to change many of the important things in my life') display DIF by age on both of the analyses, and control item e ('I often feel helpless in dealing with the problems of life') displays DIF by gender on the initial analysis.

Stage 3: Item selection

Two items (control a: 'There is really no way I can solve some of the problems I have' and control c: 'I have little control over things that happen to me') remain for selection. Of these, control c covers more of the severe end of the severity spectrum and displays larger spread at logit 0 so was therefore selected for the classification system

Stigma:

Stage 2: Item elimination

The stigma domain contains 3 items. Stigma item a ('I feel some people are uncomfortable with me') has been eliminated due to poor fit to the Rasch model and this is consistent across both halves of the analysis.

Stage 3: Item selection

On both the initial and validation analyses two stigma items remain for selection (Stigma b: 'I feel some people treat me like an inferior person' and stigma c: 'I feel some people would prefer to avoid me'). Both of the remaining items have a high ceiling effect, though item b displays slightly better severity coverage and overall item fit. Clinical input also suggests that item b may be the more able to discriminate between severity levels and therefore this item was selected for the health state classification.

Worry about seizures:

Stage 2: Item elimination:

Of the two items included in the worry factor, both are ordered on both the initial and validation analyses. Neither item displayed DIF and both displayed goodness of fit to the model so neither was eliminated at this stage.

Stage 3: Item selection:

Both of the items cover the full severity range in terms of item range and spread. As both items are valid for the health state classification it was decided to select the item assessing worry about future seizures as this is the most relevant for those with newly diagnosed epilepsy.

Stage 4: Item level ordering

The item level ordering and response distribution across the levels was investigated for each of the selected items. Each of the six items has 4 response levels and this was maintained for the classification system as Rasch analysis demonstrated that of all the categories were ordered on the logit scale, and would therefore be amenable to health state valuation.

Final health state classification:

The final six dimension health state classification is displayed in figure 1, with the final health states developed in accordance with the response levels assigned to the original item. The classification system generates a possible 4096 (4^6) health states, a selection of which will be valued by a general population and patient sample.

Discussion:

This technical report describes the development of a health state classification system for epilepsy from NEWQOL, a condition specific measure of HRQL. This was carried outusing a combination of classical psychometric techniques and Rasch analysis. We have completed the first stage of the process of developing an epilepsy specific preference based measure by identifying a tool that can now be valued using a standard preference elicitation technique. This work is the first attempt to derive a condition specific classification system for epilepsy for the purposes of cost utility analysis

using QALYs. It is also the first attempt to derive a classification system from a battery that includes a variety of standardised measures that were developed for epilepsy (for example ABNAS and AEP) and also a widely used measure (HADS) that was not specifically developed for use in epilepsy but is an accepted measure of depression and anxiety across a range of conditions (Bjellend et al., 2002). This is also the first reported study to develop a classification system for a neuropsychological condition.

This study has built on previous work by members of the research group using the results of Rasch and psychometric analysis to develop condition specific classification systems for over active bladder (Young et al, 2009), asthma (Young et al.,2010) and dementia (Mulhern et al., 2010). Again it has been demonstrated that these analyses can help guide the selection of items for a reduced health state classification system. The analysis quantifies the performance of the items and clinical input during the selection process maximises face validity and enables the best item to be selected if the Rasch statistics of a number of items are similar.

The use of generic measures such as the EQ-5D and SF-6D in epilepsy has been criticised as it has been found that they do not fully reflect the impact of the condition, and may not cover all of the relevant domains (Selai et al, 2000; Stavem et al, 2001). Therefore the cost utility estimations gained using generic measures may not be fully accurate. The final instrument that will be available following valuation may address some of these concerns. Further work should use both generic and condition specific PBMs in intervention trials both to increase the strength of the conclusions relating to the cost utility analysis and to subsequently assess the performance of the measures.

