Health Economics Review

RESEARCH

Open Access

A comparison of measurement properties between EORTC QLU-C10D and FACT-8D in patients with hematological malignances

Yiyin Cao[1†], Haofei Li[1†], Ling Jie Cheng[2], Madeleine T. King[3], Georg Kemmler[4], David Cella[5], Hongjuan Yu[6*], Weidong Huang[1*] and Nan Luo[2]

Abstract

Objective To perform a comparison of the measurement properties of two cancer-specific Multi-Attribute Utility Instruments (MAUIs), EORTC QLU-C10D and FACT-8D, in Chinese patients with hematologic malignancies (HM).

Methods We conducted a longitudinal study on patients with HM in China, using QLU-C10D and FACT-8D at baseline and follow-up (3–4 months from baseline). We assessed: (i) convergent validity using Spearman's rank correlation test (r) with EQ-5D-5L; (ii) clinical-groups validity by differentiating cancer stages, overall health assessment (OHA), Eastern Cancer Oncology Group (ECOG) performance status, and mental health status. We also examined clinical validity with effect size (ES) and relative efficiency (RE); (iii) responsiveness to changes in patient self-perception using receiver operating characteristics (ROC) curves and area under the curves (AUC); and (iv) agreement using intraclass correlation coefficients (ICC) and visualized with Bland-Altman plot.

Results Among the 308 patients with HM at baseline, 131 completed the follow-up survey. Agreement between the two measures was high (ICC = 0.76). Both measures were highly correlated with EQ-5D-5 L and significantly differentiated ($p < 0.001$) among groups categorized by cancer stage, OHA performance status, and mental health. ESs for QLU-C10D were numerically higher for cancer stage, OHA, and performance status (ES = 0.53–1.49), whereas ES was higher for FACT-8D and mental health status (ES = 1.35). Responsiveness was higher for QLU-C10D (AUC = 0.84) compared to FACT-8D (AUC = 0.78).

Conclusion Both QLU-C10D and FACT-8D are valid cancer-specific MAUIs for evaluating patients with HM. However, scholars should consider their slight differences in focus when choosing between the two measures.

Keywords QLU-C10D, FACT-8D, Measurement properties, Cancer, Hematological malignance

†Yiyin Cao and Haofei Li contributed equally to this work and shared first authorship.

*Correspondence:
Hongjuan Yu
yuhongjuan2008@163.com
Weidong Huang
huangweidong@hrbmu.edu.cn
[1]School of Health Management, Harbin Medical University, Harbin 150081, China
[2]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore
[3]School of Psychology, University of Sydney, Sydney 2006, Australia
[4]Department of Psychiatry and Psychotherapy, Innsbruck Medical University, Innsbruck, Austria
[5]Feinberg School of Medicine, Northwestern University, Chicago 60601, USA
[6]Department of Hematology, The First Affiliated Hospital of Harbin Medical University, Harbin, China

Introduction

Cancer, one of the most lethal diseases, causes approximately 10 million deaths globally annually [1, 2]. The economic burden of cancer is substantial, with projections indicating the global economic burden of cancer could reach $25.2 trillion over the next 30 years, starting from 2020. This is equivalent to an annual 0.55% tax on the global gross domestic product [3].

Health technology assessment (HTA), particularly cost-utility analysis (CUA), is gaining importance in cancer care due to the influx of new medical technologies and financial constraints [4]. The quality-adjusted life-year (QALY) is the recommended health outcome metric for CUA [5]. The quality adjustment metric in QALYs is health utility. Consequently, tools for quantifying utility values, including both generic multi-attribute utility instruments (MAUIs) (e.g., EQ-5D-5 L) and cancer-specific MAUIs (e.g., QLU-C10D), have been developed to measure health utility in patients with cancer.

Although generic MAUIs are widely used in HTA of cancer interventions [6], they have notable limitations, including a lack of sensitivity to cancer-specific symptoms [7]. This insensitivity can obscure small yet critical differences in treatments outcomes across different cancer stages, impacting both QALY estimations and CUA results [8–10]. To address these issues, the Multi-Attribute Utility in Cancer Consortium (MAUCaC) used the European Organization for Research and Treatment of Cancer (EORTC) Core Quality of Life Questionnaire (QLQ-C30) [11], to create the Quality of Life Utility Measure-Core 10 Dimensions (EORTC QLU-C10D) [12]. Additionally, MAUCaC created the Functional Assessment of Cancer Therapy 8 Dimensions (FACT-8D) [13] adapted from the Functional Assessment of Cancer Therapy-General (FACT-G) [14]. As the QLQ-C30 and the FACT-G are the most widely used health-related quality of life (HRQoL) questionnaires in cancer clinical trials, accounting for 91% of HRQoL assessment in published cancer clinical trials [15], the QLU-C10D and FACT-8D together facilitate the inclusion of the majority of HRQoL data from cancer clinical trials in HTA.

QLU-C10D and FACT-8D share some strengths. First, they include cancer-relevant dimensions like nausea and sleep problems, thereby better reflecting HRQoL impacts on QALY estimations. Second, multiple countries have completed or are undertaking valuation studies for QLU-C10D and FACT-8D with 14 value sets for QLU-C10D [16–26] and three for FACT-8D [13, 27, 28] currently published. These country-specific value sets, developed with a standard valuation protocol in collaboration with MAUCaC [29], facilitate international research comparability. Third, both instruments can derive utilities from their parent instruments (QLQ-C30 and FACT-G), either prospectively or retrospectively. This approach helps reduce additional response burden, particularly given that their parent instruments are widely recognized condition-specific HROoL measures in cancer research [15].

Evidence to date suggests that both QLU-C10D and FACT-8D may offer favorable measurement properties compared to popular generic MAUIs like EQ-5D [30–35]. These studies demonstrate the advantage of these cancer- specific MAUIs in terms of responsiveness, but evidence about the advantage in terms of construct validity is mixed. For example, Shaw et al. [34] and Gamper et al. [32] demonstrated the superiority of QLU-C10D over EQ-5D-3 L with respect to construct validity in patients with multiple solid tumors and myelodysplastic syndromes, respectively, but Pan et al. [30] demonstrated that EQ-5D-5 L had better discriminative power in gastric cancer patients. To date, there has been no direct comparison of the measurement properties of QLU-C10D and FACT-8D, either cross-sectionally or longitudinally, limiting the evidence base for selecting between these two cancer-specific MAUIs in economic evaluations and impacting the precision of CUAs in various oncology settings.

Therefore, this study aimed to compare the measurement properties of QLU-C10D and FACT-8D in Chinese patients with hematological malignances, specifically focusing on construct validity (convergent and clinical known-groups), responsiveness, and agreement.

