Sense and Sensibility: Selecting Health States that Make Sense using Rasch Analysis

Tracey Young*, John Brazier, Donna Rowen

Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, S1 4DA

Joint Meeting of the UK Health Economists' Study Group and the Nordic Health Economists' Study Group.

University of Aberdeen August 2009

Acknowledgements

This study is funded by Merck &Co. Inc. John Brazier is funded by the Medical Research Council Health Services Research Collaboration. The usual disclaimer applies.

* Correspondence to: Tracey Young, Health Economics and Decision Science, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA, UK. Telephone: +44114 222 0837. Fax: +44114 272 4095. Email: t.a.young@sheffield.ac.uk

ABSTRACT

AIMS:

Previous studies have used Rasch analysis to derive a preference-based measure (PBM) from an existing non-PBM by selecting a series of items that best represent the non-PBM. Once a PBM has been created, health states are traditionally selected for valuation using statistical modelling techniques, eg orthogonal, cyclical or D-optimal selection. However, these methods can result in a proportion of the selected health states being implausible, and some non-PBM do not have a clear multi-dimensional structure, where the items of such instruments are essentially different ways of tapping into the same structure (i.e. one-dimensional). A lack of independence between dimensions in a health state classification creates problems because many of the states that need to be valued for the modelling, using for example an orthogonal design, also involve combinations of dimension levels that would not be credible (eg feeling down hearted and low most of the time and happy most of the time). Typically these states can be included in health state valuations and valued alongside other plausible states. Below we propose an alternative approach using Rasch analysis to develop a plausible series of health states amenable to valuation in the development of a PBM.

METHODS:

After establishing a PBM from a non-PBM using Rasch analysis, item threshold maps, a map of where item levels occur on a latent (logit) scale, were examined. Each empirically plausible health state observed in the threshold map was recorded and these states were used in the valuation survey.

DATA:

The vignette approach is illustrated using a PBM derived from responses to a non-PBM, the Flushing Symptoms Questionnaire (FSQ). In this study the respondents who took niacin medications and experienced a common side effect of flushing were asked to complete the FSQ.

RESULTS:

The health states selected for valuation were based on the item map for the PBM derived from 1418 responders to the FSQ. The selected items for the PBM asked about symptoms of flushing related to redness, warmth, tingling, itching and sleep. The item response map identified a total of 16 possible health states which ranged from no problems for all five items to extreme problems for all five items. These sixteen states covered 79% of responses to the FSQ.

CONCLUSIONS:

The proposed approach identifies a series of states based on actual responses to items selected for the PBM. The advantage of the Rasch approach over traditional statistical methods for selecting health states for valuation is that implausible states are not included and subsequently valued. A valuation survey of the general population is currently being conducted using the sixteen states and the responses from the survey will be mapped onto the Rasch latent variable scale in order to value all possible health states.

1. INTRODUCTION

The use of preference-based measures (PBM) of health has grown considerably over the last decade with the increasing use of economic evaluation to inform health policy, such as through the establishment of bodies such as the National Institute for Clinical Excellence in England and Wales, the Health Technology Board in Scotland and similar agencies in Australia and Canada.

PBM have become a common means of generating health state values for calculating Quality Adjusted Life Years (QALYs). A widely used method to obtaining health state values is to administer one of the generic PBM of health in a clinical study [Drummond *et al*, 2005], such as the EQ-5D [Dolan, 1997], HUI3 [Feeny *et al*, 2002] or SF-6D [Brazier *et al*, 2002]. However, general measures of health have been found to be inappropriate or insensitive for some medical conditions [Brazier *et al*, 1999] and therefore many clinicians and researchers prefer condition specific measures. Most condition specific measures are not preference-based and cannot be used to derive the 'quality adjustment weight' for use in QALYs. Thus, there has been increasing interest in developing health state values from condition specific measures.

The usual method for deriving health state values from a non-PBM of health has been to apply the methods originally developed in the estimation of a PBM of health from the SF-36 [Brazier *et al*, 1998; Brazier *et al,* 2002]. This method has also successfully been applied to a number of condition specific instruments including the SQOL for sexual quality of life [Ratcliffe *et al*, 2006] and the AQLQ for asthma [Yang *et al*, 2006]. This approach involves the use of Rasch analysis to derive a health state classification system from a quality of life questionnaire that is amenable to valuation using a recognised preference elicitation technique [Young *et al*, 2007; Young *et al*, 2008]. A sample of health states defined by the classification system are valued using a valuation survey on members of the general population and subsequently a regression model is estimated to predict values for all states defined by the health state classification system.

