REVIEW   **Open Access**

Estimating the minimally important difference for the EQ-5D-5L and EORTC QLQ-C30 in cancer

Siobhan Bourke[1*], Bryan Bennett[2†], Yemi Oluboyede[1], Tara Li[1†], Louise Longworth[1†], Sian Bissell O'Sullivan[1†], Julia Braverman[3†], Ioana-Alexandra Soare[1†] and James W. Shaw[3†]

Abstract

Background The minimal important difference (MID) is a useful tool to interpret changes in patients' health-related quality of life. This study aims to estimate MIDs for interpreting within-patient change for both components of the EQ-5D-5L questionnaire [EQ-Visual Analogue Scale (EQ-VAS) and utility index] and domains of the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 (EORTC QLQ-C30) for cancer patients.

Methods Data were obtained from the Cancer 2015 dataset, a longitudinal cohort of Australian cancer patients. Anchor-based approaches were used to estimate MIDs for the EQ-5D-5L index-based utility index [Australia and the United States (US) tariff sets], EQ-VAS scores, and the EORTC QLQ-C30. Clinical [Eastern Cooperative Oncology Group (ECOG) performance status] and patient-reported (items 29 and 30 of the EORTC QLQ-C30 and the EQ-VAS) anchors were assessed for appropriateness by their correlation strength. Clinical change groups (CCGs) were defined a priori for improvement and deterioration based on estimates used in previous literature. MIDs were estimated via linear regression and distribution-based methods.

Results For the index-based utility scores in Australia, the anchor-defined MID estimates were 0.01 to 0.06 for improvement and − 0.04 to -0.03 for deterioration, with a weighted value of 0.03 for improvement and deterioration. The EQ-VAS MID estimate was 5 points for both improvement and deterioration. For the EORTC QLQ-C30, changes of at least 3.64 (improvement) and − 4.28 (deterioration) units on the physical functioning scale, 6.31 (improvement) and − 7.11 (deterioration) units on the role functioning scale, 4.65 (improvement) and − 3.41 (deterioration) units on the emotional functioning scale, and 5.41 (improvement) and − 5.56 (deterioration) units on the social functioning scale were estimated to be meaningful.

Conclusion This study identified lower MIDs for the EQ-5D-5L utility index, EQ-VAS, and EORTC QLQ-C30 domain scores, than those reported previously. The use of a real-world cancer-specific panel dataset may reflect smaller MID

---

†Bryan Bennett, Ioana-Alexandra Soare, Louise Longworth, James Shaw, Tara Li, Sian Bissell O'Sullivan and Julia Braverman: formerly.

*Correspondence:
Siobhan Bourke
siobhan.bourke@putassoc.com

Full list of author information is available at the end of the article

Background

Health-related quality of life (HRQoL) is an important complement to other clinical endpoints for people living with cancer as it can provide a means of capturing personal and social contexts of the disease experience. It extends beyond the absence of disease and is the subjective analysis of the impact of physical and emotional wellbeing on patients' quality of life [1]. HRQoL can be used to understand the patient experience, particularly the impact of treatment on patients' functioning. In addition, it can help to identify from patients' perspective whether they experience a minimal or clinically meaningful change in their wellbeing [2]. Two commonly used HRQoL measures in cancer patients are the EQ-5D-5L, a well-established generic preference-based instrument in health outcomes research, and the European Organisation for Research and Treatment of Cancer Quality of life Questionnaire Core 30 (EORTC QLQ-C30), a cancer-specific health profile. The EQ-5D has been increasingly used as a clinical outcome assessment (COA) tool to complement clinical measures in regulatory and drug technology submissions to understand patient outcomes and needs [3, 4].

Submission of HRQoL data is common for reimbursement submissions especially using the EQ-5D, which is the preferred measure for many reimbursement agencies. For regulatory approvals, the inclusion of HRQoL measures is not mandatory for submissions to the US Food and Drug Administration (FDA) [2], European Medicines Agency (EMA), or Australian Therapeutic Goods Administration (TGA) [5]. A review of health technology assessment (HTA) and regulatory submissions found that 14% of HTA and 58% of regulatory submissions used the EQ-5D measure as a COA with 37% of HTA and 16% of regulatory submissions also referencing minimal important difference (MID) values for the instrument [6]. Using the EQ-5D in this way requires the estimation of robust MID values for the interpretation of the magnitude of change similar to the interpretation of other COA measures.

Although most reimbursement agencies use the relative effectiveness approach for interpreting HRQoL, regulatory agencies prefer interpreting HRQoL scores that reflect a clinically relevant change. This value goes beyond statistical significance because it detects important outcomes in patients and indicates whether individual patients have experienced meaningful clinical benefit from the treatment [7].

There is currently a lack of consensus between regulatory agencies regarding the use of terminology for clinically relevant changes; the FDA refers to these changes as a meaningful within-patient change [8], whereas the EMA and TGA refer to these changes as the MID [5]. The EMA definition focuses on the minimal amount of change; however, the FDA suggests that a minimal change does not necessarily imply that the change is meaningful to patients, and thus, their definition of meaningful change avoids using this terminology. Another important difference that the FDA highlights is that the MID terminology can be used interchangeably to indicate either 'group-level' mean differences or 'individual-level within patient change' differences. The FDA distinguishes these groups by stating that a treatment difference, as calculated in the 'between group difference' accounts for the difference between two trial arms, is separate to an 'individual within-patient change' where the analysis evaluates whether a meaningful score change is observed. Hence, the FDA's preference being individual-level responses over group level [2].

