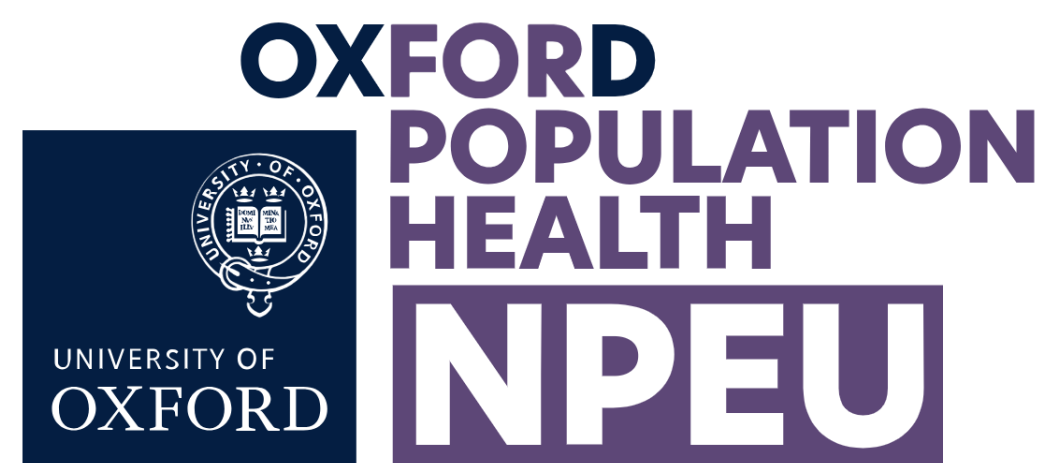


# Differential responsiveness and minimally important differences of patient-reported outcomes: truth or artefact?



Ling Jie CHENG<sup>1,2,3</sup>, Gabriel Ka Po LIU<sup>4,5</sup>, Hwee Weng Dennis HEY<sup>4,5</sup>, Nan LUO<sup>3</sup>

<sup>1</sup> National Perinatal Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom; <sup>2</sup> Alice Lee Centre for Nursing Studies, Yong Loo Lin School of Medicine, National University of Singapore, Singapore; <sup>3</sup> Saw Swee Hock School of Public Health, National University of Singapore, Singapore; <sup>4</sup> University Spine Centre, Department of Orthopaedic Surgery, National University Hospital; <sup>5</sup> Department of Orthopaedic Surgery, Yong Loo Lin School of Medicine, National University of Singapore

## BACKGROUND

- Responsiveness and minimally important difference (MID) are essential for interpreting EQ-5D scores in the evaluation of treatment effects.
- A recent review (Cheng et al., 2024) reported consistently larger anchor-based MIDs among surgical than non-surgical patients. However, it remains unclear whether this difference is due to baseline scores, treatment type, or patient characteristics.

## AIMS

- To compare the responsiveness and MID of EQ-5D and oDI in the perioperative and 1-year postoperative periods among patients followed up by a spine surgery registry.

## METHODS

- We analysed longitudinal EQ-5D-3L, ODI (a spine-specific PRO measure), and SF-36 (item 1) data collected from a cohort of 603 patients undergoing lumbar spine surgery (mean age: 57.1 years; 54.4% female) at four timepoints: preoperative (TP0), 6-month (TP1), 1 year (TP2), and 2 years postoperative (TP3).
- We assessed the responsiveness of the EQ index, VAS, and ODI index, and estimated their MIDs for the perioperative and postoperative periods using data from TP0–TP1 and TP2–TP3, respectively.
- In both analyses, SF-36 item 1 was used to define changes in health status; responsiveness was evaluated using standardised response means (SRMs), and MIDs were estimated using both the mean change and logistic regression methods.
- Subgroup analyses were conducted among patients with low baseline EQ index ( $\leq 0.5$ ), VAS ( $\leq 50$ ), or high ODI index scores ( $\geq 40$ ).

## Responsiveness

- EQ VAS demonstrated responsiveness to both improvement and deterioration in both the perioperative (SRM: 1.25 and -0.76) and postoperative periods (SRM: 0.61 and -0.98).
- While EQ index showed responsiveness in the postoperative period, it indicated improvements even among patients who reported deterioration or no change in overall health after surgery in the perioperative period (SRM: 1.15, 0.88, and 0.54; for improvement, no change, and deterioration, respectively).

## MIDs

- Mean change-derived MIDs were substantially higher in the perioperative period than the postoperative period for both EQ index (0.40 vs 0.10) and VAS (17.5 vs 6.9).
- MIDs in the perioperative period remained elevated for EQ index (0.80) and VAS (25.7) compared to MIDs in the postoperative period (EQ index: 0.04; VAS: 9.1) when only patients with low baseline scores were analysed.
- ODI results mirrored EQ-5D trends, and logistic regression-derived MIDs showed similar patterns.

