

# Differential Item Functioning and the EQ-5D: Evidence from the UK Hospital Episode Statistics



**Dr Adam B Smith**

York Health Economics Consortium, Level 2 Market Square, University of York, Heslington, York, YO10 5NH, UK

## Background and Objectives

The EuroQol-5D (EQ-5D) is a generic or preference-based, patient-reported outcome measure (PROM) allowing comparisons to be made across different diseases and conditions. The instrument has been used in the UK's National Health Service (NHS) since 2008 to collect data from patients to assess the effectiveness of a number of surgical interventions and is used by the National Institute for Health and Care Excellence (NICE) for cost-effectiveness calculations. However, despite the widespread use of the EQ-5D it has not been subjected to rigorous scrutiny to determine whether its psychometric properties hold across different patient populations. The aim of this study was to investigate the psychometric properties of the EQ-5D using the Rasch Model.

## Methods

The data were derived from published Hospital Episode Statistics (HES) for April 2012 to March 2013. The EQ-5D had been completed by patients undergoing four surgical procedures: groin hernia repair (GH, N=21831), hip (HR, N=37800) and knee replacement (KR, N=40429) and varicose vein repair (VV, N=4681). The partial credit model<sup>1</sup> was applied to the individual datasets. Category disordering, fit statistics (infit mean square (MNSQ) < 1.30), unidimensionality (principal components analysis, eigenvalue 1st contrast < 2.0), and person separation / reliability were determined (criterion >2 and >0.8 respectively). A random sample (N=5000) was drawn from each dataset to assess differential item functioning (DIF) across the four medical conditions (criterion difference > 0.5 logits).

## Results

No category disordering was observed. Misfit was observed for the Anxiety/Depression domain for GH (MNSQ = 1.39) and VV repair (MNSQ = 1.38) groups (Table 1a and b). The eigenvalues for the 1st contrast were: 1.5 (GH), 1.4 (HR), 1.3 (KR) and 1.4 (VV) indicating there was no further dimensionality present. Person separation / reliability was 0.73 / 0.35 for GH, 1.35 / 0.63 HR, 1.06 / 0.53 KR and 0.91 / 0.45 VV suggesting that the EQ-5D is not sensitive enough to differentiate between different levels of latent trait. There was significant DIF between the 4 treatment groups: 53% (16/30) of all possible contrasts showed DIF (Table 2). The only domain not affected by DIF was Discomfort/Pain. There was DIF present in 67% (4/6) of the contrasts for Anxiety/Depression, Mobility and Self-care and in 50% (3/6 contrasts) of the Usual Activities domain.

**Table 1a: Item misfit for the EQ-5D across 4 interventions**

Intervention	Measure	Infit	Infit	Outfit	Outfit
		MNSQ	ZSTD	MNSQ	ZSTD
<b>Knee:</b>					
Mobility	0.94	0.95	-2.8	1.55	8.1
Self-care	3.45	0.94	-9.1	0.86	-5.9
Usual activities	-1.59	0.93	-7.3	0.91	-6.5
Pain/Discomfort	-4.88	0.98	-2.5	1.02	1
Anxiety/Depression	2.08	1.07	9.3	3.07	9.9
<b>Hip:</b>					
Mobility	1.23	0.93	-4	1.2	3
Self-care	2.83	0.95	-7.2	1.01	0.3
Usual activities	-1.88	0.92	-8.6	0.88	-8.3
Pain/Discomfort	-4.39	0.94	-9.6	0.99	-0.4
Anxiety/Depression	2.21	1.1	9.9	3.99	9.9

Key: MNSQ = mean square; ZSTD = standardized chi-squared

## Contact Us

adam.smith@york.ac.uk  
Telephone: +44 1904 324831  
Website: www.yhec.co.uk

 <http://tinyurl.com/yhec-facebook>

 <http://twitter.com/YHEC1>

 <http://www.minerva-network.com/>

 <http://tinyurl.com/YHEC-LinkedIn>

Providing Consultancy & Research in Health Economics

THE UNIVERSITY of York

INVESTORS IN PEOPLE



 **YHEC**  
York Health Economics Consortium

**Table 1b: Item misfit for the EQ-5D across 4 interventions**

Intervention	Measure	Infit	Infit	Outfit	Outfit
		MNSQ	ZSTD	MNSQ	ZSTD
<b>Groin:</b>					
Mobility	1.52	0.83	-9.90	0.76	-7.80
Self-care	2.16	0.87	-3.90	0.57	-5.70
Usual activities	-0.64	0.76	-9.90	0.59	-9.90
Pain/Discomfort	-3.42	1.04	2.00	1.20	7.10
Anxiety/Depression	0.38	1.39	9.90	2.17	9.90
<b>Varicose Vein:</b>					
Mobility	1.58	0.80	-8.80	0.69	-6.60
Self-care	2.78	0.87	-2.20	0.61	-3.10
Usual activities	-0.31	0.77	-9.60	0.61	-9.60
Pain/Discomfort	-3.75	1.04	1.00	1.21	3.90
Anxiety/Depression	-0.29	1.38	9.90	1.74	9.50

Key: MNSQ = mean square; ZSTD = standardized chi-squared

**Table 2: Differential item functioning of the EQ-5D by surgical intervention**

Intervention 1	Intervention 2	t	P	Domain
Groin	Hip	-2.31	-37.30	< 0.001
Groin	Knee	-2.09	-33.30	< 0.001
Groin	Varicose	0.29	4.20	< 0.001
Hip	Knee	0.23	4.72	< 0.001
Hip	Varicose	2.61	45.38	< 0.001
Knee	Varicose	2.38	40.97	< 0.001
Groin	Hip	1.58	18.85	< 0.001
Groin	Knee	2.10	24.93	< 0.001
Groin	Varicose	0.36	4.93	< 0.001
Hip	Knee	0.52	5.63	< 0.001
Hip	Varicose	-1.22	-14.90	< 0.001
Knee	Varicose	-1.74	-21.10	< 0.001
Groin	Hip	-0.23	-3.55	0.0004
Groin	Knee	0.26	4.10	< 0.001
Groin	Varicose	-0.06	-0.80	0.421
Hip	Knee	0.49	9.50	< 0.001
Hip	Varicose	0.17	2.73	0.0064
Knee	Varicose	-0.32	-5.18	< 0.001
Groin	Hip	1.13	11.43	< 0.001
Groin	Knee	0.15	1.49	0.1371
Groin	Varicose	-0.44	-3.44	0.0006
Hip	Knee	-0.98	-19.10	< 0.001
Hip	Varicose	-1.57	-16.90	< 0.001
Knee	Varicose	-0.58	-6.21	< 0.001
Groin	Hip	0.27	4.61	< 0.001
Groin	Knee	0.25	4.23	< 0.001
Groin	Varicose	-0.53	-7.96	< 0.001

## Conclusions

The most important result of this study is the finding that the EQ-5D performs differentially depending on the patient group meaning that, at least for the four patient groups investigated here, the instrument should not be used to draw comparisons across different surgical interventions. This has potentially significant ramifications for the use of the instrument as a measure of efficacy in the NHS and for cost-effectiveness by NICE.

## References

1. Masters GN. A Rasch model for partial credit scoring. *Psychometrika*, 1982; 47: 149-74.