Beyond regulators, authors have highlighted that there is a difference in terminology between values generated by anchor based and distribution-based methods. Distribution based methods focus on the "smallest change that can be detected by the instrument beyond measurement error" or a minimal detectable change (MDC). Where anchor-based methods estimate "the smallest difference in the score of a patient-reported outcome (PRO) measure or clinical outcome that is meaningful to a patient" i.e. a MID [9].

Currently, the EORTC is undertaking a project to establish MIDs for all QLQ-C30 scales according to cancer sites, using individual patient data from archived EORTC trials [10]. For the EQ-5D-5L, MID estimates currently exist for the utility index in 7 disease areas [11–13], but not in cancer. For cancer, the only MID value available is for the EQ-5D-3 L questionnaire estimated from a cross-sectional dataset [14]. Previous studies estimating the MID for the EQ-VAS reported a value of between 7 and 10 points on the 0 to100 scale.

Methods

In this study, we calculated the individual within-patient change; however, given the adoption of the EMA definition of MID by the TGA, we used this MID definition, to ensure consistency between the regulator's definition and the data source in Australia [5]. This study used a

longitudinal population-based cohort study to estimate the individual within-patient MID and the instrument MDC scores for the EQ-5D-5L and EORTC QLQ-C30. To our best knowledge, this study is the first to harness the use of observational data for the estimation of MID in cancer, providing a real world MID for the EORTC QLQ-C30 and EQ-5D measures beyond those captured in a clinical trial study.

Study design and population

A retrospective analysis of the Cancer 2015 dataset was performed. Cancer 2015 was a longitudinal population-based cohort study [15] that collected data in patients with a new diagnosis of cancer for over 20 different tumour types at all stages of the disease. Information on patient sociodemographic, patient and family history, patients' HRQoL, and their EQ-5D and EORTC QLQ-C30 measures was collected. These measures were repeated at 6- and 12-month follow-ups or for those with advanced disease, and the first follow-up data collection was at 3 months. For the first 4 years, data on the EQ-5D-3 L were collected; however, following the development of the EQ-5D-5L [16] in late 2015, data were collected using the 5 L version, for newly enrolled participants and for those undergoing follow-ups.

Inclusion/exclusion criteria

To ensure the estimation of robust MIDs, the study required complete data for the selected anchors and the 2 outcome measures, EQ-5D-5L and EORTC QLQ-C30. Therefore, data was requested from the Cancer 2015 data custodian for patients who had complete/non-missing EQ-5D-5L and EORTC QLQ C30 observations for at least 3 data timepoints – with a maximum of 6 timepoints across the analysis period including baseline. A total of 799 patients in the Cancer2015 database met these criteria for either EQ-5D-5L or EORTC QLQ-C30. Given the switch from EQ-5D-3 L to 5 L version in the data base, we created two analytical samples. For the EORTC QLQ-C30 the full data set of 799 patients were used for the analysis, and a subset of 464 patients were used for the analysis of the EQ-5D-5L. This data was pooled across all timepoints and combined into one dataset, to provide sufficient a sufficient sample size for estimating the MID/ MCD for both the EQ-5D-5L components and EORTC QLQ-C30 domains.

Measures

EQ-5D-5L

The EQ-5D-5L is a generic measure of HRQoL. The first part includes 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression), each being characterised into 5 levels of severity (no problems, some problems, moderate problems, severe problems, and extreme problems/unable to that creates a health profile. The second part is a visual analogue scale (EQ-VAS) which asks respondents to rate their health today on a scale ranging from 0 (worst imaginable health) to 100 (best imaginable health). Each health profile described by the EQ-5D-5L can be linked to a utility index value anchored on a scale, where 1 represents full health and 0 is equivalent to being dead, with negative health states representing health states worse than being dead [17]. The Australian value set was used to value the EQ-5D-5L utility index [18] in the base-case analysis, and the US utility index [19] was applied in the sensitivity analysis (Supplementary material).

EORTC QLQ-C30

The EORTC QLQ-C30 is a cancer-specific HRQoL questionnaire that contains 30 items evaluating 5 functional domains (physical, emotional, social, role, and cognitive), 8 symptoms (fatigue, pain, nausea/vomiting, constipation, diarrhoea, insomnia, dyspnoea, and appetite loss), global health status and quality of life, and the financial impact of cancer. All items have a 4-point scale (not at all, a little, quite a bit, and very much), except for the global health status and overall QoL scales. Each of the 13 dimensions are scored on a 0-100 scale [20]. To ensure consistency in the interpretation of MID/ MCD scores, we deviated from the standard scoring procedure by scoring all scales such that 0 represents the worst possible score and 100 the best possible score.

Statistical analysis

Anchor and clinical change groups

Anchor-based approaches use an external indicator (or anchor) to assign patients into different clinical change groups (CCGs) determined prior to analysis [21].A review of anchors commonly used in MID studies was completed prior to requesting the data. This was done to ensure that the Cancer 2015 dataset had suitable anchors that could be used to estimate MID values. From the review we identified the candidate anchors and cut off points in Table 1 which further describes the anchors and the CCG cutoffs. The CCGs for each anchor reflected levels of change, defined as small positive change, small negative change, and no change. Change scores either lower or higher than the pre-defined CCG cut offs, within-patient thresholds (shown in column 3 of Table 1) were excluded from the analysis, as these were not considered to have experienced a 'minimally important' change. For example, they either experienced a change that was too low to be considered important or too high to be considered 'minimal'. Four external indicators were chosen to be included in the analysis, one clinical [Eastern Cooperative Oncology Group (ECOG)] and 3 patient-reported (EQ-VAS as well as items Q29 and Q30