It is accepted that it is impractical to value all possible health states from a health state classification at the valuation stage of deriving a PBM, as there would be too much information for responders to process in a valuation task. Thus, health states are traditionally selected for valuation using statistical modelling techniques, for example orthogonal design or balanced design selection. These methods are used to randomly select a series of health states and the states selected can consist of any combination of dimension levels. This selection process is applicable for multi-dimensional health state structures, where health state dimensions are independent of each other.

Using multiattribute utility theory to value health state classifications requires the valuation of corner states as part of the valuation survey [Feeny *et al,* 2002]. When valuing corner states one dimension is held at the extreme level, usually indicating extreme problems with the dimension whilst the other dimensions are held at the other extreme (no problems), this assumes independence (multi-dimensionality) between dimensions [Brazier *et al,* 2007]. However, some non-PBM do not have a clear multi-dimensional structure, where items are essentially different ways of tapping into the same concept (i.e. one-dimensional). A lack of independence between items in a non-PBM health state classification presents problems when creating a PBM because many of the health states that need to be valued for the modelling are implausible, for example feeling downhearted and low most of the time <u>and</u> happy most of the time.

An alternative approach is to construct a sample of representative health states without using a health state classification. Early approaches of deriving health states for specific conditions often involved valuing of specific vignettes or scenarios [Torrance, 1976; Sackett & Torrance, 1978; Llwellyn-Thomas *et al,* 1984; Gerrard *et al,* 1993]. This involves defining health states that represent patients with particular severity levels of health problem. Sugar and colleagues pioneered an approach using *k*-means cluster analysis to break up the data into states [Sugar *et al*, 1999]. In one study they identified patterns of the physical and mental health summary scores of the SF-12 into models with varying numbers of discrete states [Sugar *et al,* 1998]. They selected 6 states from the SF-12 data from a sample of depressed patients (i.e. near normal, mild mental and physical health impairment, severe physical health impairment, severe mental health impairment, severe mental and moderate physical impairment and severe mental and physical impairment). These were defined in terms of scores, so a process of turning the score distributions of each state into words taken from the original 12 items to define the states had to be developed based on expert judgement. This is an interesting approach, but it suffers from three limitations. Firstly, the derivation of the states uses essentially arbitrary cut-offs in the cluster analysis. Secondly, it uses dimensions scores that then need to be related to the item descriptions to generate the states and this uses expert judgement. While some expert judgement is always needed in this type of work, it should be minimized where possible. Finally, this method has only been used to value a small sample of states and it may not be possible to allocate all patients to these states (in contrast to the inclusive approach of the health state classification).

This paper proposes an alternative approach, using Rasch analysis, to develop a plausible series of health states amenable to valuation in the development of a PBM where the items included in the measure do not have a clear multi-dimensional structure. The proposed methods are illustrated using data from the flushing symptoms questionnaire (FSQ).

2. METHODS

Flushing and the FSQ

Flushing, also known as blushing, is a redness caused by increased blood flow, although it can be unpleasant, it is not harmful or dangerous. Symptoms of flushing include an uncomfortable feeling of warmth or heat, prickly (itching or tingling sensation) and reddening of the skin anywhere on the body, but most commonly to the face, neck, chest or back. People may experience one symptom or several symptoms of flushing. Flushing can occur naturally, i.e. when a person is embarrassed but is also a common and well recognised side effect of niacin medications.

The flushing symptoms questionnaire (FSQ) was constructed to "further characterise" the symptoms of flushing as a side effect of taking niacin medications [Norquist *et al,* 2007]. The questionnaire has been well validated. FSQ asks a series of 11 questions about flushing symptoms, of these seven ask respondents to rate their flushing symptoms on a 0 to 10 scale and as such are amenable for use in a PBM. The seven questions ask about overall flushing symptoms, bother of flushing symptoms, redness, warmth, tingling, itching and bother during the night. The authors of the questionnaire further suggest that responses on the ten-point scale can be categorised as: not at all bothersome (score 0), slightly bothersome (score 1 to 3), bothersome (score 4 to 6), very bothersome (score 7 to 9) and extremely bothersome (score 10) for the two bother related items and: did not have (score 0), mild (score 1 to 3), moderate (score 4 to 6), severe (score 7 to 9) and extreme (score 10) for the remaining five items.

FSQ is not currently a PBM, therefore the aim of this study was to derive a PBM based on the seven symptom relaated questions. It is recognised that the FSQ is not multidimensional and the items of the FSQ correlate with each other, therefore traditional methods for selecting health states for valuation were not applicable in the development of a PBM.