Materials & methods

Study design and patients

From August 2022 to December 2023, we conducted a longitudinal survey on Chinese patients with HM in three tertiary hospitals in Harbin, the capital of Heilongjiang Province, China. Doctors or nurses from these hospitals screened and selected participants for this study. Inclusion criteria were: (1) clinical diagnosis of HM; (2) expected survival time of more than 1 year; (3) at least 18 years old; and (4) able to read and communicate in Chinese and ability to complete questionnaires. Trained interviewers obtained informed consent and conducted face-to-face interviews with consenting patients in the hematology ward, recording responses on paper questionnaires.

The baseline questionnaire included the assessment of QLQ-C30, FACT-G, EQ-5D-5 L, Eastern Cancer Oncology Group (ECOG) [36], Kessler Psychological Distress Scale (Kessler-10) [37, 38] and an overall health assessment (OHA) question ("How is your overall health today?" [39], with five response options: "excellent", "good", "fair", "poor", "very poor"). Additionally, the interviewer also collected socio-demographic and clinical characteristics such as gender, age, frequency of health check-ups, cancer type and cancer stage from medical records.

After a three-month interval, participants were contacted for a follow-up interview. Via the interviewer, participants completed the QLQ-C30 and FACT-G, and were also asked about their health transition, responding to the question, "How is your overall health now compared to the last time you were asked to answer these questionnaires?" The response options included "better than before", "about the same as before", and "worse than before".

Instruments

QLU-C10D

The QLU-C10D, developed by MAUCaC 2010–2016 [12, 16, 29], is a derivative of the QLQ-C30, encompassing 10 health dimensions (physical functioning, role functioning, social functioning, emotional functioning, pain, fatigue, sleep problems, appetite, nausea, and bowel problems). Each dimension has four levels of severity (not at all, a little, quite a bit, and very much). Thus, the QLU-C10D can represent 4^{10}= 1,048,576 health states. In this study, the Chinese version of QLQ-C30 was applied. Utility values were calculated based on the original QLQ-C30 scores, and the Australian QLU-C10D value set [16] was used because the Chinese QLU-C10D value set has not yet been published. Furthermore, Australia pioneered the development of a country-specific QLU-C10D value set, which has since become a reference standard for other nations, with theoretical utility values based on the Australian population ranging from −0.10 to 1 [29].

FACT-8D

The FACT-8D, developed by MAUCaC in 2014–2020 [13], originates from FACT-G and evaluates eight dimensions (pain, fatigue, nausea, sleep problems, work problems, social support problems, sadness, and future health worry), each with five levels of severity (none, a little bit, some, quite a bit, and very serious). Thus, the FACT-8D can describe 5^8 = 390,625 health states. The Chinese version of FACT-G was applied in this study. FACT-8D dimensions scores were derived from the nine FACT-G source scores (one for each dimension except social support problems, which contains two FACT-G items), standardized such that 1 represents the best level and 5 the worst level. FACT-8D utility values were calculated using the Australian population tariff, with values ranging from −0.54 to 1 [13] .Since there is no current FACT-8D value set in China and Australia developed the valuation protocol and first country-specific FACT-8D value set [13].

EQ-5D-5L

The EQ-5D-5L, developed by the EuroQol Group, comprises five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) with five severity levels each, which can describe 5^5 = 3125 health states [40]. The EQ-5D-5L questionnaire used in this study was the official, EuroQol-approved Chinese (for China) version. Previous studies have validated EQ-5D-5L's superior measurement properties over EQ-5D-3 L in Chinese hematologic disease patients [41]. Although the Chinese EQ-5D-5L value set has been developed [42], the Australian EQ-5D-5L value set was applied to ensure comparability with other instruments, with values ranging from −0.30 to 1 [43].

Statistical methods

Descriptive statistics were generated for all socio-demographic characteristics at baseline and follow-up. Pearson's chi-square test compared patients completing follow-up interviews with those who did not, assessing selection bias. Box plots and percentile distribution plots illustrated QLU-C10D and FACT-8D utility scores and dimension scores at baseline.

For convergent validity, we evaluated convergent validity by correlating QLU-C10D and FACT-8D utility scores using scatterplots and Pearson correlation coefficient. We also analyzed correlations between QLU-C10D/FACT-8D and EQ-5D-5 L, in both utility scores and dimensions, using Spearman rank correlation for categorical variables. All dimension scores were calculated based on raw scores, except for "sleep problems," "work problems," and "social support problems" in the FACT-8D, which were reverse-scored as outlined in previous study [13]. Correlation strength was categorized as weak (0.2–0.34), moderate (0.35–0.50), or strong (above 0.5) [30].We hypothesized strong, positive correlations in dimensions that are similar or capture the same concept, such as pain (included in all three instruments); mobility (EQ-5D-5 L) and physical functioning (QLU-10D); anxiety/depression (EQ-5D-5 L), emotional functioning (QLU-10D), and sadness and future health worry (FACT-8D).

Clinical validity (also called 'known-groups' validity) was assessed by examining the differences between patient groups based on cancer stages (stage I-IV), OHA status (excellent-very poor), ECOG performance (grade 0–4) [36], and mental health (Kessler-10 total scores: 10–19 points [low risk)], 20–24 points [relatively low risk], 25–29 points [relatively high risk], and 30–50 points [high risk]) [44]. We hypothesized that patients with more advanced cancer stage, poorer OHA status, poorer ECOG performance, or poorer mental health would have lower utility values [45–47]. Mean differences between groups were analyzed using two-sample t-tests and Mann-Whitney U-tests. Effect sizes (ES) were calculated for both the QLU-C10D and the FACT-8D by dividing the mean difference in utility scores between groups by the pooled standard deviation (SD). ES were interpreted as small (≥0.20), medium (≥0.50), and large (≥0.80) based on Cohen's criterion [48]. To compare their

discriminative capacity, relative efficiency (RE) was calculated based on the ratios of square of t-statistics from the t-tests [49].

Responsiveness was assessed using the ROC curve and AUC [50] to evaluate the ability of the QLU-C10D and of the FACT-8D to accurately categorize patients into improved and unchanged or worsened groups based on self-reported health changes (the answer to the health transition question). Change sensitivity and change specificity refer to the ability to categorize between disease improved and unimproved, respectively [51]. An AUC>0.7 was considered sufficiently responsive, with an AUC range between 0.5 (no accuracy responsiveness) and 1.0 (full accuracy responsiveness) [52].

Lastly, agreement between QLU-C10D and FACT-8D utility scores was assessed using intraclass correlation coefficients (ICC), categorized as low (ICC<0.39), moderate (ICC=0.40–0.74), or high (ICC>0.75) [53]. A Bland-Altman plot was constructed to visualize agreement, allowing for the identification of the correlation between measurement error and the most accurate estimate of the true value [54].