## RESULTS

**Table 1** Mean change and responsiveness of the EQ-5D index, EQ VAS, and ODI scores across perioperative and postoperative periods, stratified by patients' reported improvement, no change, or deterioration in general health

	Perioperative period <sup>a</sup>				Postoperative period <sup>b</sup>			
	Improvement ( $\geq 1$ )	No change (0)	Deterioration ( $\leq -1$ )	F-statistics/ p-value	Improvement ( $\geq 1$ )	No change (0)	Deterioration ( $\leq -1$ )	F-statistics/ p-value
<b>EQ-5D Index</b>								
n	224	242	129		126	328	141	
Mean change (95% CI)	0.44 (0.39, 0.49)	0.32 (0.27, 0.36)	0.25 (0.17, 0.33)	11.5/<0.001**	0.12 (0.07, 0.18)	-0.004 (-0.03, 0.02)	-0.10 (-0.14, -0.06)	29.2/<0.001**
SRM (95% CI)	1.15 (1.04, 1.27)	0.88 (0.78, 0.98)	0.54 (0.37, 0.72)	*	0.43 (0.28, 0.57)	-0.02 (-0.13, 0.09)	-0.40 (-0.56, -0.25)	*
<b>EQ VAS</b>								
n	224	242	129		126	328	141	
Mean change (95% CI)	21.1 (18.9, 23.4)	4.1 (2.2, 6.0)	-10.5 (-12.9, -8.1)	182.1/<0.001**	9.1 (6.5, 11.7)	-2.4 (-3.9, -0.9)	-14.3 (-16.7, -11.9)	91.6/<0.001**
SRM (95% CI)	1.25 (1.08, 1.43)	0.28 (0.14, 0.41)	-0.76 (-0.94, -0.58)	**	0.61 (0.43, 0.79)	-0.18 (-0.28, -0.07)	-0.98 (-1.18, -0.78)	*
<b>ODI</b>								
n	224	242	129		126	328	141	
Mean change (95% CI)	-27.5 (-30.5, -24.6)	-21.8 (-24.8, -18.9)	-15.7 (-20.6, -10.8)	10.1/<0.001**	-6.7 (-10.3, -3.0)	-1.5 (-3.1, 0.2)	5.9 (2.9, 9.0)	18.6/<0.001**
SRM (95% CI)	-1.23 (-1.39, -1.07)	-0.94 (-1.09, -0.78)	-0.56 (-0.74, -0.38)	*	-0.32 (-0.48, -0.16)	-0.10 (-0.20, 0.01)	0.32 (0.16, 0.49)	*

SRM: Standardised response mean; n: sample size; CI: Confidence interval; \*: Baseline to 6-month; #: 1-year to 2-year; \*<0.05; \*\*<0.01; \*\*\*<0.001; Values in *italic* font denote cases where the estimate was expected to reflect deterioration or no change but instead indicates improvement.

**Table 2** Anchor-based estimates of MIDs for EQ-5D index, EQ VAS, and ODI by period, direction of change, and baseline score

	Perioperative period <sup>a</sup>				Postoperative period <sup>b</sup>			
	n	Mean change (MID <sub>Mean</sub> )	n	Logistics (MID <sub>Logistics</sub> )	n	Mean change (MID <sub>Mean</sub> )	n	Logistics (MID <sub>Logistics</sub> )
<b>MID<sub>Improvement</sub></b>								
EQ index	172	0.40	595	0.36	108	0.10	595	0.04
Low baseline score <sup>a</sup>	57	0.80			21	0.04		
High baseline score <sup>b</sup>	118	0.21			87	0.11		
EQ VAS	172	17.5	595	9.7	108	6.9	595	1.6
Low baseline score <sup>a</sup>	79	25.7			32	9.1		
High baseline score <sup>b</sup>	93	10.5			76	5.9		
ODI	105	25.6	595	23.6	108	4.5	595	2.8
Low baseline score <sup>a</sup>	113	32.1			67	4.9		
High baseline score <sup>b</sup>	59	13.0			41	4.0		
<b>MID<sub>Deterioration</sub></b>								
EQ index	105	0.28	595	0.31	119	0.09	595	0.03
Low baseline score <sup>a</sup>	29	0.81			34	0.04		
High baseline score <sup>b</sup>	76	0.08			85	0.11		
EQ VAS	172	9.1	595	0.6	119	13.9	595	6.4
Low baseline score <sup>a</sup>	10	2.0			26	14.8		
High baseline score <sup>b</sup>	95	9.8			93	13.6		
ODI	105	16.4	595	20.1	119	5.5	595	1.5
Low baseline score <sup>a</sup>	54	29.7			76	5.2		
High baseline score <sup>b</sup>	51	2.3			43	6.2		

n: sample size; \*: Baseline to 6-month; #: 1-year to 2-year; <sup>a</sup>EQ-index <0.5, EQ VAS  $\leq 50$  or ODI  $\geq 40$ ; <sup>b</sup>EQ-index  $\geq 0.5$ , EQ VAS >50 or ODI <40; Values in *italic* font denote cases where the estimate was expected to reflect deterioration but instead indicates improvement.

## CONCLUSIONS

- This study suggests that both treatment type and baseline score affect responsiveness evaluation and MID estimates when self-rated global health questions are used to define changes in health status. While the effect of baseline score could be due to ceiling effects, the treatment effect is likely driven by a change in patients' internal standards before or after surgery, manifested as response shift. There are two possible mechanisms: (1) elevated standards for good functioning before surgery, and (2) elevated standards for good overall health after surgery.
- These findings highlight the need for caution when using self-rated global health questions as anchors in assessing responsiveness or estimating MIDs of function-based HRQoL measures, such as the EQ-5D index, among surgical patients.

## REFERENCES

Cheng, L. J., Chen, L. A., Cheng, J. Y., Herdman, M., & Luo, N. (2024). Systematic review reveals that EQ-5D minimally important differences vary with treatment type and may decrease with increasing baseline score. *Journal of clinical epidemiology*, 174, 111487. <https://doi.org/10.1016/j.jclinepi.2024.111487>