Rasch Analysis

Rasch analysis is a mathematical technique that converts qualitative (categorical) responses to a continuous (unmeasured) latent scale using a logit model and can be conceptualized as a “statistical approach to the measure of human performance, attitudes and perceptions” [Tesio, 2003]. Rasch analysis is increasingly being used in health related quality of life (HRQL) studies to create new instruments or validate existing instruments. The underlying concept with Rasch modelling is that the categorical (item) responses measure an underlying, unidimensional, latent trait, for example physical functioning, pain or general HRQL. Therefore, under the Rasch model, each item included in a HRQL domain or instrument contributes towards the underlying overall domain or instrument scale.

In previous studies we have successfully used Rasch analysis alongside more traditional psychometric analysis to aid in the construction of health state classifications [Young *et al,* 2007; Young *et al,* 2008]. These techniques will firstly be used here to validate whether the seven FSQ items unidimensionally measure flushing symptoms. The process of validating a HRQL measure using Rasch analysis is briefly summarised below.

The first stage in validating a HRQL measure using Rasch modelling is to establish where responders are able to distinguish between the item level categories. This process is achieved by fitting an initial Rasch model and examining categorical or threshold probability curves for each item included in the model (Figure 1). Figure 1 presents the categorical and threshold probability curves for an unordered item. Items that are unordered exhibit an item level curve that fails to display a unique peak i.e. the peak is masked under other item level curves, as can be seen in Figure 1 for level 1 of the item. Alternatively, the item threshold probability curve can be examined and items are unordered when item levels do not exhibit a natural order. In Figure 1 the order of levels, from left to right, is 2,1,3,4 where as, if ordered, should appear in the order 1,2,3,4). For unordered items the unordered levels should be collapsed with adjacent item levels and the Rasch model refitted.

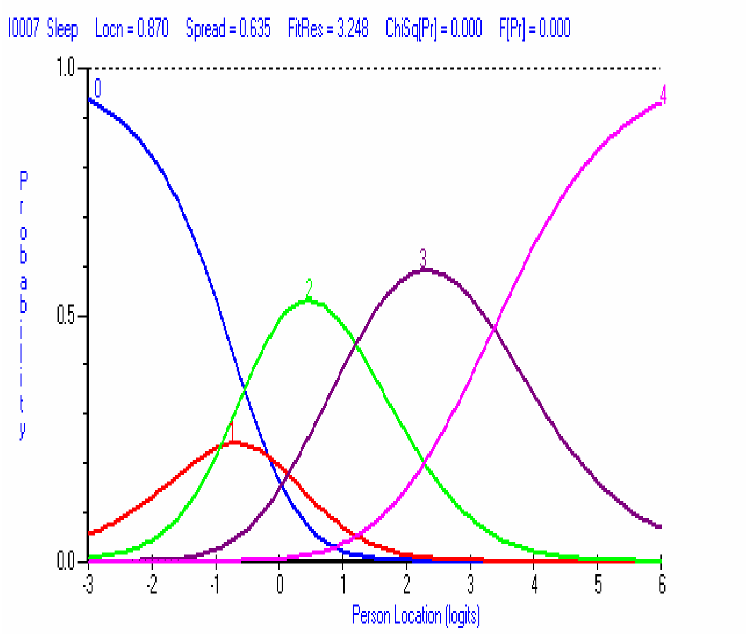
Once item level ordering has been achieved for all items then next stage is to establish unidimensionality under the Rasch model, this is achieved by examining the overall Rasch model goodness of fit statistics and individual item goodness of fit statistics. For models where unidimensionality does not hold ($p <= 0.01$) then individual item goodness of fit statistics are examined and items that fail to fit the underlying model are removed until unidimensionality is achieved. Unidimensionality is achieved when the overall Rasch model goodness of fit, measured using the chi-squared statistic, is non-significant ($p > 0.01$) indicating that all items fit the underlying Rasch model.

Using Rasch Analysis to Select Health States for Valuation

This section describes how Rasch item threshold maps (Figure 2) can be used to derive health states for the valuation stage of the development of a PBM. The item threshold map shows the points between adjacent item levels and the point where the map changes colour for each item shows the point where each item level is equally likely, the map is presented on the logit scale. This map can be used to create a set of health states across the items selected for inclusion in the PBM. For example at logit -5 respondents are most likely to be in state 00000, whereas at logit 0 they are most likely to be in state 22220. By moving from left to right across the item threshold map we can construct a set of plausible health states that are amenable for the valuation stage of creating the PBM. These states are logical and are based on the natural occurrence of states for people suffering from the underlying disease/condition e.g. flushing.