Statistical analysis was performed with Statistical Package for Social Sciences version 24.0, Stata version 13, and R version 4.0.5. Significance was defined at $p<0.05$.

Table 1 Demographic characteristics of patients with hematological malignances

Variable	Baseline (*N*=308)	3–4 Months Follow-up		
		Yes (*N*=131)	No (*N*=177)	*p*-value
Gender				
Male	144 (46.8%)	55 (42.0%)	89 (50.3%)	
Female	164 (53.2%)	76 (58.0%)	88 (49.7%)	0.166
Age				
≤29	25 (8.1%)	10 (7.6%)	15 (8.5%)	
30–39	41 (13.3%)	19 (14.6%)	22 (12.4%)	
40–49	55 (17.9%)	24 (18.3%)	31 (17.5%)	0.943
≥50	187 (60.7%)	78 (59.5%)	109 (61.6%)	
Registered residence				
City	144 (46.8%)	55 (42.0%)	89 (50.3%)	
Rural	164 (53.2%)	76 (58.0%)	88 (49.7%)	0.166
Marital status				
Unmarried	34 (11.0%)	14 (10.7%)	20 (11.3%)	
Married	232 (75.4%)	104 (79.4%)	128 (72.3%)	
Other	42 (13.6%)	13 (9.9%)	29 (16.4%)	0.243
Employment status				
Employed	187 (60.7%)	81 (61.8%)	106 (59.9%)	
Repaired	76 (24.7%)	30 (22.9%)	46 (26.0%)	
Unemployed	45 (14.6%)	20 (15.3%)	25 (14.1%)	0.817
Frequency of health check-ups				
Regularly	111 (36.0%)	43 (32.8%)	68(38.4%)	
Occasionally	113 (36.7%)	49 (37.4%)	64 (36.2%)	
Hardly ever	84 (27.3%)	39 (29.8%)	45 (25.4%)	0.547
Family care				
Yes	285 (92.5%)	122 (93.1%)	163 (92.1%)	
No	23 (7.5%)	9 (6.9%)	14 (7.9%)	0.828
Economic pressure				
No	22 (7.1%)	8 (6.2%)	14 (7.9%)	
Mild	58 (18.8%)	24 (18.3%)	34 (19.2%)	
Moderate	112 (36.4%)	51 (38.9%)	61 (34.5%)	
Severe	116 (37.7%)	48 (36.6%)	68 (38.4%)	0.837
Disease type				
Lymphoma	171 (55.5%)	68 (51.9%)	103 (58.2%)	
Leukemia	75 (24.4%)	35 (26.7%)	40 (22.6%)	
Myeloma	62 (20.1%)	28 (21.4%)	34 (19.2%)	0.539
Mean response time (months)		3.9		

Results

Demographic characteristics

Table 1 presents the characteristics of patients with HM at baseline and those who participated in the follow-up survey. Among the 308 patients recruited at baseline, 53.2% were female, 60.7% were aged over 50. Lymphoma was the most common disease type, accounting for 55.5% of cases. A total of 131 (42.5%) patients completed the follow-up, mean 3.9 months for response time. Importantly, no significant demographic differences were found between follow-up completers and non-completers.

Utility and dimensional distribution

Figure 1 illustrates the distributions of utility and dimension scores for both QLU-C10D and FACT-8D. FACT-8D had a broader utility range (-0.20-1.0), with higher mean (0.63) and median (0.66) utility scores compared to QLU-C10D (0.61, 0.60). In terms of dimensional distribution, QLU-C10D physical functioning was the most evenly distributed, whereas patients reported more severe problems in FACT-8D sleep and work dimensions. 5.8% reported full health with QLU-C10D, whereas only 0.3% did with FACT-8D. Detailed response frequencies per dimension are in Table A1-A2 in the Appendix.

Convergent validity

The Pearson correlation coefficient between the utility scores of QLU-C10D and FACT-8D was 0.76 (95% CI=0.70–0.81), indicating strong positive association. Both FACT-8D and QLU-C10D utility scores also exhibited a negative skew, as confirmed by the Shapiro-Wilk test (Fig. 2). At the dimension level, the dimensions related to the similarity of QLU-C10D and FACT-8D exhibit a strong or moderate correlation (Table A3 in the Appendix).

The utility scores of QLU-C10D were highly positively correlated with EQ-5D-5 L (r=0.65, 95% CI=0.59–0.71), with dimensions correlating from 0.22 to 0.69 (Fig. 3).

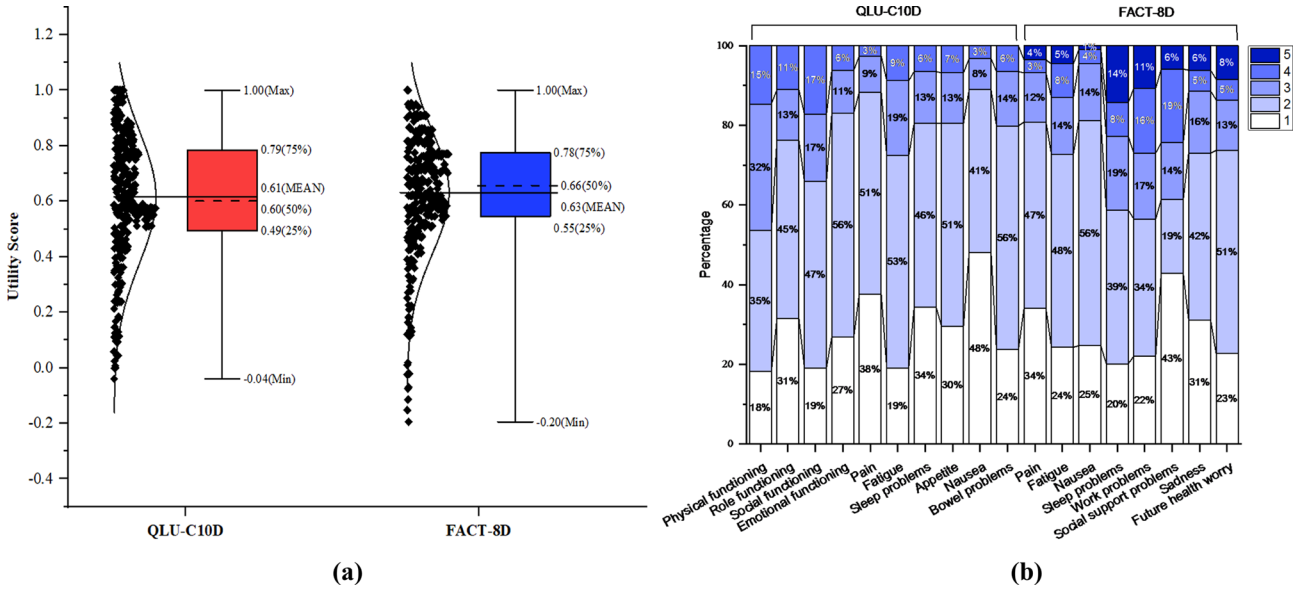

Fig. 1 **a**. Distributions of utility **b**. Distributions of dimension scores Distributions of utility and dimension scores for QLU-C10D and FACT-8D Note: Figure B: Color coding reflects the dimensions levels of the instruments; for both, Level 1 is the best health level; for QLU-C10, the worst health level is 4; for FACT-G, the worst health level is 5