The classification system that has been developed may possibly be criticised for not covering all of the relevant HRQL domains related to epilepsy. For example social and activity restrictions due to

epilepsy is not covered. This is a consequence of the social activity limitation items included in the NEWQOL which cannot be used in a health state classification system because of the wording of the item and response options. Further work may investigate the possibility of using bolt-on dimensions to cover omissions and this possibility is been investigated in asthma (Brazier et al, 2010). A possible bolt on for this instrument may be the usual activities domain of the EQ-5D. It may also be possible to add a dimension by applying the techniques described here to epilepsy specific items from other instruments. This may particularly fit here as the items included in the classification system are already drawn from a variety of measures, and as such may be amenable to bolt-on dimensions.

We have completed the first stage of the development of a condition specific preference based measure, which is the generation of a health state classification system from a parent instrument using a combination of analysis techniques. The next stage of the process involves the valuation of a set of the health states generated using the preference elicitation technique Time Trade-Off (TTO) (for example see Yang et al, 2008). Health states will be valued by both the general population and by epilepsy patients, and the general population preference weights will be used in economic evaluations to complement those gained by using generic measures such as the EQ-5D. This will be initially tested by application to the SANAD dataset (Marson et al., 2005, 2007), to calculate the incremental costs per QALY and compare with the EQ-5D-based estimates. The new instrument may also be used when generic utility scores are not available and will help to address some of the concerns around using generic PBMs in epilepsy. The measure will provide a useful tool for those concerned with the allocation of resources to epilepsy interventions.

References:

Abetz L, Jacoby A, Baker GA, McNulty P. (2000). Patient-based assessments of quality of life in newly diagnosed epilepsy patients: validation of the NEWQOL. *Epilepsia, 41(9)*, 1119-28.

Aldenkamp AP, Baker GA. (1997). The Neurotoxicity Scale II**:** results of a patient-based scale assessing neurotoxicity in patients with epilepsy. *Epilepsy Research, 27*, 165-73.

Baker GA, Jacoby A**,** Francis P, Chadwick DW. (1995). The Liverpool adverse drug events profile. *Epilepsia, 36*, S59.

Birbeck GL, Hays RD, Cui XP, et al. (2002). Seizure reduction and quality of life improvements in people with epilepsy. *Epilepsia, 43(5)*, 535–538.

Bjelland I, Dahl AA, Haug TT, Neckelmann D. (2002). The validity of the Hospital Anxiety and Depression Scale – An updated literature review. *Journal of Psychosomatic Research, 52(2)*, 69-77.

Brazier J, Czoski-Murray C, Roberts J, Brown MC, Symonds S, Kelleher C. (2008). Estimation of a preference-based index from a condition specific measure: The King's Health Questionnaire. *Medical Decision Making, 28(1)*, 113-26.

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

Brazier, J.E., Rowen, D., Tsuchiya, A., Yang, Y., Young, T. (2010). What a pain: adding a generic dimension to a condition-specific preference-based measure. *Health Economics Study Group,* Cork, June 2010.

Brooks R. (1996). EuroQol: The current state of play. *Health Policy, 37(1)*, 53-72.

Dolan P. (1997). Modeling valuations for EuroQol health states. *Medical Care, 35(11)*, 1095-108.

Ettinger A, Reed M, Cramer J. (2004). Depression and comorbidity in community-based patients with epilepsy or asthma. *Neurology, 63(6)*, 1008–1014.

Ferguson E, Cox T. (1993). Exploratory factor analysis: A user's guide. *International Journal of Selection and Assessment 1(2)*, 84-94.

Jacoby A. (1992). Epilepsy and the quality of everyday life: findings from a study of people with well-controlled epilepsy. *Social Science Medicine, 34(6)*, 657–66.

Jacoby A**,** Johnson AL, Chadwick DW. (1992). Psychosocial outcomes of antiepileptic drug discontinuation. *Epilepsia, 33*, 123-31.