Table 1 Anchor and clinical change group descriptions

Anchor	Description	Clinical change group
ECOG	The ECOG is a clinical measure with 5 grades of functioning that range from 0 to 5, where 0 indicates that a patient is *fully active*, 4 suggests that a patient is *completely disabled*, and 5 represents *death*.	A positive/negative change in the score of 1 grade on the scale of 0–5 represented the *minimal expected change* [22].
EQ-VAS	Patient self-reported general health question on a scale ranging from 0, which indicates *the worst health you can imagine*, to 100, which represents *the best health you can imagine*.	A positive/negative change in the score in the range of 7–10 points on the VAS scale of 0-100 [14].
EORTC QLQ-C30 item 29	Item 29 of the QLQ-C30 questionnaire asks respondents to self-report their overall health during the past week on a scale that ranges from 1, which indicates *very poor*, to 7, which represents *excellent*.	A positive/negative change in the score of 1 point on the scale of 1–7 [23]
EORTC QLQ-C30 item 29	Item 30 of the QLQ-C30 questionnaire asks respondents to self-report their quality of life during the past week on a scale that ranges from 1, which indicates *very poor*, to 7, which represents *excellent*.	A positive/negative change in the score of 1 point on the scale of 1–7 [23]

ECOG, Eastern Cooperative Oncology Group; EORTC QLQ-C30, European Organisation for Research and Treatment of Cancer Quality of life Questionnaire Core 30; EQ-VAS, EuroQol Visual Analogue Scale

of the EORTC QLQ-C30). The EQ-VAS was not used to facilitate anchor-based analyses for itself.

The anchor and outcome measures should measure the same or similar underlying constructs, and therefore, should be appreciably correlated; thus, there should be an empirical association between each anchor and the EQ-5D-5L (utility index score and EQ-VAS) and each of the EORTC QLQ-C30 domains. To be interpretable for the estimation of MID, each anchor must correlate at least moderately (>|0.30|) with the HRQoL outcome measure [21]. The financial EORTC QLQ-C30 scale item was excluded from the analysis as it was presumed to be weakly correlated with the anchors [24]. Thus, no MID was estimated for this scale.

The regression analysis method quantified MID as the coefficient of the CCG group indicator obtained from fitting 2 regression models, one estimating improvement and one estimating deterioration with the change in HRQoL index score as the dependent variable, and the anchor as an explanatory variable [24, 25]. Hence, they compare the improvement group with the no-change group, and the deterioration group with the no-change group, respectively. The improvement and deterioration groups (anchors) are determined based on small positive and negative changes, respectively, in overall health score between consecutive time points for each patient. Note

this is an unbalanced panel dataset with some patients having more data points (and hence, change points) than others.

The analysis utilises a fixed-effects regression model, controlling for unobserved heterogeneity by allowing each patient to have its own intercept. This includes an adjustment by clustering on patients to provide robust standard errors that account for potential heteroskedasticity and autocorrelation within clusters.

Potential confounding factors, such as patient demographic and clinical characteristics (e.g., age, sex, primary cancer site, and cancer stage), were included in the regression models. A weighted MID anchor value was calculated based on the strength of the correlation between the anchor and the outcome measure [26].

Distribution- based methods
Complementary to the main study MID results, a distributional based MCD was estimated. Two proportions of standard deviation (0.3, 0.5) were used as well as the standard error of measurement (SEM). SEM is calculated as $SD * \sqrt{1-r}$, where r is the reliability coefficient (EORTC QLQ-C30 0.85 [27], EQ-5D-5L index 0.85 [28]) for the HRQoL instrument.

The effect size (ES) was calculated by dividing the MID/MCD estimate by the SD of the overall HRQoL score. For interpretation, a valid MID/MCD represented an ES ranging between ≥0.2 and <0.5 based on Cohen's recommendation [29].

Results
Table 2 presents a summary of the demographic and clinical characteristics of patients in each of the EQ-5D-5L and EORTC QLQ-C30 samples. In both analytic samples, the mean age of participants at baseline was 63 years. There were 56% and 55% female participants in the EQ-5D-5L and EORTC QLQ-C30 samples, respectively. The most common cancer type reported was breast cancer (35%) followed by genitourinary cancer [EORTC QLQ-C30 (22%) and EQ-5D-5L (23%)]. The EORTC QLQ-C30 sample had more respondents who had distant metastases (6%) in comparison with the EQ-5D-5L sample (3%) Baseline HRQoL is detailed in the supplementary material.

A total of 4 potential anchors were initially assessed for both the EQ-5D-5L (utility index score and EQ-VAS) and EORTC QLQ-C30 domain scales. Table 3 provides estimates of the correlation between the EQ-5D-5L outcomes and each anchor. The correlation between the HRQoL scales and anchors ranged from −0.09 to 0.78, it was determined that the EQ-5D-5L utility index scores were at least moderately correlated (|r| ≥ 0.30, highlighted in bold) with the selected patient-reported health anchors. However, the correlation between the EQ-VAS

Table 2 EQ-5D-5L and EORTC QLQ-C30 sample demographic and clinical characteristics