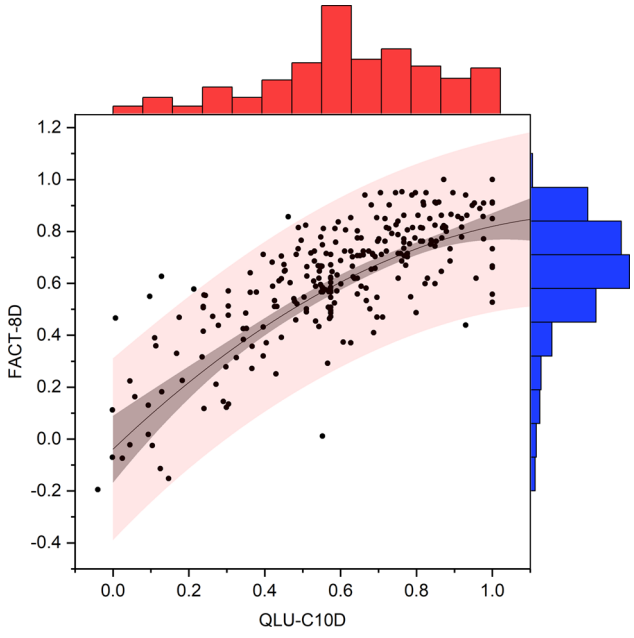

Fig. 2 Correlation and distribution of QLU-C10D and FACT-8D. Note The line shows a loess smoother with 95% confidence interval

In QLU-C10D, dimensions that contain similar concepts to the EQ-5D-5 L displayed strong positive correlations: pain (EQ-5D-5 L and QLU-C10D) (r=0.69); physical functioning (QLU-C10D) and mobility (EQ-5D-5 L) (r=0.63); emotional functioning (QLU-C10D) and anxiety/depression (EQ-5D-5 L) (r=0.52). Functioning dimensions had a stronger positive correlation with EQ-5D-5 L compared to symptom dimensions.

For the FACT-8D the utility scores showed a highly positive correlation with EQ-5D-5 L (r=0.58 95% CI=0.48–0.66), and the correlation of dimensions ranged from 0.02 to 0.68. Pain (FACT-8D) displayed a strong positive correlation with all dimensions of EQ-5D-5 L except self-care. Similarly, sadness (FACT-8D) exhibited a strong positive correlation with anxiety/depression (EQ-5D-5 L) (r=0.50).

Clinical validity

In Table 2, both the QLU-C10D and FACT-8D utility values were significantly different (p<0.01) among groups categorized by cancer stage, OHA, ECOG and mental health, with ES ranging from 0.53 to 1.49 for the QLU-C10D, and 0.47 to 1.42 for the FACT-8D. Clinical-group validity hypotheses were met in all tested groups. The Mann-Whitney U test yielded similar results (Table A4 in the Appendix). RE estimates indicated QLU-C10D better differentiated between cancer stage, OHA and ECOG, while FACT-8D better differentiated between mental health levels, and the mean RE is 1.48.

Responsiveness

Based on patients' self-reports, of 131 patients who completed 3–4 months follow-up, 87 (66.4%) patients were categorized as improved overall health group, and 44 (33.6%) were categorized as unchanged or worsened health group. Figure 4 shows the ROC curves for the utility and dimension of QLU-C10D and FACT-8D scores. Table A5 in the Appendix provides AUC, sensitivity and specificity estimates. The AUC values for the QLU-C10D utility values and dimensions ranged from 0.62

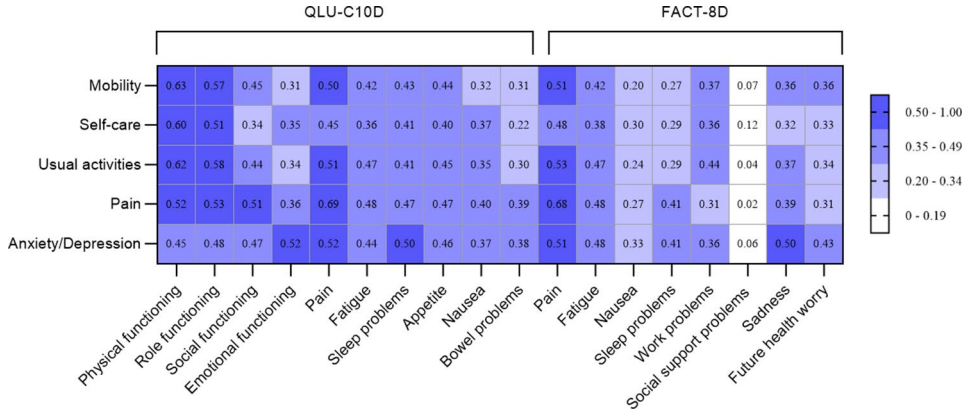

Fig. 3 Convergent validity of QLU-C10D and FACT-8D. In the colored cells, all dimensions are significantly correlated ($P<0.05$). In the five colorless cells, only social support problems (FACT-8D) and self-care (EQ-5D-5 L) were significantly correlated, while the others show no significant correlation. Note: Based on Spearman correlation

Table 2 Clinical validity of QLU-C10D and FACT-8D ($n=308$)

Instruments	Cancer stage	Number of patients	Mean (SD)	*p*-value	ES	RE
QLU-C10D	I/II	215	0.63 (0.19)			
	III/IV	93	0.51 (0.26)	<0.001	0.53	
FACT-8D	I/II	215	0.65(0.18)			
	III/IV	93	0.55(0.25)	0.001	0.47	1.27
Instruments	**OHA**	**Number of patients**	**Mean (SD)**	***p*-value**	**ES**	**RE**
QLU-C10D	Excellent/good	24	0.64 (0.22)			
	Fair/poor/very poor	284	0.29 (0.25)	<0.001	1.49	
FACT-8D	Excellent/good	24	0.66(0.19)			
	Fair/poor/very poor	284	0.32(0.29)	<0.001	1.42	1.48
Instruments	**ECOG**	**Number of patients**	**Mean (SD)**	***p*-value**	**ES**	**RE**
QLU-C10D	Grade 0/1	219	0.70 (0.18)			
	Grade 2–4	89	0.40 (0.23)	<0.001	1.46	
FACT-8D	Grade 0/1	219	0.69(0.16)			
	Grade 2–4	89	0.49(0.27)	<0.001	0.93	2.45
Instruments	**Mental health**	**Number of patients**	**Mean (SD)**	***p*-value**	**ES**	**RE**
QLU-C10D	Low risk/relatively low risk	187	0.71 (0.18)			
	Relatively high risk/high risk	121	0.47 (0.24)	<0.001	1.14	
FACT-8D	Low risk/relatively low risk	187	0.73 (0.14)			
	Relatively high risk/high risk	121	0.48 (0.23)	<0.001	1.35	0.71