Figure 1: Example of categorical and threshold probability curves for an unordered item.

Categorical probability curve	Threshold probability curve



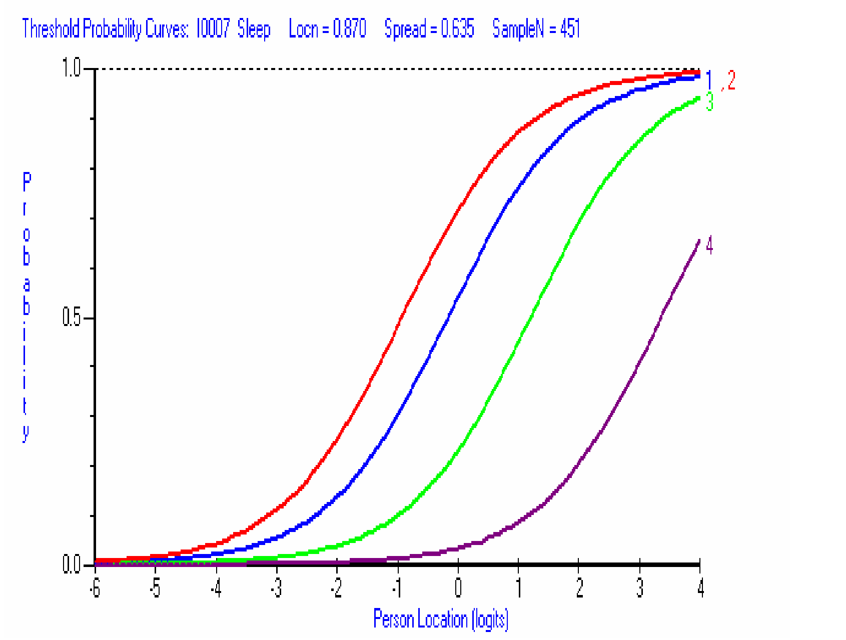
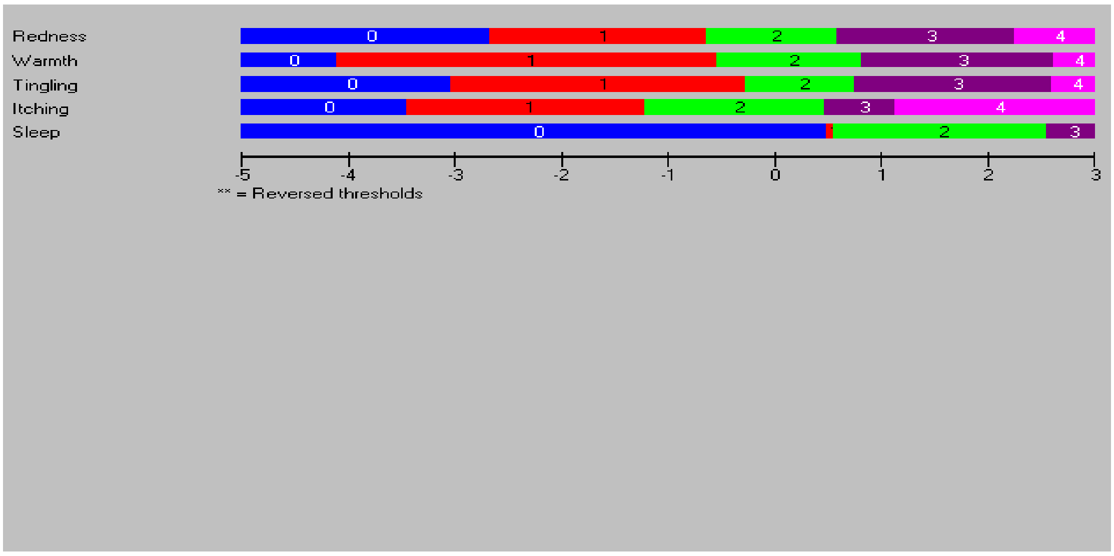


Figure 2: Item Threshold Map



For redness, warmth, tingling and itching: 0 = none, 1 = mild, 2 = moderate, 3 = severe, 4 = extreme: For sleep 0 = not at all or slightly bothersome, 1 = bothersome, 2 = very bothersome, 3 = extremely bothersome.

Alternative Methods of Selecting Health States

In order to illustrate the impracticalities of using traditional methods of selecting health states for the evaluation stage of creating a PBM the health states selected using Rasch analysis were compared with those derived using an orthogonal block design. The number of states generated by this method was set so that the number was the same as that created from the Rasch analysis.