Jacoby, A., Baker, G.A., Smith, D.F. et al (1993) Measuring the impact of epilepsy: the development of a novel scale. *Epilepsy Research*, *16*, 83-88

Jacoby A. (1994). Felt versus enacted stigma: a concept revisited: evidence from a study of people with epilepsy in remission. *Social Science Medicine, 38*, 269-74.

Jacoby A, Baker GA, Steen N et al. (1996). The clinical course of epilepsy and it's psychosocial correlates: findings from a U.K. community study. *Epilepsia, 37*, 148-161.

Jacoby A, Baker GA. (2008). Quality of life trajectories in epilepsy: A review of the literature. *Epilepsy and Behavior, 12(4)*, 557-571.

Jacoby A, Snape D, Baker GA. (2009). Determinants of quality of life in people with epilepsy. *Neurologic Clinics, 27(4)*, 843-863.

Linacre JM. (1999). Investigating rating scale category utility. *Journal of Outcome Measurement, 3(2)*, 103-122.

Loring DW, Meador KJ, Lee GP. (2004). Determinants of quality of life in epilepsy. *Epilepsy and Behaviour, 5(6)*, 976–80.

Marson AG, Appleton R, Baker GA, Chadwick DW, Doughty J, Eaton B, Gamble C, Jacoby A, Shackley Smith DF, Smith PEM, Smith CT, Tudur-Smith C, Vanoli A, Williamson PR. (2007). A randomised controlled trial examining the longer-term outcomes of standard versus new antiepileptic drugs. The SANAD trial. *Health Technology Assessment, 11*,37.

Marson AG, Al-Kharusi AM, Alwaidh M, Appleton R, Baker GA, Chadwick DW, Cramp C, Cockerell OC, Cooper PN, Doughty J, Eaton B, Gamble C, Goulding PJ, Howell SJL, Hughes A, Jackson M, Jacoby A, Kellett M, Lawson GR, Leach JP, Nicolaides P, Roberts R, Shackley P, Shen J, Smith DF, Smith PEM, Smith CT, Tudur-Smith C, Vanoli A, Williamson PR. (2007). The SANAD study of effectiveness of carbamazepine, gabapentin, lamotrigine, oxcarbazeprine, or topiramate for treatment of partial epilepsy: an unblended randomised controlled trial. *Lancet, 369*, 1000-15.

Mensah SA, Beavis JM, Thapar AK, Keer MP. (2007). A community study of the presence of anxiety disorder in people with epilepsy. *Epilepsy and Behaviour, 11*, 118-124.

Mulhern B, Smith SC, Rowen D, Brazier J, Knapp M, Lamping DL, Loftus V, Young TA, Howard RJ, Banerjee S. (2010). Improving the measurement of QALYs in dementia: Developing patient- and carer-reported health state classification systems using Rasch analysis. *HEDS Discussion Paper 10/13.*

National Health Service (2008). Epilepsy – Introduction. http://www.nhs.uk/epilepsy. Accessed 09/08/2010.

National Institute for Health and Clinical Excellence (2008). *Guide to the Methods of Technology Appraisal.* London: National Institute for Health and Clinical Excellence.

Pearlin L, Schooler C. (1978). The structure of coping. *Journal of Health and Social Behaviour, 19*, 2-21.

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

Rasch Unidimensional Measurement Models (RUMM) 2020©. (1997-2004). RUMM Laboratory Pty Ltd.

Selai CE, Elstner K, Trimble MR. (2000). Quality of life pre and post epilepsy surgery. *Epilepsy Research, 38*, 67-74.

SPSS for windows. (2007). Release 16.0. Chicago: SPSS Inc.

Stavem K, Bjornaes H, Lossius MI. (2001). Properties of the 15D and EQ-5D utility measures in a community sample of people with epilepsy. *Epilepsy Research, 44*, 179-189.

Suurmeijer TPBM, Reuvekamp MF, Aldenkamp BP. (2001). Social functioning, psychological functioning, and quality of life in epilepsy. *Epilepsia, 42(9)*, 1160–8.