ES: Effect size; RE: Relative efficiency; SD: Standard deviations

Note: In the RE calculation, the denominator is the square of *t*-statistics for QLU-C10D, and the numerator is the square of *t*-statistics for FACT-8D. A RE value > 1 indicates superior clinical-groups validity for QLU-C10D, while a RE value < 1 suggests the opposite

The analyses of cancer stage included only lymphoma and myeloma cases ($N=233$), while all other analyses were based on the entire sample ($N=308$)

to 0.84, with the highest responsiveness observed in the four functional dimensions: physical, role, social, and emotional functioning. The AUC values for the FACT-8D utility values and dimensions were somewhat lower, ranging from 0.57 to 0.78. Except for FACT-8D social support problems, all other AUC values indicated significant responsiveness Figure 5.

Agreement

The ICC of QLU-C10D and FACT-8D utility scores was 0.76 (95% CI=0.71–0.80), indicating a high degree of agreement between the two measures. The Bland-Altman plot showed the same proportional bias, and the 95% limits of agreement were within the range of -0.33 and 0.29.

Discussion

Since the development of QLU-C10D and FACT-8D, few studies have evaluated their measurement properties in cancer populations. To our knowledge, this is the first study to conduct a formal comparison of QLU-C10D and FACT-8D. Our findings reveal the favorable performance of both QLU-C10D and FACT-8D in terms of construct validity and responsiveness. Additionally, there is good agreement between the two measures. Importantly, these

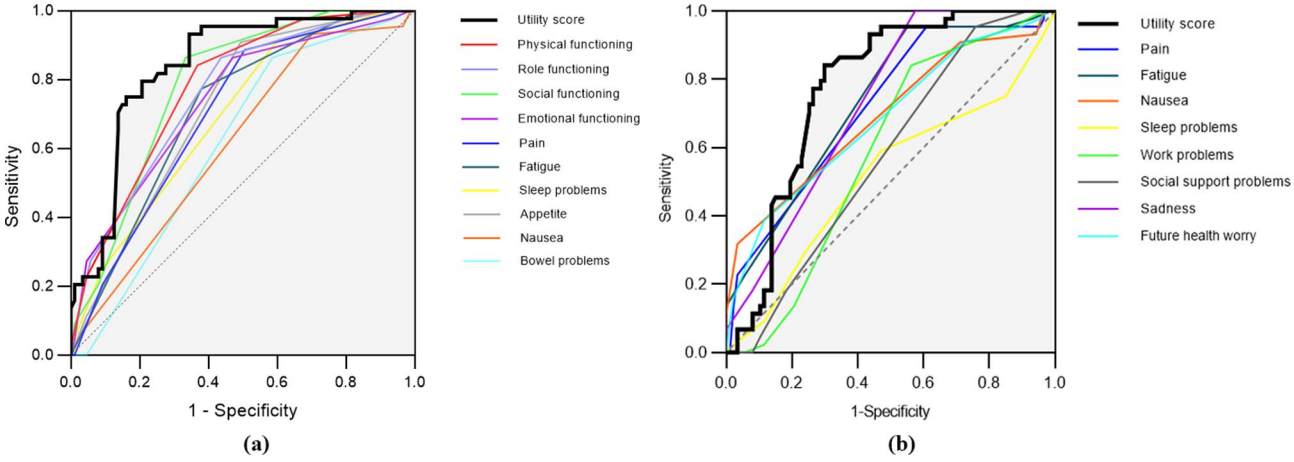

Fig. 4 **a**. QLU-C10D. **b**. FACT-8D ROC curves for the change scores for the QLU-C10D (panel A) and FACT-8D (panel B) and their dimensions ($n = 131$)

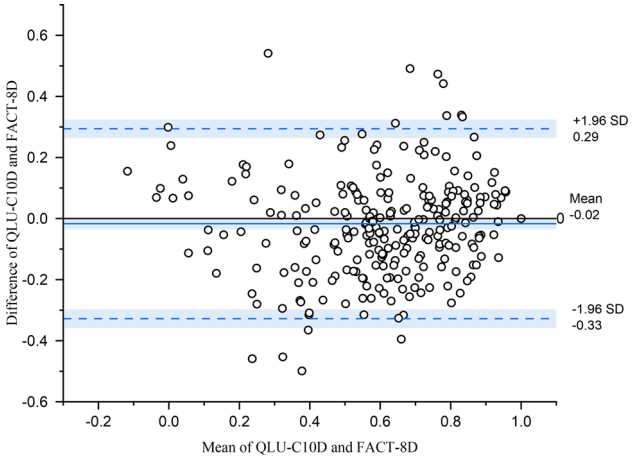

Fig. 5 Bland-Altman plot of QLU-C10D and FACT-8D utility scores SD: Standard deviations

results affirm the success of these two new preference-based measures, indicating that they have inherited the robust measurement properties observed in QLQ-C30 and FACT-G [55]. The similarity may also be attributed to the shared or similar health dimensions in both measures, as well as the use of similar valuation methods and study protocols [13, 29]. However, they exhibit their own relative merits, which contribute to the choice between these two cancer-specific MAUIs in a specific setting, further to considerations in choosing between the source questionnaires, QLQ-C30 and the FACT-G [55].

While not significantly different on mean utility scores, the FACT-8D yielded a broader range of utility variation towards lower values, and a smaller ceiling effect. Several factors may have contributed to this. First, the FACT-8D dimensions each have five levels, which allow more detailed reporting of poorer health compared to the four levels in QLU-C10D, given that each level deviation from full health results in a utility decrement [32]. Notably, the Australian FACT-8D value set [13] used in this study showed substantial utility decrements (-0.398 to -0.112) for the worst levels in each dimension, which were significantly lower than those in the Australian QLU-C10D (-0.25 to -0.037) [16]. The second factor is the intrinsic nature of the dimensions. In the same category of patients, the FACT-8D's dimensions of social support and future health worry reported more problems, while these dimensions have no related concepts in the QLU-C10D.The third factor is the utility weights assigned by the Australian population to these intrinsically different set of dimensions. The FACT-8D also exhibited a broader range of measurements at the upper end of the instrument, indicating that it might be more advantageous for HTA in patient cohorts with milder conditions.