The Flushing Datasets

Rasch models were applied to a dataset consisting of 1270 responders to the FSQ seven days after taking niacin medications to illustrate the process of selecting a series of health states amenable to valuation. Our experience of valuation surveys has shown that respondents will struggle to distinguish between a large number of item levels. Therefore, rather than using the original ten point FSQ item scale the categorised version of the scale, suggested by the authors, was used in the construction of the PBM.

Spearman's correlation and factor analysis were used to establish the level of unidimensionality between the seven FSQ items. The coverage of the selected health states created by Rasch modelling and orthogonal block design were compared using chi-squared statistics.

RUMM2020 [RUMM2020 ©, 2004] was used for all Rasch analysis, SPSS version 14.0 was used to create an orthogonal block design [SPSS Inc, 2005] and STATA version 10.0 for all other statistical analysis [STATA Corp, 2008].

3. RESULTS

Table 1 presents Spearman's correlation matrix for the seven items amenable for inclusion in a preference-based flushing symptoms measure. The table shows high levels of correlation between the seven items. Table 2 presents a similar correlation matrix to that in Table 1 but for 410 patients with osteoarthritis responding to the EQ-5D. A comparison of the results between Table 1 and Table 2 shows that the EQ-5D, a multidimensional measure, has much lower correlations between domains than the correlations between the flushing symptoms items.

Table 1: Spearman's correlation matrix for the flushing symptoms measure (N = 1270)

	Overall	Bothersome	Redness	Warmth	Tingling	Itching	Sleep
Overall	1.00	0.96	0.83	0.90	0.81	0.84	0.69
Bothersome		1.00	0.82	0.88	0.80	0.84	0.70
Redness			1.00	0.84	0.77	0.78	0.71
Warmth				1.00	0.78	0.77	0.71
Tingling					1.00	0.80	0.69
Itching						1.00	0.70
Sleep							1.00

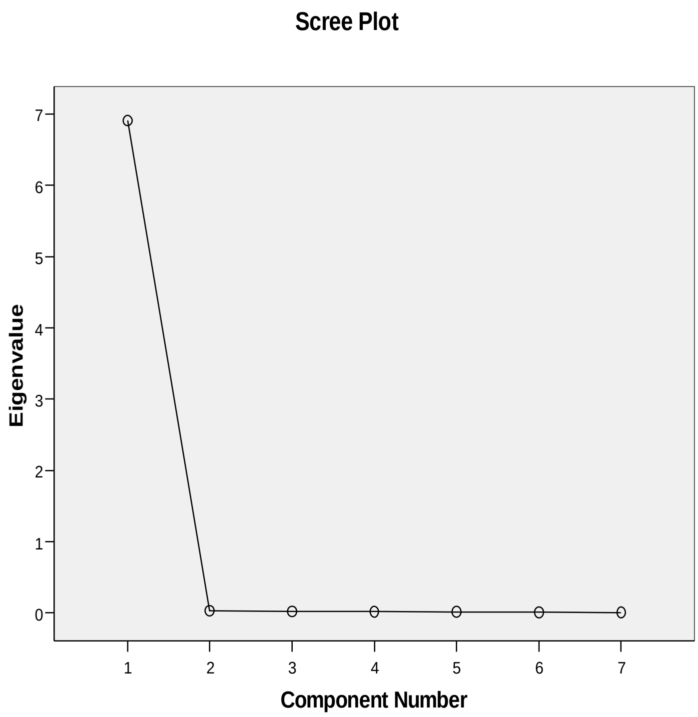
Table 2: Spearman's correlation matrix for osteoarthritis patients (N = 410)

	Mobility	Self-care	Usual activities	Pain	Anxiety and depression
Mobility	1.00	0.29	0.43	0.32	0.20
Self-care		1.00	0.38	0.33	0.35
Usual activities			1.00	0.51	0.35
Pain				1.00	0.43
Anxiety and depression					1.00

Brazier et al. (1999) dataset of osteoarthritis patients.

Figure 3 presents the results of the factor (principle component) analysis, where a sharp fall after principle component 1 strongly suggests that the seven items all belong to the same flushing symptoms dimension.

Figure 3: Results of Principle Component Analysis (Scree Plot)



Rasch Analysis

A Rasch model was fitted to the 1270 responders to the seven FSQ items and Rasch analysis was first used to validate the FSQ and a check was undertaken to verify whether respondents could distinguish between different item levels for each item (item level ordering). Item threshold curves and category probability curves were examined for each item; all items except sleep bother were ordered. Figure 1 presents the item threshold curve and the category probability curve for sleep bother, this shows an issue with item level “slightly bothersome” which could either be merged with “not at all bothersome” or “bothersome.” Rasch analysis showed that levels for “not at all bothersome” and “slightly bothersome” should be merged together.