Tennant A, McKenna SP, Hagell P. (2004). Application of Rasch analysis in the development and application of quality of life instruments. *Value in Health, 7(Suppl. 1)*, S22-26.

Torrance GW, Thomas W, Sackett D. (1972). A utility maximization model for evaluation of health care programmes. *Health Services Research, 7*, 118-133.

Yang Y, Brazier J, Tsuchiya A, Coyne K. (2008). Estimating a preference based index from the Overactive Bladder questionnaire. *Value in Health, 12(1)*, 159-166.

Young T, Yang YL, Brazier JE, Tsuchiya A, Coyne K. (2009). The first stage of developing preference-based measures: constructing a health-state classification using Rasch analysis. *Quality of Life Research, 18(2)*, 253-65.

Young T, Rowen D, Norquist J, Brazier J. (2010). Developing preference-based health measures: using Rasch analysis to generate health state values. *Quality of Life Research, 19(6)*, 907-917.

Young TA, Yang Y, Brazier JE, Tsuchiya A. (2010). The use of Rasch analysis in reducing a large condition specific instrument for preference valuation: The case of moving from AQLQ to AQL-5D. *Medical Decision Making, early view.*

Zeber JE, Copeland LA, Amuan M, et al. (2007). The role of comorbid psychiatric conditions in health status in epilepsy. *Epilepsy and Behaviour, 10(4)*, 539–546.

Zigmond, A.S. & Snaith, R.P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica, 67*, 361-370

Table 1: Items included in NEWQOL subset (adapted from Abetz et al., 2000)

Scale name	Subscales and items	Scale definition
Worry about seizures scale	2 items	Measures worry about past and possible future seizures using a 4-point scale
Liverpool Adverse Events profile	19 items	Measures a range of adverse events using a 4-point frequency scale
Hospital Anxiety and Depression Scale	14 items, anxiety (7) and depression (7) subscales	Identifies clinical cases of anxiety and depression using 5-point scales
Social limitations	1 item	Assesses the extent of perceived limitation of social activities on a 4-point scale
Mastery/locus of control scale	7 items	Measures degree of internal vs. external locus of control using a 4-point scale
Stigma of epilepsy scale	3 items	Measures perceived level of stigma associated with epilepsy using a 4-point scale
Impact of seizures scale	12 items	Measures the perceived impact of seizures on a range of life areas using a 5-point scale
AB Neuropsychological Assessment Schedule	24 items, fatigue (8), memory (5), concentration (6), motor (3) and reading (2) subscales	Measures *5* aspects of cognitive function using a 4-point scale