Characteristics	EQ-5D-5L analytic sample (*n* = 464)		EORTC QLQ-C30 analytic sample (*N* = 799)	
	Mean	SD	Mean	SD
Age	63	11	63	12
	n	% of total N of patients	n	% of total N of patients
Sex				
Female	260	56%	442	55%
Number of follow-ups				
1	27	6%	45	6%
2	70	15%	154	19%
3	117	25%	219	27%
4	177	38%	206	26%
5+	73	16%	170	22%
Cancer site at first diagnosis				
Breast	161	35%	278	35%
Genitourinary	106	23%	178	22%
Head and neck	63	14%	95	12%
Colorectal	60	13%	101	13%
Lung	25	5%	45	6%
All others	46	10%	99	12%
Missing	3	0.7%	3	0.38%
Cancer stage				
Unknown	8	2%	16	2%
Distant metastases	12	3%	50	6%
Regional lymph nodes	108	23%	188	24%
Invasion of adjacent tissue or organs	44	9%	70	9%
Localised to the tissue of origin	262	56%	437	55%
Missing	30	6%	38	5%

EORTC QLQ-C30, European Organisation for Research and Treatment of Cancer Quality of life Questionnaire Core 30; EQ-5D-5L, EuroQol 5-Dimension 5-Level Questionnaire; SD, standard deviation

Table 3 Correlations between anchors and the EQ-5D-5L utility score and EQ-VAS score

Anchor	EQ-5D-5L utility score	EQ-VAS score
	Correlation	Correlation
ECOG	-0.10	-0.09
EQ-VAS	**0.60**	n/a
Q29 (EORTC QLQ-C30 overall health)	**0.57**	**0.78**
Q30 (EORTC QLQ-C30 quality of life)	**0.57**	**0.73**

ECOG, Eastern Cooperative Oncology Group; EORTC QLQ-C30, European Organisation for Research and Treatment of Cancer Quality of life Questionnaire Core 30; EQ-VAS, EuroQol Visual Analogue Scale

Numbers highlighted in **bold** meet the threshold (r >| 0.30|) for inclusion for the minimal important difference estimation

Table 4 Base-case correlations between EORTC QLQ-C30 and external anchors

Scale	ECOG	EQ-VAS	Q29 (Overall health)	Q30 (QLQ-C30 quality of life item)
PF	**-0.48**	**0.61**	**0.6**	**0.58**
RF	**-0.41**	**0.64**	**0.63**	**0.63**
EF	-0.17	**0.49**	**0.49**	**0.52**
CF	-0.24	**0.47**	**0.46**	**0.46**
SF	**-0.35**	**0.58**	**0.57**	**0.62**
FA	**-0.38**	**0.66**	**0.67**	**0.64**
NV	-0.24	**0.36**	**0.37**	**0.37**
PA	-0.25	**0.56**	**0.58**	**0.55**
DY	-0.25	**0.43**	**0.44**	**0.44**
SL	-0.17	**0.4**	**0.42**	**0.42**
AP	**-0.30**	**0.43**	**0.44**	**0.43**
CO	-0.17	0.27	0.27	0.25
DI	-0.12	0.2	0.23	0.22
QL	**-0.34**	-0.11	n/a	n/a

AP, appetite loss; CF, cognitive functioning; CO, constipation; DI, diarrhoea; DY, dyspnoea; ECOG, Eastern Cooperative Oncology Group; EF, emotional functioning; EORTC QLQ-C30, European Organisation for Research and Treatment of Cancer Quality of life Questionnaire Core 30; EQ-5D-5LEQ-VAS, EuroQol Visual Analogue Scale; FA, fatigue; NV, nausea/vomiting; PA, pain; PF, physical function; QL, global health status/quality of life; RF, role functioning; SF, social functioning; SL, insomnia

Bold figures satisfy the threshold (|r| > 0.30)

HRQoL scales and anchors ranged from 0.11 [EQ-VAS/ global health status/quality of life (QL)] to 0.67 [EORTC QLQ-C30 item Q29/Fatigue (FA)] in absolute value. Generally, for both measures, patient-reported anchors showed greater correlations with the HRQoL items when compared with the clinical anchor.

Table 5 presents the MID estimates for the anchors that met the responder threshold criteria for the EQ-5D-5L utility index score. The valid MID estimate for the EQ-5D-5L utility index for improvement ranged from 0.01 (anchor used Q30 QLQ-C30, estimated using the linear regression model) to 0.06 (anchor used EQ-VAS, estimated with and without baseline HRQoL). Table 5 also presents a single average MID weighted summary based on the correlation values. This resulted in an MID of 0.03 for both the improvement and deterioration using the regression model.

Table 6 shows the results of the EQ-VAS MID analysis. The EQ-VAS MID estimate of improvement was between 4 (anchor Q30, estimated using the regression model accounting for baseline HRQoL) and 6 (anchor Q29, estimated using the regression model with and without baseline). The deterioration ranged from −4.00 (estimated using anchor Q30 and regression with and without baseline) to -5.00 (estimated using anchor Q29 and regression with and without baseline). Table 7 also provides the MCD estimates for the distribution approach, with the

and the index scores and the clinical anchor (ECOG performance status) were poor.

Table 4 provides the correlations between each EORTC QLQ-C30 scales and anchors. The correlation between

Table 5 EQ-5D-5L utility index MID using linear regression models and distribution -based MCD

Anchors	Linear Regression Model		Linear Regression Model (baseline EQ-5D-5L utility index score)		Distribution-based Approach (MCD)		
	Improve	Deteriorate	Improve	Deteriorate	0.5	0.3	SEM
EQ-VAS	0.06	-0.04[a]	0.05[a, b]	-0.04[a]	0.07	0.04	0.06
	(0.41)	**(-0.27)**	**(0.37)**	**(-0.26)**			
Q29 (QLQ-C30 overall health)	0.03[a]	-0.03[a]	0.03[a]	-0.03[a]			
	(0.23)	**(-0.21)**	**(0.23)**	**(-0.21)**			
Q30 (QLQ-C30 quality of life)	0.01	-0.04[a]	0.01	-0.03[a, b]			
	(0.09)	**(-0.25)**	**(0.08)**	**(-0.25)**			
Weighted EQ-5D-5L MID	0.03	-0.03	0.03	-0.03			
	(0.25)	**(-0.25)**	**(0.23)**	**(-0.24)**			

