ELSEVIER

Journal of Clinical Epidemiology 187 (2025) 111965

Journal of Clinical Epidemiology

ORIGINAL RESEARCH

EORTC QLU-C10D was similarly valid and sensitive as EQ-5D-5L but more responsive to cancer patients' health deterioration

Luying Wang[a], Mihir Gandhi[b,c,d], Ravindran Kanesvaran[e], Mohamad Farid Bin Harunal Rashid[e], Dawn Qingqing Chong[e], Wen-Yee Chay[e], Richard Norman[f], Madeleine T. King[g], Nan Luo[h,*]

[a]*School of International Pharmaceutical Business, China Pharmaceutical University, Nanjing, China*

[b]*Biostatistics Department, Singapore Clinical Research Institute, Consortium for Clinical Research and Innovation Singapore, Singapore, Singapore*

[c]*Centre for Quantitative Medicine and Lien Centre for Palliative Care, Duke-NUS Medical School, Singapore, Singapore*

[d]*Tampere Center for Child, Adolescent, and Maternal Health Research: Global Health Group, Tampere University, Tampere, Finland*

[e]*Division of Medical Oncology, National Cancer Centre Singapore, Singapore, Singapore*

[f]*School of Population Health, Curtin University, Perth, Australia*

[g]*School of Psychology, The University of Sydney, Sydney, Australia*

[h]*Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore*

Accepted 26 August 2025; Published online 30 August 2025

Abstract

Objectives: The psychometric performance of the Quality of Life Utility-Core 10 Dimensions (QLU-C10D), a recently developed disease-specific preference-based measure (PBM) for patients with cancer, is not well understood yet. This study aimed to compare the construct validity, sensitivity, and responsiveness of QLU-C10D with that of 5-level EQ-5D (EQ-5D-5L), a generic PBM.

Study Design and Setting: We recruited patients with cancer from outpatient clinics of a tertiary cancer hospital then interviewed them face-to-face at two consecutive clinic visits using both QLU-C10D and EQ-5D-5L questionnaires. Construct validity was assessed through known-group comparisons and correlation analysis, agreement of utility scores was examined using intraclass correlation coefficient (ICC), and sensitivity and responsiveness compared using effect sizes (ESs) derived from known-group comparisons and standardized response means (SRMs) derived from within-group comparisons, respectively.

Results: We surveyed 626 patients; 280 of whom also completed the follow-up survey. Mean baseline QLU-C10D and EQ-5D-5L utility scores were 0.799 (SD: 0.224) and 0.916 (SD: 0.156), respectively. Both utility scores demonstrated known-groups and convergent/discriminant validity and their agreement was moderate (ICC: 0.60). EQ-5D-5L's ES was slightly higher than QLU-C10D's ES for cancer stage, whereas QLU-C10D's ES was higher for Eastern Cooperative Oncology Group. The SRM of QLU-C10D was considerably higher than that of EQ-5D-5L for deteriorated patients (−0.43 vs −0.02), whereas their SRMs were similar for the improved patients.

Conclusion: Although QLU-C10D and EQ-5D-5L utility scores are both valid in patients with cancer, they may not be used interchangeably, and the QLU-C10D scores appear to be more responsive to deterioration in health status than the EQ-5D-5Lscores.

Keywords: EORTC QLU-C10D; EQ-5D-5L; Validity; Responsiveness; Disease-specific preference-based measures

1. Introduction

The Quality of Life Utility-Core 10 Dimensions (QLU-C10D) [1,2] is a preference-based measure (PBM) developed by the European Organization for Research and Treatment of Cancer (EORTC) for converting data collected using the Quality of Life Questionnaire-Core 30 (QLQ-C30) [3], a profile-based measure of health-related quality of life (HRQL) of patients with cancer, into utility scores that can be used to calculate quality-adjusted life years (QALYs) in economic evaluations. The QLU-C10D describes patients' health status in terms of ten domains. As the QLQ-C30 has been widely used in clinical studies of patients with cancer and treatment, and there are 18

Funding: This work was supported by the Singapore Cancer Society and the National Medical Research Council Singapore (NMRC/HSRG/0084/2017).

* Corresponding author. Saw Swee Hock School of Public Health, National University of Singapore, Tahir Foundation Building, 12 Science Drive 2, Singapore, Singapore 117549.