Convergent validity analysis showed that QLU-C10D demonstrated a stronger correlation with EQ-5D-5 L compared to FACT-8D in both utility and dimensions numerically. It is worth highlighting that the correlations between the functional dimensions of QLU-C10D and EQ-5D-5 L were significantly higher than those of the symptom dimensions, aligning with previous research [31]. This reflects that the symptom dimensions of QLU-C10D identify specific symptoms in patients with cancer not recognized by EQ-5D-5 L. As for FACT-8D, only the dimensions of pain and sadness were highly correlated with EQ-5D-5 L, with other dimensions showing weak-to-moderate correlations, reflecting less overlap in conceptual content coverage. Surprisingly, the future health worry of FACT-8D correlated moderately with the anxiety/depression of EQ-5D-5 L, rather than being strongly correlated. We speculate that future health worry may also be influenced by external factors (e.g. disease status, economic situation) and differs conceptually from the immediate psychological feelings represented by the anxiety/depression in both QLU-C10D and EQ-5D-5 L, and sadness in FACT-8D. Future research should explore the impact of external factors on the future health worry

dimension among patients with cancer, through comparative studies involving individuals with varying disease characteristics or socioeconomic backgrounds, in order to better understand how these factors affect measurement performance. Our study highlights EQ-5D-5 L 's limited capacity to capture cancer-specific symptoms, underscoring the necessity for cancer-specific instruments in assessing the HRQoL of cancer populations [31].

Our analyses show that both instruments effectively identify patient groups with different health statuses based on four criteria, supporting their clinical validity. The QLU-C10D demonstrated superior ESs in distinguishing patients' conditions related to cancer stage, overall health, and ECOG, while the FACT-8D exhibited greater ES specifically in distinguishing patients' mental health. Several factors contribute to these differences. First, the QLU-C10D has a more comprehensive description system, giving it an advantage in distinguishing between well-defined clinical groups [32]. Second, QLU-C10D includes two cancer-specific dimensions, appetite and bowel problems, not present in FACT-8D, which may aid in differentiating patients' physical health. Third, FACT-8D places a stronger emphasis on assessing social interactions and mental health over physical health. Particularly noteworthy is the pronounced ES advantage of QLU-C10D in differentiating ECOG, where the dimension of physical functioning and role functioning, representing the capacity for activities of daily living, contributes substantially. And QLU-C10D further categorizes physical functioning into "short distance walking" and "long distance walking", allowing precise differentiation of patients' physical abilities. Physical functioning, critical in clinical trials for assessing treatment efficacy [56, 57], , also holds independent prognostic value for survival across various cancers [58], highlighting the significance of this advantage in QLU-C10D. In contrast, FACT-8D has a slightly superior ES performance in discriminating mental health, arguably due to its inclusion of two mental health-related dimensions—sadness and future health worry (25% of its dimensions), compared to QLU-C10D's single emotional functioning dimension (10% of its dimensions), providing an enhanced capacity for discriminating mental health.

One strength of our study is using an anchor-based approach, measuring changes in self-perceived overall health over 3–4 months to assess the responsiveness of QLU-C10D and FACT-8D. Although we didn't directly compare their responsiveness with EQ-5D-5L, previous studies have demonstrated their superior responsiveness compared to the EQ-5D-5L [31, 34, 35]. Our findings indicated that both QLU-C10D and FACT-8D exhibit acceptable responsiveness, rendering both instruments suitable for assessing the economic evaluation of interventions for patients with HM, however, and numerically higher for QLU-C10D as compared to FACT-8D. Classical test theory suggests that instruments with more dimensions should be more responsive, assuming each dimension is responsive [59]. In this study, all dimensions of QLU-C10D were significantly sensitive, with all functional dimensions and three symptom dimensions demonstrating high responsiveness. We speculated that as patients' health improves through therapeutic or rehabilitative interventions, the restoration or enhancement of patients' fundamental functioning becomes more readily perceptible and observable, while it may take longer for patients to adapt or resolve these symptoms. Therefore, the functional dimensions of QLU-C10D are crucially significant. In FACT-8D, all dimensions except the social support problems exhibited significant responsiveness, which has a similar performance in FACT-G [60]. Social support problems, as external evaluative criteria, offer objectivity and short-term stability, enhancing patient-centricity but introducing some evaluation 'noise' in clinical trials; additionally, the negative correlation between social support problems and role functioning was an unexpected finding. One possible explanation is that individuals facing social support problems may need to develop stronger coping strategies or self-reliance. This personal development could enhance their performance in various roles. Moreover, the HM patients included in this study reported fewer social support problems, with over 40% reporting "no problem", likely due to China's predominantly family-oriented caregiving context for patients with cancer [61], potentially influencing its responsiveness outcomes.

This study has several other strengths. Firstly, it represents the first head-to-head comparison between QLU-C10D and FACT-8D, advancing knowledge for selecting cancer-specific MAUIs in oncology and enhancing the precision of economic assessment decisions in cancer care. It included a comprehensiveness assessment of measurement properties, including responsiveness, enabled by the longitudinal study design and the inclusion of a self-assessed health change anchor question. Additionally, this study applied the value set of the same country, any differences in measurement properties can be attributed to each MAUI's content and utility weights without confounding by country. This study also had some limitations. Firstly, QLU-C10D was administered first, followed by the FACT-8D, and this ordering could potentially impact the performance of both instruments. Secondly, due to the lack of a Chinese value set for FACT-8D (currently in development), Australian preference weights were used for all index scores in this study to maintain instrument stability. However, it is recommended to use the local country's value set when evaluating MAUIs, as applying the Australian value set may

influence the measurement properties of the instruments in this study. Upon completion of the Chinese FACT-8D value set development, further refinement of relevant research will be pursued.

Conclusion

The evidence in this study supports that both QLU-C10D and FACT-8D are valid cancer-specific MAUIs for evaluating patients with HM. However, both have their relative merits and scholars should take note of these slight differences in focus when selecting between the two measures.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13561-024-00560-0.

Supplementary Material 1

Acknowledgements
The authors are especially grateful to the participants who have suffered from cancer and have never given up. We also acknowledge the interviewers for helping with the data collection.

Author contributions
All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by HL and LC. The first draft of the manuscript was written by YC, and all authors commented on previous versions of the manuscript. Supervision and validation were provided by NL, HY, WH, and DC. Methodological input was provided by NL, LC, HY, MK, GK, and DC. Funding acquisition was managed by WH. All authors read and approved the final manuscript.

Funding
This work was supported by the National Natural Science Foundation of China (Grant No. 72274045, 71974048).