EQ-VAS, EuroQol Visual Analogue Scale; MID, minimally important difference; MCD, minimal detectable change; **Bold** figures reflect effect size (ES)

[a]MID change vs. no change is statistically significant at 0.01

[b]Baseline EQ-5D-5L utility significant

Table 6 EQ-VAS MID using linear regression models and distribution -based MCD

Anchors	Linear Regression Model		Linear Regression Model (baseline EQ-5D-5L utility)		Distribution-based Approach (MCD)		
	Improve	Deteriorate	Improve	Deteriorate	0.5	0.3	SEM
Q29 (QLQ-C30 overall health)	6.00	-5.00	6.00[a]	-5.00[a]	8.00	5.00	7.00
	(0.40)	**(0.30)**	**(0.40)**	**(-0.30)**			
Q30 (QLQ-C30 quality of life)	5.00[a]	-4.00[a]	4.00[a]	-4.00			
	(0.3)	**(0.3)**	**(0.30)**	**(0.20)**			
Weighted EQ-VAS MID	5.00	-5.00	5.00	-5.00			
	(0.30)	**(0.30)**	**(0.30)**	**(-0.30)**			

EQ-VAS, EuroQol Visual Analogue Scale; MID, minimally important difference; MCD, minimal detectable change; QLQ-C30, Quality of life Questionnaire Core 30

Bold figures reflect effect size (ES)

[a]MID change vs. no change is statistically significant at 0.01

results of the MID analysis ranging from 5.00 (estimated using 1/3 SD) to 8.00 (1/2 SD).

Table 7 presents the MID anchor estimates for the EORTC QLQ-C30 subscales. The MID estimate signs were consistent with expectation, i.e., positive for improvement and negative for deterioration. The results of the range of MID estimated using the different anchors and weighted MID values are presented as bold figures in Table 7.

Discussion

This study provides estimated MID values for the EQ-5D-5L utility index, EQ-VAS and EORTC QLQ-C30 for a population of people with cancer. Estimated MID values were consistent with the expected signs for all instruments, whereby deterioration in health had negative values and improvements in health positive values. For the index-based utility scores, the anchor-defined MID estimates were 0.01 to 0.03 for improvement and −0.04 to -0.03 for deterioration, with a weighted value of 0.03 for both improvement and deterioration. The MID for the EQ-VAS was 5 points for both improvement and deterioration. MID values were estimated for the EORTC QLQ-C30 functioning subscales and ranged from 2.44

to 6.31 for improvements, and from −3.41 to -7.11 for deterioration.

Our estimated MID anchor values were generally smaller than estimates previously reported for populations of people with cancer. For example, for the EORTC-QLQ-C30 we found similar MID values to the study by Cocks et al., MID estimates in this study for 4 of the scales (role functioning, social functioning, pain, and insomnia) were within the range of 5 to 8-unit change. However, estimates for 5 of the scales (cognitive functioning, nausea vomiting, dyspnoea, and appetite loss) were lower than the reported range by Cocks et al. [30]. Published estimates for the EQ-5D-5L in people with cancer are not available with which to compare. However, previously estimated MIDs in cancer patients for the EQ-5D-3 L in the UK and US are higher than our estimates (0.08 for UK-index scores, 0.06 for US-index scores, [14]). Published EQ-5D-5L MID estimates in chronic conditions are also lower than reported here (0.044 in Canadian population with chronic conditions [31], and 0.07 in a Chinese population with hypertension and diabetes [32]).

MID values can vary by disease, severity, patient baseline status, direction of change, demographic factors and sample characteristics [22, 30]. Our study used

Table 7 Base-case EORTC QLQ-C30 MID estimates

	Linear Regression model		Distribution-based approach (MCD)		
Scale	Improve	Deteriorate	0.3	0.5	SEM
PF[a]	2.75 to 6.37 **(3.64)**	-6.58 to -3.01 **(-4.28)**	5.37 4	7.59	7.59
RF[a]	5.39 to 9.25 **(6.31)**	-8.12 to -6.43 **(-7.11)**	13.63	8.2	11.57
EF	4.39 to 5.08 **(4.65)**	-4.96 to -1.96 **(-3.41)**	10.72	6.43	9.09
CF	2.18 to 2.68 **(2.44)**[b]	-3.13 to -2.90 **(-3.03)**	10.11	6.06	8.58
SF	4.98 to 5.81 **(5.41)**	-7.04 to -4.12 **(-5.56)**	12.93	7.77	10.97
FA[a]	4.95 to 7.23 **(5.76)**	-8.27 to -6.47 **(-6.93)**	11.48	6.89	9.76
NV	1.16 to 2.71 **(1.86)**	-2.26 to -2.12 **(-2.19)**	5.86	3.52	4.98
PA	3.39 to 5.37 **(4.31)**	-7.36 to -5.48 **(-6.54)**	12	7.2	10.18
DY	0.25 to 2.11 **(1.39)**[b]	-4.48 to -3.29 **(-4.08)**	10.49	6.29	8.9
SL	4.43 to 7.01 **(5.63)**	-6.18 to -5.05 **(-5.70)**	14.64	8.78	12.42
AP	3.00 to 4.85 **(3.84)**	-4.87 to -2.32 **(-3.42)**	11.71	7.03	9.94