Table 2: NEWQOL 5 factor structure

	Factor	Item	Original measure	Loading
1.	Memory and cognition	My thinking has slowed down	ABNAS (h)	0.806
		I forget all kinds of things, for example an appointment or where I put an object	ABNAS (i)	0.780
		I have difficulty concentrating on the things that I am doing.	ABNAS (j)	0.780
		I get confused and forget what I was doing	ABNAS (t)	0.771
		I forget things people have said to me	ABNAS (o)	0.763
		My mind does not work as fast as it should	ABNAS (b)	0.760
		I can't concentrate for more than a short period of time	ABNAS (p)	0.747
		I feel I react too slowly to things that are said to me	ABNAS (w)	0.745
		I have problems finding the correct word	ABNAS (f)	0.744
		I sometimes stutter or am unable to find the correct words	ABNAS (v)	0.708
		I get distracted easily	ABNAS (u)	0.707
		I have difficulties remembering the names of people	ABNAS (c)	0.674
		I feel clumsy	ABNAS (e)	0.665
		I am less capable of getting started on doing things	ABNAS (g)	0.658
		I have difficulties in following a book or film	ABNAS (d)	0.653
		I cannot keep an activity going for long	ABNAS (x)	0.651
		Memory problems	AEP (18)	0.636
		I have problems understanding what I read	ABNAS (l)	0.614
		I constantly bump against tables, doorposts, etc	ABNAS (q)	0.547
		Unsteadiness	AEP (1)	0.433
		Dizziness	AEP (15)	0.421
		I cannot use a pen or pencil accurately	ABNAS (k)	0.421
2.	Mental health	I feel cheerful	HADS (6)	0.679
		I look forward with enjoyment to things	HADS (12)	0.623
		I can sit at ease and feel relaxed	HADS (7)	0.622
		I can laugh and see the funny side of things	HADS (4)	0.618
		I still enjoy the things I used to enjoy	HADS (2)	0.599
		I feel tense or 'wound up'	HADS (1)	0.560
		Depression	AEP (17)	0.559
		I can enjoy a good book or radio or TV program	HADS (14)	0.515
		Worrying thoughts go through my mind	HADS (5)	0.503
		Feelings of aggression	AEP (4)	0.436
3.	Impact	Attacks effect the kind of paid work you can do	Impact (e)	0.744
		Attacks effect whether or not you are able to work in paid employment	Impact (d)	0.708
		Attacks effect your standard of living	Impact (j)	0.645
		Attacks effect your plans and ambitions for the future	Impact (i)	0.603
		Attacks effect your social life and social activities	Impact (c)	0.526
		Social activity restriction level	Social (1)	-0.517
		Attacks effect your the level of your independence	Impact (l)	0.494
		Attacks effect your health overall	Impact (f)	0.485
		Attacks effect the way you feel about yourself	Impact (h)	0.469
4.	Control	I have little control over things that happen to me	Control (c)	-0.654
		I often feel helpless in dealing with the problems of life	Control (e)	-0.571
		I sometimes feel that I am pushed around in my life	Control (b)	-0.557

	There is really no way I can solve some of the problems I have	Control (a)	-0.553
	There is little I can do to change many of the important things in my life	Control (g)	-0.546
5. Stigmatism	I feel some people would prefer to avoid me	Stigma (c)	0.798
	I feel some people treat me like an inferior person	Stigma (b)	0.710
	I feel some people are uncomfortable with me	Stigma (a)	0.662

Table 3:

Dimension	Item
Cognition	My mind does not work as fast as it should
	I have difficulties in following a book or film
	I have problems finding the correct word
	My thinking has slowed down
	I have difficulty concentrating on the things that I am doing.
	I have problems understanding what I read
	I can't concentrate for more than a short period of time
	I get distracted easily
	I sometimes stutter or am unable to find the correct words
	I feel I react too slowly to things that are said to me
Memory	Memory problems
	I have difficulties remembering the names of people
	I forget all kinds of things, for example an appointment or where I put an object
	I forget things people have said to me
	I get confused and forget what I was doing
Depression	Depression
	I still enjoy the things I used to enjoy
	I can laugh and see the funny side of things
	I feel cheerful
	I look forward with enjoyment to things
	I can enjoy a good book or radio or TV program
Control	There is really no way I can solve some of the problems I have
	I sometimes feel that I am pushed around in my life
	I have little control over things that happen to me
	I often feel helpless in dealing with the problems of life
	There is little I can do to change many of the important things in my life
Stigmatism	I feel some people are uncomfortable with me
	I feel some people treat me like an inferior person
	I feel some people would prefer to avoid me
Worry	How worried are you about the attacks you have had?
	How worried are you that you might have another attack?