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics Statement and consent to participate
The protocol of this cohort study received approval from the Ethics Committee of Harbin Medical University (HMUIRB2023005). Informed consent was obtained from all individual participants included in the study.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 28 June 2024 / Accepted: 17 September 2024
Published online: 01 October 2024

References

1. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of Disease Study 2019. Lancet. 2020;396(10258):1204–22.
2. GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the global burden of Disease Study 2017. Lancet. 2018;392(10159):1736–88.
3. Chen S, Cao Z, Prettner K, Kuhn M, Yang J, Jiao L, et al. Estimates and projections of the Global Economic cost of 29 cancers in 204 countries and territories from 2020 to 2050. JAMA Oncol. 2023;9(4):465–72.
4. Maynou L, Cairns J. What is driving HTA decision-making? Evidence from cancer drug reimbursement decisions from 6 European countries. Health Policy. 2019;123(2):130–9.
5. Wang Y, Qiu T, Zhou J, Francois C, Toumi M. Which Criteria are considered and how are they evaluated in Health Technology assessments? A review of methodological guidelines used in Western and Asian countries. Appl Health Econ Health Policy. 2021;19(3):281–304.
6. Churruca K, Pomare C, Ellis LA, Long JC, Henderson SB, Murphy LED, et al. Patient-reported outcome measures (PROMs): a review of generic and condition-specific measures and a discussion of trends and issues. Health Expect. 2021;24(4):1015–24.
7. Dowie J. Decision validity should determine whether a generic or condition-specific HRQOL measure is used in health care decisions. Health Econ. 2002;11(1):1–8.
8. Teckle P, Peacock S, McTaggart-Cowan H, van der Hoek K, Chia S, Melosky B, et al. The ability of cancer-specific and generic preference-based instruments to discriminate across clinical and self-reported measures of cancer severities. Health Qual Life Outcomes. 2011;9:106.
9. Rowen D, Brazier J, Ara R, Azzabi Zouraq I. The role of Condition-Specific preference-based measures in Health Technology Assessment. Pharmacoeconomics. 2017;35(Suppl 1):33–41.
10. Rowen D, Young T, Brazier J, Gaugris S. Comparison of generic, condition-specific, and mapped health state utility values for multiple myeloma cancer. Value Health. 2012;15(8):1059–68.
11. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst. 1993;85(5):365–76.
12. King MT, Costa DS, Aaronson NK, Brazier JE, Cella DF, Fayers PM, et al. QLU-C10D: a health state classification system for a multi-attribute utility measure based on the EORTC QLQ-C30. Qual Life Res. 2016;25(3):625–36.
13. King MT, Norman R, Mercieca-Bebber R, Costa DSJ, McTaggart-Cowan H, Peacock S, et al. The Functional Assessment of Cancer Therapy eight dimension (FACT-8D), a Multi-attribute Utility Instrument Derived from the Cancer-Specific FACT-General (FACT-G) quality of Life Questionnaire: Development and Australian Value Set. Value Health. 2021;24(6):862–73.
14. Cella DF, Tulsky DS, Gray G, Sarafian B, Linn E, Bonomi A, et al. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. J Clin Oncol. 1993;11(3):570–9.
15. Giesinger JM, Efficace F, Aaronson N, Calvert M, Kyte D, Cottone F, et al. Past and current practice of patient-reported outcome measurement in Randomized Cancer clinical trials: a systematic review. Value Health. 2021;24(4):585–91.
16. King MT, Viney R, Simon Pickard A, Rowen D, Aaronson NK, Brazier JE, et al. Australian utility weights for the EORTC QLU-C10D, a Multi-attribute Utility Instrument Derived from the Cancer-Specific Quality of Life Questionnaire, EORTC QLQ-C30. PharmacoEconomics. 2018;36(2):225–38.
17. McTaggart-Cowan H, King MT, Norman R, Costa DSJ, Pickard AS, Regier DA, et al. The EORTC QLU-C10D: the Canadian valuation study and algorithm to Derive Cancer-Specific Utilities from the EORTC QLQ-C30. MDM Policy Pract. 2019;4(1):2381468319842532.
18. Revicki DA, King MT, Viney R, Pickard AS, Mercieca-Bebber R, Shaw JW, et al. United States Utility Algorithm for the EORTC QLU-C10D, a Multiattribute Utility Instrument based on a Cancer-specific quality-of-life instrument. Med Decis Mak. 2021;41(4):485–501.
19. Kemmler G, Gamper E, Nerich V, Norman R, Viney R, Holzner B, et al. German value sets for the EORTC QLU-C10D, a cancer-specific utility instrument based on the EORTC QLQ-C30. Qual Life Res. 2019;28(12):3197–211.
20. Norman R, Mercieca-Bebber R, Rowen D, Brazier JE, Cella D, Pickard AS, et al. U.K. utility weights for the EORTC QLU-C10D. Health Econ. 2019;28(12):1385–401.
21. Gamper EM, King MT, Norman R, Efficace F, Cottone F, Holzner B, et al. EORTC QLU-C10D value sets for Austria, Italy, and Poland. Qual Life Res. 2020;29(9):2485–95.
22. Nerich V, Gamper EM, Norman R, King M, Holzner B, Viney R, et al. French Value-Set of the QLU-C10D, a Cancer-specific utility measure derived from the QLQ-C30. Appl Health Econ Health Policy. 2021;19(2):191–202.