AP, appetite loss; CF, cognitive functioning; DY, dyspnoea; ECOG, Eastern Cooperative Oncology Group; EF, emotional functioning; EORTC QLQ-C30, European Organisation for Research and Treatment of Cancer Quality of life Questionnaire Core 30; EQ-VAS, EuroQol Visual Analogue Scale; FA, fatigue; MID, minimally important difference; MCD, minimal detectable change; NV, nausea/vomiting; PA, pain; PF, physical function; QL, global health status/quality of life; RF, role functioning; SF, social functioning; SL, insomnia

Bold figures reflect EORTC QLQ-C30 subscale weighted MID estimates

aECOG anchor available for estimation for these subscales

bFailed to meet the Cohen's criteria, a true MID estimate effect size is required to be between ≥0.2 and ≤0.8

real-world data rather than data from clinical trials which may explain some of the differences between our estimates of MID and the published literature. First, given the aim of trials is to identify HRQoL differences between treatment and control group the impact of the treatment exposure may lead to greater HRQoL changes compared to those reported in a observational sample [33]. In addition, observational data may include the HRQoL values that are infrequently identified in trials given that clinical trials often aim to recruit a homogenous patient group. As such, the results from this study can be seen as conservative estimates of the MID compared to those found in clinical trial settings. Future research could determine whether observational cohorts may require different MID thresholds for different subsets of populations. For example, separate estimates would be required for those categorised as stable versus progressive disease to ensure heterogeneity. As this study utilised a cancer-specific longitudinal dataset in the estimation of MID for both HRQoL measures this approach may be more applicable to the real-world cancer population and provide better estimates for reimbursement decision-making agencies. Consequently, more patients with cancer who experience small changes in the EQ-5D-5L utility index or VAS score

would be categorised as having a meaningful improvement using our estimation of MID than if using MID estimates previously estimated using clinical trial data.

It should also be noted that differences in the MIDs are reflected from different anchors that reflect the specific contexts of the health aspects being measured. As such, the MIDs may be sensitive to the bluntness or sensitivity of the chosen anchors. The ECOG performance status, which captures broader physical functioning changes, may produce different MID estimates compared to the more subjective measures of quality of life in QLQ-C30 items. In addition, a consideration must be made for the magnitude of change in ECOG performance status where a change between each level is substantial enough to demonstrate clear, objective improvement. While this can offer greater clarity compared to the subjective nature of the magnitude change in PRO instrument anchor, it may not fully capture the nuanced personal perspective of the patient with cancer, beyond the physical health that a PRO measure can.

Another reason for the smaller estimates may be due to the high baseline HRQoL values in our sample making it more difficult to detect improvement but easier to detect the deterioration. We can hypothesise that

because of these high HRQoL levels at baseline, improvement occurs at a lower magnitude in our sample, which may explain our lower improvement estimates. There is strong support for this as in a previous study where the magnitude of change was greater when baseline HRQoL scores were low; thus, the capacity for change was greater for the improvement MID resulting in greater than expected values in the improvement [23].

Finally, the difference in results may be linked to the differences in the measurement properties of the EQ-5D-3 L and EQ-5D-5L. The 5 L version of the EQ-5D can detect smaller changes and has fewer ceiling effects compared with the 3 L version, and both measures produce different values for changes in health [34]. Thus, the 3 L and 5 L MID values may not be comparable based on their magnitude of change.

The Institute for Quality and Efficiency in Healthcare (IQWiG), Germany, has implemented a new universal approach for all PRO measures [35], specifying a responder threshold of 15% of the PRO scale under study for determination of treatment efficacy. For an individual to have a perceived meaningful change in both the QLQ-C30 functional scales (0-100–point scale) and the EQ-VAS (0-100–point scale), this would require the improvements in both scales to be equal to or greater than 15 points. It should be noted that the IQWiG recommendation of 15% related to a patient-level (responder) threshold, whereas for group-level data the recommendation is for a standardised mean difference of 0.20. As this analysis is at patient-level, the reference value of 15% is used. Considering our results, where we found an MID of 5 points for both improvement and deterioration in the EQ-VAS, this universal approach may overestimate the response needed for a treatment by 3 times the value of what is considered a meaningful minimal difference in the VAS scale. Previously used MIDs were a 10% point change in the QLQ-C30 [7] and 7–10 points for the EQ-VAS [14]. Additional important contextual characteristics (i.e., condition, patient population, and stage of disease) are also not considered in the determination of this 15-point threshold. Moreover, the baseline health status reported by patients also needs to be carefully considered. If the health status score is high at the baseline, such as the EQ-VAS found in this study, a 15-point improvement could be difficult or impossible to achieve. Further research to determine the relevance of the IQWiG responder threshold to patients is required, and further exploration is needed to determine whether a 15% improvement is a meaningful change perceived by the cancer patients.

This study followed best practice recommendations and guidelines from the FDA[1] and the EMA [5, 8]. We used multiple anchors with 1 objective clinical measure (ECOG) and 3 PROs (EQ-VAS, EORTC, and QLQ-C30 overall health and quality of life items). The PRO measures were valid and reliable and helped to support our results. In addition, there is agreement in the literature that the use of a patient-reported anchor-based approach is the optimal way to determine the MID, as it directly captures patients' preferences and is now considered the gold standard approach [36, 37]. Future research could consider using a predictive model based on logistic regression analysis to determine MID values. While distribution-based methods provided additional estimates, this approach has often been criticized for not being directly related to clinical relevance. These methods are arbitrary effect sizes or indicators of measurement precision and may not necessarily reflect meaningful changes from a patient's perspective. As such, they are complementary to anchor-based methods, which are more closely tied to clinical significance through their reliance on external criteria, or 'anchors', that represent meaningful changes in health status. Given these limitations, the distribution-based MCD values should be viewed as supplementary information.