The next stage in validating the FSQ using Rasch analysis was to eliminate items that did not fit the Rasch model (not unidimensional) this was achieve by examining the overall model goodness of fit statistic – $p > 0.01$ indicates that all items fit the underlying Rasch model.

The initial Rasch model fit for all seven items (overall flushing symptoms, bothersome, redness, warmth, tingling, itching and sleep bother) gave a significant chi-squared statistic and an overall mean item residual with a high standard deviation (mean = -1.07, standard deviation = 3.79). This suggested that one or more items could be removed from the model, as they did not fit it well. Table 3 presents the item fit statistics and residuals for the seven FSQ items, as with the overall fit a $p > 0.01$ indicates a poorly fitting item. Examination of item fit statistics suggested that four items did not fit the Rasch model: overall flushing symptoms, bothersome, tingling and itching. Further, overall symptoms and bothersome had high negative fit residuals, where residuals > 2.5 or < -2.5 are considered to be high, suggesting that these items were redundant when included with the remaining five items in the model, of the two items overall flushing symptoms had the highest residual and chi-squared statistic and was removed from the Rasch model.

Removal of the item asking about overall symptoms improved the overall goodness of fit of the Rasch model, although the model fit was still significant ($\chi^2_{42}=102.65$, $p < 0.001$, mean = -0.70, standard deviation = 2.06). The individual item goodness of fit and residuals were again examined and this time only the bothersome item failed to fit the Rasch model (Table 3). As before, this item was removed, the Rasch model was refitted and overall model goodness of fit was achieved ($\chi^2_{30}=44.64$, $p = 0.042$, mean = -0.70, standard deviation = 1.14). Therefore, the final health state classification to be developed into a PBM contained five items asking about redness, warmth, tingling, itching and sleep bother (Figure 4).

Table 3: FSQ Rasch Model Item Residual and Goodness of Fit Statistics

	Rasch Model all 7 items included			Rasch Model: Excluding overall symptoms			Rasch Model: Excluding overall symptoms and bothersome		
	Residual	c_7^2	P-value	Residual	c_7^2	P-value	Residual	c_7^2	P-value
Overall	-6.50	37.74	< 0.01						
Bothersome	-5.63	28.12	< 0.01	-4.06	38.43	< 0.01			
Redness	-0.39	16.48	0.02	-1.09	23.05	< 0.01	-2.42	17.44	0.01
Warmth	-0.10	2.32	0.94	-0.52	5.40	0.61	-0.46	3.23	0.78
Tingling	3.93	28.12	< 0.01	2.21	13.46	0.06	0.60	6.17	0.40
Itching	2.00	38.54	< 0.01	0.36	14.10	0.05	-0.01	8.98	0.17
Sleep	-0.78	5.54	0.60	-1.11	8.21	0.31	-1.27	8.83	0.18

Note: Residuals > 2.5 or < -2.5 are considered high; p > 0.01 indicates items that do not meet Rasch item fit criteria

Figure 4: Flushing Symptoms Health State Classification

Redness of skin

1. No redness of skin as a result of flushing
2. Mild redness of skin as a result of flushing
3. Moderate redness of skin as a result of flushing
4. Severe redness of skin as a result of flushing
5. Extreme redness of skin as a result of flushing

Warmth

1. No warmth as a result of flushing
2. Mild warmth as a result of flushing
3. Moderate warmth as a result of flushing
4. Severe warmth as a result of flushing
5. Extreme warmth as a result of flushing

Tingling

1. No tingling as a result of flushing
2. Mild tingling as a result of flushing
3. Moderate tingling as a result of flushing
4. Severe tingling as a result of flushing
5. Extreme tingling as a result of flushing

Itching

1. No itching as a result of flushing
2. Mild itching as a result of flushing
3. Moderate itching as a result of flushing
4. Severe itching as a result of flushing
5. Extreme itching as a result of flushing

Sleeping

1. No difficulty sleeping as a result of flushing, or difficulty sleeping as a result of flushing is not at all bothersome or slightly bothersome
2. Difficulty sleeping as a result of flushing is bothersome
3. Difficulty sleeping as a result of flushing is very bothersome
4. Difficulty sleeping as a result of flushing is extremely bothersome

Using Rasch Analysis to Select Health States for Valuation

The item threshold map was examined for the five remaining FSQ items (Figure 2), this identified a total of 16 possible health states ranging from full health (00000) to worst possible health with flushing symptoms ("pits" state 44443) (Table 4). These states cover 76% of responses to the FSQ for the 1270 responders.