Table 4: Rasch analysis for data half 1

Half 1 analysis		**Classical**				**Rasch**						
Factor	Item	Factor loading	% floor	% ceiling	Missing	SRM[1]	Disordered	Item range	Fit resid	Chi p value	Spread at logit 0	DIF	Poor fit
Cognition													
	My mind does not work as fast as it should	0.76	12	34	1.9	-0.01						Age	
	I have difficulties in following a book or film	0.65	7	60	2.2	-0.06		-0.39 to 1.55	2.14	0.12	0.18 to 0.60		
	I have problems finding the correct word	0.74	10	40	2.0	-0.17							Yes
	My thinking has slowed down	0.81	11	38	2.0	-0.04						Age	
	I have difficulty concentrating on the things that I am doing.	0.78	10	41	2.3	-0.01							Yes
	I have problems understanding what I read	0.61	4	68	2.5	-0.09		0.13 to 2.51	1.04	0.69	0.07 to 0.47		
	I can't concentrate for more than a short period of time	0.74	8	47	1.7	-0.09		-1.42 to 1.78	-2.33	0.02	0.14 to 0.80		
	I get distracted easily	0.71	9	43	2.0	-0.06		-1.65 to 1.61	-0.43	0.01	0.17 to 0.84		
	I sometimes stutter or am unable to find the correct words	0.71	9	46	1.6	-0.04							Yes
	I feel I react too slowly to things that are said to me	0.75	7	52	1.6	0.02		-1.05 to 2.17	-1.64	0.11	0.10 to 0.74		
Memory													
	Memory problems	0.64	22	23	1.8	-0.08		-2.37 to 0.58	1.22	0.58	0.36 to 0.91		
	I have difficulties remembering the names of people	0.67	11	44	2.0	-0.24						Age	
	I forget all kinds of things, for example an appointment	0.78	16	36	2.2	-0.10		-1.78 to 0.80	-0.95	0.01	0.31 to 0.86		
	I forget things that people have said to me	0.76	11	33	1.6	0.01							Yes
	I get confused and forget what I was doing	0.77	10	49	1.8	0.02						Age	
Depression													
	Depression	0.559	13	36	2.2	0.14		-2.04 to 0.64	1.63	0.43	0.35 to 0.88		
	I still enjoy the things I used to enjoy	0.599	9	42	0.9	0.03	Yes (Not quite so much/only a little)					Age	
	I can laugh and see the funny side of things	0.618	2	61	0.8	0.02		-0.71 to 2.63	0.41	0.05	0.07 to 0.67		
	I feel cheerful	0.679	3	48	0.7	-0.07		-1.53 to 2.20	-0.38	0.08	0.10 to 0.82		
	I look forward with enjoyment to things	0.623	4	51	1.1	0.03						Age	
	I can enjoy a good book or radio or TV program	0.515	5	60	0.7	0.00	Yes (sometimes/not often)						Yes
Control													
	There is really no way I can solve some of the problems I have	-0.553	16	17	2.5	-0.01		-1.96 to 0.74	1.96	0.46	0.32 to 0.88		
	I sometimes feel that I am pushed around in my	-0.557	7	25	2.0	-0.02						Age	

Half 1 analysis		Classical					Rasch						
Factor	Item	Factor loading	% floor	% ceiling	Missing	SRM[1]	Disordered	Item range	Fit resid	Chi p value	Spread at logit 0	DIF	Poor fit
	life												
	I have little control over things that happen to me	-0.654	10	20	2.6	-0.14		-1.86 to 1.57	-1.57	0.01	0.17 to 0.87		
	I often feel helpless in dealing with the problems of life	-0.571	10	19	2.0	-0.07						Gender	
	There is little I can do to change many of the important things	-0.546	10	19	2.2	-0.08						Age	
Stigmatism													
	I feel some people are uncomfortable with me	0.662	9	52	1.7	0.10							Yes
	I feel some people treat me like an inferior person	0.710	5	68	2.4	0.02		-1.54 to 1.38	0.93	0.44	0.20 to 0.82		
	I feel some people would prefer to avoid me	0.798	5	69	2.6	-0.03		-1.40 to 1.39	0.84	0.28	0.20 to 0.80		
Worry													
	Worried about attacks you have had		33	7	0.6	-0.85		-3.36 to 3.64	-0.12	0.55	0.03 to 0.97		
	Worried might have another attack		38	6	0.4	-0.70		-3.78 to 2.96	-0.38	0.11	0.05 to 0.98		