E-mail address: ephln@nus.edu.sg (N. Luo).

https://doi.org/10.1016/j.jclinepi.2025.111965

0895-4356/

What is new?

Key findings

- The EQ-5D-5L appeared to discriminate between cancer stages better than Quality of Life Utility-Core 10 Dimensions (QLU-C10D), whereas QLU-C10D appeared to better discriminate between Eastern Cancer Oncology Group performance statuses.
- QLU-C10D seemed to be more able to detect health deterioration of patients with cancer than EQ-5D-5L, although their ability to detect improvement in health was similar.

What this adds to what is known?

- Cancer-specific preference-weighted quality of life measures are likely to be more responsive to change in health outcomes than generic preference-weighted quality of life measures.

What is the implication and what should change now?

- More research is warranted to guide the selection between generic and disease-specific preference-weighted quality of life measures for patients with cancer.

country-specific value sets available [4−7], QLU-C10D represents an accessible tool for deriving preference-based quality of life data.

In contemporary health technology assessment, generic PBMs are preferred in many jurisdictions [8−10]. However, disease-specific PBMs such as QLU-C10D may be valuable when there are concerns about the adequacy of generic PBMs, such as EQ-5D, for patients with cancer [11,12]. Indeed, several studies have shown that QLU-C10D was more sensitive than the 3-level EQ-5D (EQ-5D-3L). For instance, Pilz et al found that QLU-C10D was more responsive and sensitive than EQ-5D-3L to performance status, working status, stage, and history classifications [13]. However, comparative studies of QLU-C10D and the 5-level EQ-5D (EQ-5D-5L) have shown inconsistent results and often lacked responsiveness assessment [14,15]. This is a significant knowledge gap as users are switching from EQ-5D-3L to EQ-5D-5L because EQ-5D-5L has shown better measurement properties [16−18].

Therefore, the aim of this study was to compare the psychometric properties of QLU-C10D and EQ-5D-5L in patients with cancer. The specific objective was to compare their construct validity, sensitivity, and responsiveness.

2. Methods

2.1. Study design and data collection

We recruited a consecutive sample of patients with cancer visiting specialist outpatient clinics at National Cancer Center Singapore, a tertiary cancer care provider to a multicultural and multiethnic Asian population of 5 million living in Singapore. Patients were included in the study if they met the following criteria: aged 21 years or older (the legal age of adulthood in Singapore), Singapore citizens or permanent residents, able to read and communicate in either English or Chinese, and capable of understanding the survey questions and providing informed consent. Face-to-face interviews were conducted immediately after recruitment to collect EQ-5D-5L and EORTC QLQ-C30 data, as well as demographic and clinical data which were also retrieved from medical records. In addition, all patients were invited to a follow-up interview at their next visit to complete EQ-5D-5L and QLQ-C30 questionnaires again, along with a question assessing general health status using a 5-point Likert scale (excellent, very good, good, fair, and poor). Ethics approval for the study was obtained from the Centralized Institutional Review Board (CIRB Ref: 2018/2345).

2.2. Measures

2.2.1. EQ-5D-5L

It comprises a descriptive system and a visual analog scale ranging from 0 to 100. The descriptive system is for describing respondents' health in five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. It comprises five sets of statements, each set defining five different functioning levels (ie, no, slight, moderate, severe, and extreme problems) for a different dimension. With this design, the descriptive system defines a total of 3125 multidimensional health states. Numerous country-specific value sets are available to assign EQ-5D-5L health states utility scores, which are anchored by 1 (full health) and 0 (dead), with negative scores for worse-than-dead health states. In this study, EQ-5D-5L utility scores were calculated using a Singaporean value set (ranging from −0.851 to 1), which was estimated using preference data collected from 500 Singaporean residents using a time trade-off method [19].

2.2.2. QLU-C10D

Its descriptive system includes ten domains: physical functioning, role functioning, social functioning, emotional functioning, pain, fatigue, sleep, appetite, nausea, and bowel problems. Each domain is described with 4 levels; the descriptive system therefore defines a total of 1,048,576 unique health states. A person's QLU-C10D health state can be determined based on their responses to QLQ-C30 items [1]. This study used a local set of QLU-

C10D scores which range from −0.821 (worst health state) to 1 (perfect health) [6].