Table 4: All possible logical states as shown in the Rasch item threshold map and 16 states (including best health and pits) selected assuming orthogonal (independent) design

States identified by Rasch analysis (% of sample in selected state)		States selected by orthogonal design (% of sample in selected state)	
1	00000 (65%)	1	00000 (65%)
2	01000 (3%)	2	01110 (2%)
3	01010 (1%)	3	03332 (0.0%)
4	01110 (2%)	4	10121 (0.0%)
5	11110 (4%)	5	13403 (0.0%)
6	11120 (0.5%)	6	14012 (0.0%)
7	21120 (0.1%)	7	20242 (0.0%)
8	22120 (0.0%)	8	21301 (0.0%)
9	22220 (0.7%)	9	22410 (0.0%)
10	22231 (0.0%)	10	23020 (0.0%)
11	32232 (0.1%)	11	30313 (0.0%)
12	32332 (0.1%)	12	33140 (0.0%)
13	33332 (0.2%)	13	41043 (0.0%)
14	33342 (0.0%)	14	42102 (0.0%)
15	43342 (0.0%)	15	44321 (0.0%)
16	44443 (0.1%)	16	44443 (0.1%)

Alternative Methods of Selecting Health States

Table 4 also contains 16 health states generated by orthogonal design; the orthogonally selected states cover 67% of responses to the FSQ, this is a significantly lower proportion of coverage than those selected using the Rasch method (c_1^2 = 28.66, $p < 0.001$). Only three of these states, best health (00000), 01110, and pits state (44443) overlap between both selection processes. The 13 states that are not common with the Rasch selected states do not cover any further responders to FSQ, whereas the 13 uncommon Rasch states cover a further 10% of responses. More importantly, given the high correlation between items (Table 1) a number of the states selected using the orthogonal approach are clearly infeasible. For example, taking state 5, it is unlikely that someone with mild redness would experience severe warmth and extreme tingling but no itching with extreme bother during sleep, or state 13 someone with extreme redness experiences no warmth and no tingling but extreme itching with extreme bother during sleep.

4. DISCUSSION

The health states selected for valuation were based on the item map for the PBM derived from 1270 responders to the FSQ. The selected items for the PBM asked about symptoms of flushing related to redness, warmth, tingling, itching and sleep bother. Rasch analysis, factor analysis and conventional correlation analysis all demonstrated that the selected items were unidimensional. The item response map identified a total of 16 possible health states which ranged from no problems for all five items to extreme problems for all five items. These sixteen states covered 76% of responses to the FSQ, compared to 16 orthogonally generated states covering 67% of responses, all of which were covered by the Rasch method, moreover some states were implausible for people with flushing symptoms.

This paper has shown that it is possible to generate a set of plausible health states, amenable to valuation, for a health state classification that is unidimensional. All of the states selected using the Rasch analysis item threshold map were states given by responders to the FSQ. Further, the Rasch process described here is more objective than the cluster analysis approach proposed by Sugar and colleagues as no expert judgement is required when selecting the health states.

The process described here was repeated on three further datasets from the FSQ, all of which confirmed the five items selected for the health state classification and the Rasch selected health states (results not shown).

Next stage: Valuation survey and value set

A valuation survey is currently being carried out; this will consist of interviews with 150 members of the general public. The representative sample of the general population will be asked firstly to complete the classification for their own health state for the instrument and secondly to undertake a warm-up ranking task and eight TTO valuations of health states. The Measurement and Valuation of Health (MVH) group version of TTO will be used to allow comparison with the EQ-5D tariff [Dolan, 1997]. Respondents will also be asked a number of background questions covering health, demographic and socio-economic characteristics.

A potential disadvantage with the Rasch-based approach, like clustering and orthogonal designs, is that it generates far fewer states than from using the full health state classification. We propose to solve this problem by estimating a relationship between the points on the preference scale and the latent variable produced by the Rasch FSQ model using regression techniques. The relationship between the two scales might not be linear and therefore a number of regression models will be fitted (such as logistic, cubic logarithmic) in order to establish the “best” relationship between the two scales. This would permit the estimation of preference values for other points on the latent variable and hence other states generated by the items.

The results described here may be unique to one particular condition, flushing, and one particular questionnaire, the FSQ, and therefore this process should be repeated for other HRQL instruments where the selected health states do not exhibit multidimensionality. Additional validation could be carried out using simulation exercises, creating a series of unidimensional health state classification datasets, and creating a series of health states based on the Rasch technique.