Table 5: Rasch analysis for data half 2

Half 2 analysis[1]		**Classical**					**Rasch**						
Factor	Item	Factor loading	% floor	% ceiling	Missing	SRM[1]	Disordered	Item range	Fit resid	Chi p value	Spread at logit	DIF	Poor fit
Cognition													
	My mind does not work as fast as it should	0.76	12	34	1.9	-0.01						Age	
	I have difficulties in following a book or film	0.65	7	60	2.2	-0.06						Age	
	I have problems finding the correct word	0.74	10	40	2.0	-0.17						Age	
	My thinking has slowed down	0.81	11	38	2.0	-0.04						Age	
	I have difficulty concentrating on the things that I am doing.	0.78	10	41	2.3	-0.01							Yes
	I have problems understanding what I read	0.61	4	68	2.5	-0.09		0.24 to 1.85	0.80	0.32	0.14 to 0.44		
	I can't concentrate for more than a short period of time	0.74	8	47	1.7	-0.09		-1.49 to 1.56	-2.46	0.03	0.17 to 0.82		
	I get distracted easily	0.71	9	43	2.0	-0.06		-1.75 to 1.39	0.01	0.47	0.20 to 0.85		
	I sometimes stutter or am unable to find the correct words	0.71	9	46	1.6	-0.04		-1.28 to 1.06	2.25	0.04	0.26 to 0.78		
	I feel I react too slowly to things that are said to me	0.75	7	52	1.6	0.02		-1.25 to 1.56	-0.97	0.19	0.17 to 0.78		
Memory													
	Memory problems	0.64	22	23	1.8	-0.08		-2.36 to 0.63	0.10	0.09	0.35 to 0.91		
	I have difficulties remembering the names of people	0.67	11	44	2.0	-0.24						Age	
	I forget all kinds of things, for example an appointment	0.78	16	36	2.2	-0.10							Yes
	I forget things that people have said to me	0.76	11	33	1.6	0.01							Yes
	I get confused and forget what I was doing	0.77	10	49	1.8	0.02		-0.64 to 1.73	0.55	0.07	0.15 to 0.65		
Depression													
	Depression	0.559	13	36	2.2	0.14		-2.20 to 0.44	2.34	0.95	0.39 to 0.90		
	I still enjoy the things I used to enjoy	0.599	9	42	0.9	0.03	Yes (not quite so much/only a little)					Age	
	I can laugh and see the funny side of things	0.618	2	61	0.8	0.02		-0.93 to 2.87	-0.44	0.04	0.05 to 0.72		
	I feel cheerful	0.679	3	48	0.7	-0.07						Age	
	I look forward with enjoyment to things	0.623	4	51	1.1	0.03						Age	
	I can enjoy a good book or radio or TV program	0.515	5	60	0.7	0.00	Yes (sometimes/not often)						Yes
Control													
	There is really no way I can solve some of the problems I have	-0.553	16	17	2.5	-0.01		-2.27 to 0.76	0.01	0.50	0.32 to 0.91		
	I sometimes feel that I am pushed around	-0.557	7	25	2.0	-0.02						Age	

Half 2 analysis[1]		Classical					Rasch							
Factor	Item	Factor loading	% floor	% ceiling	Missing	SRM[1]	Disordered	Item range	Fit resid	Chi p value	Spread at logit	DIF		Poor fit
	in my life													
	I have little control over things that happen to me	-0.654	10	20	2.6	-0.14		-2.11 to 1.35	-0.50	0.10	0.21 to 0.89			
	I often feel helpless in dealing with the problems of life	-0.571	10	19	2.0	-0.07		-2.05 to 1.49	0.01	0.22	0.18 to 0.89			
	There is little I can do to change many of the important things	-0.546	10	19	2.2	-0.08						Age		
Stigmatism														
	I feel some people are uncomfortable with me	0.662	9	52	1.7	0.10								Yes
	I feel some people treat me like an inferior person	0.710	5	68	2.4	0.02		-1.23 to 1.24	0.93	0.26	0.22 to 0.77			
	I feel some people would prefer to avoid me	0.798	5	69	2.6	-0.03		-1.25 to 0.69	0.88	0.29	0.34 to 0.78			
Worry														
	Worried about the attacks you have had		33	7	0.6	-0.85		-3.07 to 3.38	0.07	0.51	0.03 to 0.96			
	Worried have another attack		38	6	0.4	-0.70		-3.34 to 2.88	-0.01	0.28	0.05 to 0.97			