The choice of a clinical anchor in the MID studies with retrospective longitudinal data is challenging. Within our study, the ECOG was selected as the clinical anchor; however, it demonstrated a poor correlation with the outcome measures. The ECOG performance status is used by the doctors and researchers to assess how a disease is progressing and affecting the patient's daily life. The weak correlation between the HRQoL measures and the anchor further affirms the current gold standard approach that the anchor selection should be based on the patient-reported measures.

Conclusion

This study estimated the minimal amount of change in scores of the EQ-5D-5L utility index, EQ-VAS, and EORTC QLQ-C30 subscales required for patients to experience a relevant minimally important change that is beneficial for determining the impact and effectiveness of treatments. The identification of robust MID in cancer patients' HRQoL like the MIDs identified in this analysis can be used as a tool to aid researchers in the determination of the sample size required for clinical trials, responder definitions that are important for

---

[1] At the time of writing, these are draft guidelines and may be updated in the future.

regulatory approval and interpretation of changes in HRQL over time using the measures.

Abbreviations

AP	Appetite loss
CCG	Clinical change group
CF	Cognitive functioning
COA	Clinical outcome assessment
DY	Dyspnoea
ECOG	Eastern Cooperative Oncology Group
EF	Emotional functioning
EMA	European Medicines Agency
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer Quality of life Questionnaire Core 30
EQ-VAS	EuroQol Visual Analogue Scale
ES	Effect size
FA	Fatigue
FDA	United States Food and Drug Administration
G-BA	Federal Joint Committee (Germany)
HRQoL	Health-related quality of life
HTA	Health technology assessment
IQWiG	Institute for Quality and Efficiency in Healthcare
MID	Minimally important difference
MCD	Minimal detectable change
NV	Nausea/vomiting
PA	Pain
PF	Physical function
PRO	Patient-reported outcome
QL	Global health status/quality of life
RF	Role functioning
SD	Standard deviation
SEM	Standard error of measurement
SF	Social functioning
SL	Insomnia
TGA	Therapeutic Goods Administration
UK	United Kingdom
US	United States

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12955-024-02294-3.

> Supplementary Material 1

Author contributions

Conceptualization of the study was done by J.W.S, B.B, L.L, Y.O, J.B and I-A.S. Methodology was defined by J.W.S, B.B, L.L, Y.O, J.B, T.L S.B.OS and I- A.S. S.B lead on the formal analysis with assistance from T.L . S.B. and Y.O wrote the original draft of the manuscript. All authors reviewed the manuscript and have approved the submitted version.

Funding

This study was sponsored by Bristol Myers Squibb (BMS).

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethical approval

The study received ethical approval as a secondary analysis of panel data by a reviewer working under the association of research managers and administrators (ARMA). This project draws on data drawn from the Cancer 2015 dataset which did not include any data that identified an individual data subject. Access to the data on the was granted by the Cancer2015 data custodians as participants consented to use of their data for other research purposes.

Consent for Publication

The authors give consent for the publication of this manuscript, which can include photograph(s) and/or videos and/or case history and/or details within the text ("Material") to be published in the above Journal and Article.

Competing interests

The authors declare no competing interests.

Author details

¹Putnam, 22-24 Torrington Place, Fitzrovia, London WC1E 7HJ, UK
²Bristol Myers Squibb, Uxbridge, England, UK
³Bristol Myers Squibb, Lawrenceville, NJ, USA

References

1. CONSTITUTION of the World Health Organization. Chron World Health Organ. 1947;1:29–43.
2. US Food and Drug Administration. Discussion Document for Patient-Focused Drug Development Public Workshop on Guidance 4: Incorporating clinical outcome assessments into endpoints for regulatory decision-making. 2019.
3. Kennedy-Martin M, Slaap B, Herdman M, van Reenen M, Kennedy-Martin T, Greiner W, Busschbach J, Boye KS. Which multi-attribute utility instruments are recommended for use in cost-utility analysis? A review of national health technology assessment (HTA) guidelines. Eur J Health Econ. 2020;21:1245–57.
4. Remenschneider AK, D'Amico L, Gray ST, Holbrook EH, Gliklich RE, Metson R. The EQ-5D: a new tool for studying clinical outcomes in chronic rhinosinusitis. Laryngoscope. 2015;125:7–15.
5. European Medicines Agency. Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man - the use of patient-reported outcome (PRO) measures in oncology studies. 2016.
6. Shaw C, Longworth L, Bennett B, Ruane P, Watson C, Francis L. J S: To what extent is EQ-5D used as a tool for clinical outcome assessment? In *ISPOR Europe (Virtual)*. 2021.
7. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol. 1998;16:139–44.
8. US Food and Drug Administration. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims (Guidance for Industry). 2009.
9. de Vet HC, Terwee CB. The minimal detectable change should not replace the minimal important difference. J Clin Epidemiol. 2010;63:804–5. author reply 806.
10. Musoro JZ, Bottomley A, Coens C, Eggermont AM, King MT, Cocks K, Sprangers MA, Groenvold M, Velikova G, Flechtner HH, Brandberg Y. Interpreting European Organisation for Research and Treatment for Cancer Quality of life questionnaire core 30 scores as minimally importantly different for patients with malignant melanoma. Eur J Cancer. 2018;104:169–81.
11. Hu X, Jing M, Zhang M, Yang P, Yan X. Responsiveness and minimal clinically important difference of the EQ-5D-5L in cervical intraepithelial neoplasia: a longitudinal study. Health Qual Life Outcomes. 2020;18:324.
12. McClure NS, Sayah FA, Ohinmaa A, Johnson JA. Minimally important difference of the EQ-5D-5L index score in adults with type 2 diabetes. Value Health. 2018;21:1090–7.
13. Tsai APY, Hur SA, Wong A, Safavi M, Assayag D, Johannson KA, Morisset J, Fell C, Fisher JH, Manganas H, et al. Minimum important difference of the EQ-5D-5L and EQ-VAS in fibrotic interstitial lung disease. Thorax. 2021;76:37–43.
14. Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. Health Qual Life Outcomes. 2007;5:70.
15. Parisot JP, Thorne H, Fellowes A, Doig K, Lucas M, McNeil JJ, Doble B, Dobrovic A, John T, James PA. Cancer 2015: a prospective, Population-Based Cancer Cohort—Phase 1: feasibility of Genomics-guided Precision Medicine in the clinic. J Pers Med. 2015;5:354–69.
16. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G, Badia X. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20:1727–36.
17. Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: past, Present and Future. Appl Health Econ Health Policy. 2017;15:127–37.