5. CONCLUSION

Condition specific measures may not always have independent items, and existing techniques of estimating health state values from condition specific measures are inappropriate when items are not independent. We have developed an alternative technique using Rasch analysis that is appropriate under these circumstances.

REFERENCES

Brazier, JE, Deverill M., Harper R., Booth A. (1999) A review of the use of Health Status measures in economic evaluation. *Health Technol Assess* 3(9)

Brazier J Harper R, Munro J, Walters S, Snaith M. Generic and condition-specific outcome measures for people with osteoarthritis of the knee. *Rheumatology 1999*; 38: 870-877

Brazier JE, Harper R, Thomas K, Jones N, Underwood T (1998) Deriving a preference based single index measure from the SF-36 *Journal of .Clinical Epidemiology* 51 (11):1115-1129

Brazier JE, Ratcliffe J, Salmon JA, Tsuchiya A (2007) Measuring and valuing health benefits for economic evaluation. Oxford University Press, Oxford

Brazier J, Roberts J, Deverill M. The estimation a preference-based single index measure for health from the SF-36. *Journal of Health Economics* 2002; 21(2):271-292

Dolan, P. (1997) Modelling valuations for EuroQol Health States. *Medical Care*, 35 (11): 1095-1108

Drummond MF, Sculpher MJ, O'Brien, BJ., Stoddart, G.L., Torrance, G.W. (2005) *Methods for the economic evaluation of health care programmes (3rd Edition).* Oxford University Press, Oxford

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

Gerard K, Dobson M, Hall K (1993) Framing and labelling effects in health descriptions: quality adjusted life years for treatment of breast cancer. Journal of Clinical Epidemiology 46: 77-84

Llewellyn-Thomas H, Sutherland HJ, Tibshirani R, Ciampi A, Till JE, Boyd NF (1984) Describing health states: methodological issues in obtaining values for health states. Medical Care 22: 543-542

Norquist J, Watson DJ, Yu QWinfen, Paolini JF, McQuarrie K, Santanello NC. (2007) Validation of a questionnaire to assess niacin-induced cutaneous flushing. Current Medical research and Opinions; 23(7):1549-1560.

Sackett DL, Torrance GW (1978) The utility of different health states as perceived by the general public. Journal of Chronic Diseases 31: 49-55

Sugar, CA, Lenert, LA, Olshen, RA, (1999) An Application of Cluster Analysis to Health Services Research: Empirically Defined Health States for Depression from the SF-12. Technical Report No. 203, Division of Biostatistics, Stanford University School of Medicine.

Sugar CA, Sturm R, Lee TT, Sherbourne CD, Olshen RA, Wells KB, Lenert LA (1998). Empirically defined health states for depression from the SF-12. *Health Services Research;* 33(1): 911-928

Ratcliffe J, Brazier JE (2006). *Estimation of a preference based single index from the sexual quality of life questionnaire (SQOL) using ordinal data*. Health Economics and Decision Science Discussion Paper 06/06, ScHARR, University of Sheffield. http://www.sheffield.ac.uk/scharr/sections/heds/discussion.html

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

SPSS Inc. Statistical Package for the Social SiencesL Release 14.0. 2005

StataCorp. Statistical Software: Release 10.0. College Station, TX: Stata Corporation. 2008

Tesio L. (2003) Measuring behaviours and perceptions: Rasch analysis as a tool for rehabilitation research. Journal of Rehabilitation Medicine 35: 105-15

Torrance GW (1976) Social preferences for health states: an empirical evaluation of three measurement techniques. Socio-Economic Planning Sciences 10(3): 129-136

Yang Y, Tsuchiya A, Brazier JE, Young TA (2006) Estimating a preference-based single index from the Asthma Quality of Life Questionnaire (AQLQ). Paper presented to the Health Economists' Study group, University of York, June 2006.

Young T, Yang Y, Brazier J, Tsuchiya A, Coyne K. (2008) Making Rasch Decisions: The use of Rasch Analysis in the Construction of Preference Based Health Related Quality of Life Instruments. HEDS Discussion Paper Series No. 08/05 http://www.shef.ac.uk/scharr/sections/heds/discussion.html

Young T, Yang Y, Brazier J, Tsuchiya A. (2007) The use of Rasch analysis as a tool in the construction of a preference based measure: the case of AQLQ. HEDS Discusion Paper Series No. 07/01 http://www.shef.ac.uk/scharr/sections/heds/discussion.html