Fig 1: Flow diagram of the process used to derive a condition specific health state classification system from a non-preference based measure

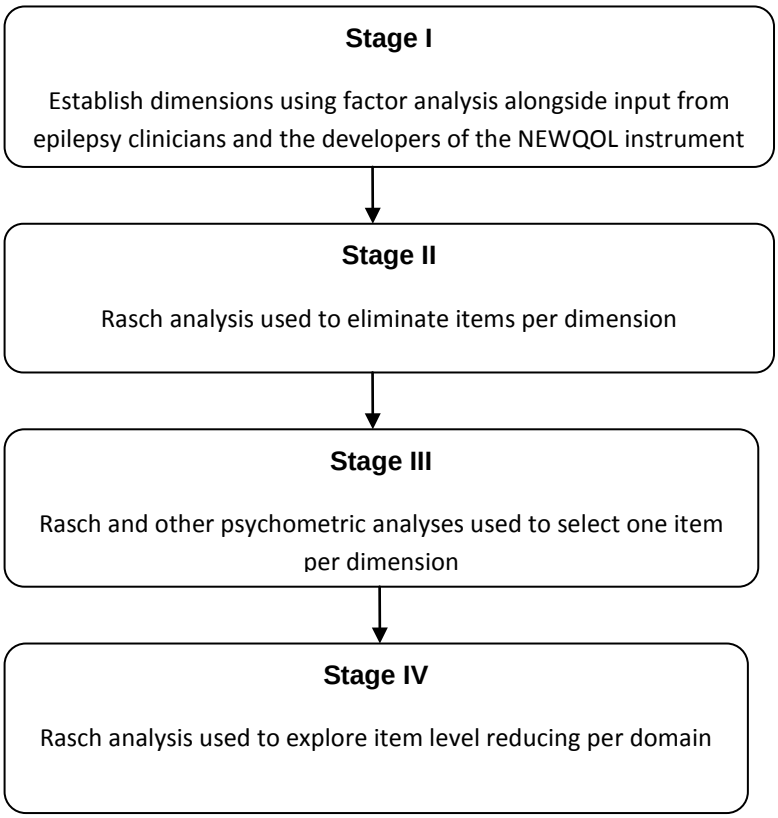

Figure 2: Final health state classification system

Worry about attacks
You are not worried at all that you might have another epileptic attack
You are a little worried that you might have another epileptic attack
You are fairly worried that you might have another epileptic attack
You are very worried that you might have another epileptic attack

Depression
You never have problems with depression
You rarely have problems with depression
You sometimes have problems with depression
You always or often have problems with depression

Memory
You never have problems with your memory
You rarely have problems with your memory
You sometimes have problems with your memory
You always or often have problems with your memory

Concentration
You have no problem concentrating for more than a short period of time
You have mild problems concentrating for more than a short period of time
You have moderate problems concentrating for more than a short period of time
You have serious problems concentrating for more than a short period of time

Control
You feel that you have complete control over things that happen to you
You feel that you have some control over things that happen to you
You feel that you have little control over things that happen to you
You feel that you have no control over things that happen to you

Stigma
You do not feel that people treat you like an inferior person
You feel that some people maybe treat you like an inferior person
You feel that some people probably treat you like an inferior person
You feel that some people definitely treat you like an inferior person