23. Finch AP, Gamper E, Norman R, Viney R, Holzner B, King M, et al. Estimation of an EORTC QLU-C10 value set for Spain using a Discrete Choice Experiment. PharmacoEconomics. 2021;39(9):1085–98.
24. Jansen F, Verdonck-de Leeuw IM, Gamper E, Norman R, Holzner B, King M, et al. Dutch utility weights for the EORTC cancer-specific utility instrument: the Dutch EORTC QLU-C10D. Qual Life Res. 2021;30(7):2009–19.
25. Xu RH, Wong EL, Luo N, Norman R, Lehmann J, Holzner B et al. The EORTC QLU-C10D: the Hong Kong valuation study. Eur J Health Econ. 2024;25(5):889-901.
26. Lehmann J, Rojas-Concha L, Petersen MA, Holzner B, Norman R, King MT et al. Danish value sets for the EORTC QLU-C10D utility instrument. Qual Life Res. 2024;33(3):831-841.
27. McTaggart-Cowan H, King MT, Norman R, Costa DSJ, Pickard AS, Viney R, et al. The FACT-8D, a new cancer-specific utility algorithm based on the Functional Assessment of Cancer therapies-General (FACT-G): a Canadian valuation study. Health Qual Life Outcomes. 2022;20(1):97.
28. King MT, Revicki DA, Norman R, Müller F, Viney RC, Pickard AS, et al. United States Value Set for the Functional Assessment of Cancer Therapy-General eight dimensions (FACT-8D), a Cancer-specific preference-based quality of Life Instrument. Pharmacoecon Open. 2024;8(1):49–63.
29. Norman R, Viney R, Aaronson NK, Brazier JE, Cella D, Costa DS, et al. Using a discrete choice experiment to value the QLU-C10D: feasibility and sensitivity to presentation format. Qual Life Res. 2016;25(3):637–49.
30. Pan CW, He JY, Zhu YB, Zhao CH, Luo N, Wang P. Comparison of EQ-5D-5L and EORTC QLU-C10D utilities in gastric cancer patients. Eur J Health Econ. 2023;24(6):885–93.
31. Bulamu NB, Vissapragada R, Chen G, Ratcliffe J, Mudge LA, Smithers BM, et al. Responsiveness and convergent validity of QLU-C10D and EQ-5D-3L in assessing short-term quality of life following esophagectomy. Health Qual Life Outcomes. 2021;19(1):233.
32. Gamper EM, Cottone F, Sommer K, Norman R, King M, Breccia M, et al. The EORTC QLU-C10D was more efficient in detecting clinical known group differences in myelodysplastic syndromes than the EQ-5D-3L. J Clin Epidemiol. 2021;137:31–44.
33. Klapproth CP, Fischer F, Rose M, Karsten MM. Health state utility differed systematically in breast cancer patients between the EORTC QLU-C10D and the PROMIS Preference score. J Clin Epidemiol. 2022;152:101–9.
34. Shaw JW, Bennett B, Trigg A, DeRosa M, Taylor F, Kiff C, et al. EQ-5D-3L, Mapping to the EQ-5D-5L, and European Organisation for Research and Treatment of Cancer Quality of Life Utility Measure-Core 10 Dimensions. Value Health. 2021;24(11):1651–9. A Comparison of Generic and Condition-Specific Preference-Based Measures Using Data From Nivolumab Trials.
35. Herdman M, Kerr C, Pavesi M, Garside J, Lloyd A, Cubi-Molla P, et al. Testing the validity and responsiveness of a new cancer-specific health utility measure (FACT-8D) in relapsed/refractory mantle cell lymphoma, and comparison to EQ-5D-5L. J Patient Rep Outcomes. 2020;4(1):22.
36. Extermann M, Overcash J, Lyman GH, Parr J, Balducci L. Comorbidity and functional status are independent in older cancer patients. J Clin Oncol. 1998;16(4):1582–7.
37. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SL, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med. 2002;32(6):959–76.
38. Kessler R, Mroczek D. Final Versions of our Non-Specific Psychological Distress Scale. 1994.
39. Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. J Health Soc Behav. 1997;38(1):21–37.
40. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med. 2001;33(5):337–43.
41. Yu H, Zeng X, Sui M, Liu R, Tan RL, Yang J, et al. A head-to-head comparison of measurement properties of the EQ-5D-3L and EQ-5D-5L in acute myeloid leukemia patients. Qual Life Res. 2021;30(3):855–66.
42. Luo N, Liu G, Li M, Guan H, Jin X, Rand-Hendriksen K. Estimating an EQ-5D-5L value set for China. Value Health. 2017;20(4):662–9.
43. Norman R, Mulhern B, Lancsar E, Lorgelly P, Ratcliffe J, Street D, et al. 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(4):427–38.
44. Andrews G, Slade T. Interpreting scores on the Kessler Psychological Distress Scale (K10). Aust N. Z J Public Health. 2001;25(6):494–7.
45. Zeng X, Sui M, Liu B, Yang H, Liu R, Tan RL, et al. Measurement Properties of the EQ-5D-5L and EQ-5D-3L in six commonly diagnosed cancers. Patient. 2021;14(2):209–22.
46. Welie AG, Stolk E, Mukuria C, Belay YB, Krahn MD, Sander B, et al. Reliability and validity of using EQ-5D-5L among healthy and adolescents with major mental health disorders in Ethiopia. Eur J Health Econ. 2022;23(7):1105–19.
47. Sun CY, Liu Y, Zhou LR, Wang MS, Zhao XM, Huang WD, et al. Comparison of EuroQol-5D-3L and short Form-6D utility scores in Family caregivers of Colorectal Cancer patients: a cross-sectional survey in China. Front Public Health. 2021;9:742332.
48. Cohen J. A power primer. Psychol Bull. 1992;112(1):155–9.
49. Liang MH, Larson MG, Cullen KE, Schwartz JA. Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. Arthritis Rheum. 1985;28(5):542–7.
50. Deyo RA, Centor RM. Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. J Chronic Dis. 1986;39(11):897–906.
51. Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PM. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. Qual Life Res. 2003;12(4):349–62.
52. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34–42.
53. Cicchetti DV. Guidelines, Criteria, and rules of Thumb for evaluating normed and standardized Assessment instruments in psychology. Psychol Assess. 1994;6(4):284–90.
54. Giavarina D. Understanding bland Altman analysis. Biochem Med (Zagreb). 2015;25(2):141–51.
55. Luckett T, King MT, Butow PN, Oguchi M, Rankin N, Price MA, et al. Choosing between the EORTC QLQ-C30 and FACT-G for measuring health-related quality of life in cancer clinical research: issues, evidence and recommendations. Ann Oncol. 2011;22(10):2179–90.
56. Gnanasakthy A, Barrett A, Evans E, D'Alessio D, Romano CD. A review of patient-reported outcomes labeling for Oncology drugs approved by the FDA and the EMA (2012–2016). Value Health. 2019;22(2):203–9.
57. Kluetz PG, Slagle A, Papadopoulos EJ, Johnson LL, Donoghue M, Kwitkowski VE, et al. Focusing on Core patient-reported outcomes in Cancer clinical trials: symptomatic adverse events, physical function, and Disease-related symptoms. Clin Cancer Res. 2016;22(7):1553–8.
58. Efficace F, Collins GS, Cottone F, Giesinger JM, Sommer K, Anota A, et al. Patient-reported outcomes as independent prognostic factors for Survival in Oncology: systematic review and Meta-analysis. Value Health. 2021;24(2):250–67.
59. Krebs DE. Measurement theory. Phys Ther. 1987;67(12):1834–9.
60. King MT, Bell ML, Costa D, Butow P, Oh B. The quality of Life Questionnaire Core 30 (QLQ-C30) and Functional Assessment of Cancer-General (FACT-G) differ in responsiveness, relative efficiency, and therefore required sample size. J Clin Epidemiol. 2014;67(1):100–7.
61. Lee J, Bell K. The impact of cancer on family relationships among Chinese patients. J Transcult Nurs. 2011;22(3):225–34.

Publisher's note