18. Norman R, Mulhern B, Lancsar E, Lorgelly P, Ratcliffe J, Street D, Viney R. The Use of a Discrete Choice Experiment Including both Duration and Dead for the development of an EQ-5D-5L value set for Australia. PharmacoEconomics. 2023;41:427–38.
19. Pickard AS, Law EH, Jiang R, Pullenayegum E, Shaw JW, Xie F, Oppe M, Boye KS, Chapman RH, Gong CL, et al. United States Valuation of EQ-5D-5L Health States using an International Protocol. Value Health. 2019;22:931–41.
20. Fayers P, Aaronson NK, Bjordal K. EORTC QLQ-C30 Scoring Manual. 3rd edition. Brussels, Belgium; 2001.
21. Mouelhi Y, Jouve E, Castelli C, Gentile S. How is the minimal clinically important difference established in health-related quality of life instruments? Review of anchors and methods. Health Qual Life Outcomes. 2020;18:136.
22. Ousmen A, Touraine C, Deliu N, Cottone F, Bonnetain F, Efficace F, Brédart A, Mollevi C, Anota A. Distribution- and anchor-based methods to determine the minimally important difference on patient-reported outcome questionnaires in oncology: a structured review. Health Qual Life Outcomes. 2018;16:228.
23. Bedard G, Zeng L, Zhang L, Lauzon N, Holden L, Tsao M, Danjoux C, Barnes E, Sahgal A, Poon M, Chow E. Minimal important differences in the EORTC QLQ-C30 in patients with advanced cancer. Asia Pac J Clin Oncol. 2014;10:109–17.
24. Musoro JZ, Coens C, Singer S, Tribius S, Oosting SF, Groenvold M, Simon C, Machiels JP, Grégoire V, Velikova G. Minimally important differences for interpreting European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 scores in patients with head and neck cancer. Head Neck. 2020;42:3141–52.
25. Nolan CM, Longworth L, Lord J, Canavan JL, Jones SE, Kon SS, Man WD. The EQ-5D-5L health status questionnaire in COPD: validity, responsiveness and minimum important difference. Thorax. 2016;71:493–500.
26. Trigg A, Griffiths P. Triangulation of multiple meaningful change thresholds for patient-reported outcome scores. Qual Life Res. 2021;30:2755–64.
27. Hjermstad MJ, Fossa SD, Bjordal K, Kaasa S. Test/retest study of the European Organization for Research and Treatment of Cancer Core Quality-of-life questionnaire. J Clin Oncol. 1995;13:1249–54.
28. Long D, Polinder S, Bonsel GJ, Haagsma JA. Test-retest reliability of the EQ-5D-5L and the reworded QOLIBRI-OS in the general population of Italy, the Netherlands, and the United Kingdom. Qual Life Res. 2021;30:2961–71.
29. Cohen J. Statistical power analysis. Curr Dir Psychol Sci. 1992;1:98–101.
30. Cocks K, King MT, Velikova G, de Castro G Jr., St-James M, Fayers PM, Brown PM. Evidence-based guidelines for interpreting change scores for the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30. Eur J Cancer. 2012;48:1713–21.
31. Tsiplova K, Pullenayegum E, Cooke T, Xie F. EQ-5D-derived health utilities and minimally important differences for chronic health conditions: 2011 Commonwealth Fund Survey of Sicker adults in Canada. Qual Life Res. 2016;25:3009–16.
32. Xu RH, Wong EL, Cheung AW. Estimation of minimally important difference of the EQ-5D-5L utility scores among patients with either hypertension or diabetes or both: a cross-sectional study in Hong Kong. BMJ Open. 2020;10:e039397.
33. Mann C. Observational research methods. Research design II: cohort, cross sectional, and case-control studies. Emerg Med J. 2003;20:54–60.
34. Hernandez G, Garin O, Dima AL, Pont A, Martí Pastor M, Alonso J, Van Ganse E, Laforest L, de Bruin M, Mayoral K, et al. EuroQol (EQ-5D-5L) validity in assessing the quality of life in adults with Asthma: cross-sectional study. J Med Internet Res. 2019;21:e10178.
35. IQWiG. General Methods Version 6.1. Köln; 2022.
36. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR. Clinical significance Consensus Meeting Group: methods to explain the clinical significance of health status measures. Mayo Clin Proc. 2002;77:371–83.
37. King MT. A point of minimal important difference (MID): a critique of terminology and methods. Expert Rev Pharmacoecon Outcomes Res. 2011;11:171–84.

Publisher's note