2.3. Statistical analysis

We described patient characteristics at both baseline and follow-up surveys using means and SDs or proportions. We assessed the known-group and convergent/discriminant validity of EQ-5D-5L and QLU-C10D separately and their agreement. Lastly, we assessed the sensitivity and responsiveness of the two instruments.

We assessed the known-group validity of QLU-C10D and EQ-5D-5L by testing four a priori hypotheses: mean utility scores are higher (i) in patients in a cancer stage of I or II compared to those in stage III or IV; (ii) in patients without chronic comorbidities compared to patients with comorbidities; (iii) in patients who reported excellent, very good or good health compared to patients who reported fair or poor health; and (iv) in patients with an Eastern Cooperative Oncology Group (ECOG) score of 0 compared to patients with an ECOG score of 1−4. The aforementioned known-group hypotheses have been commonly used in previous psychometric studies of HRQL instruments for assessing patients with cancer [13,17,20]. T-statistics were used to test the differences between the groups, with *P* values less than 0.05 indicating statistically significant differences.

We used a multitrait-multimethod matrix of Spearman's rank correlation coefficients (r) to test the convergent and discriminant validity of the two instruments [21]. In this analysis, the methods were the QLU-C10D and EQ-5D-5L, and the traits were the HRQL dimensions. We arranged the matrix such that correlations between same or similar dimensions of the two instruments fall along its diagonal. On either side of the diagonal are the heterotrait-heteromethod triangles. We hypothesized that the correlation between similar health dimensions (eg, QLU-C10D physical functioning and EQ-5D-5L mobility, QLU-C10D role functioning and EQ-5D-5L usual activities, QLU-C10D social functioning and EQ-5D-5L usual activities, QLU-C10D emotional functioning and EQ-5D-5L anxiety or depression, QLU-C10D pain and EQ-5D-5L pain, or discomfort) would be moderate to strong (ie, convergent validity testing) and higher than those between less obviously related dimensions which would be weak to moderate only (ie, discriminant validity testing). According to Cohen's rules, Spearman's correlation coefficient from 0.1 to 0.3 was considered a "weak effect", a correlation coefficient ranging from 0.3 to 0.5 was considered a "moderate effect", and a correlation coefficient above 0.5 was considered a "strong effect" [22].

The degree of agreement between the QLU-C10D and EQ-5D-5L scores was assessed using the intraclass correlation coefficient (ICC) and examined using the Bland−Altman plot. ICC values ranged from 0 to 0.49, 0.50−0.75, 0.76−0.90, and 0.91−1 were considered as poor, moderate, good, and excellent, respectively [23].

The sensitivity of the two utility scores to discriminate between groups known to differ with each other was assessed using Cohen's effect size (ES) [22] and relative efficiency (RE) as the ratio of F-statistics in the analysis of variance tests of QLU-C10D over EQ-5D-5L [24]. Bias corrected bootstrapping was used to estimate the 95% CIs for ES and RE estimates. Larger ES indicates better known-group validity and sensitivity, and RE greater than 1 indicates that QLU-C10D is more efficient than EQ-5D-5L and vice versa.

Responsiveness of QLU-C10D and EQ-5D-5L scores was examined by analyzing their changes in patients with improved and deteriorated health status separately. Improvements and deteriorations were defined as reporting better or worse health with the general health question at the follow-up visit compared to the baseline visit. Standardized ES (SES), defined as the difference between the mean score at baseline and the mean at follow-up divided by the baseline SD, standardized response mean (SRM), defined as the mean change between baseline and the follow-up scores divided by the SD of change scores, and RE were calculated, and the 95% CIs of SES, SRM, and RE were calculated using the bias corrected bootstrap method, respectively. Larger SES or SRM indicate higher responsiveness to changes in health status [25].

Data were analyzed using STATA SE (standard edition 18.0).

3. Results

A total of 626 patients with cancer completed the baseline survey. The mean age of the participants was 59.5 years. The majority of the patients were female (58%) and Chinese (84%). Breast cancer (31%) was the most prevalent cancer type among the patients, followed by digestive or gastrointestinal (21%) and head and neck cancer (14%). The proportion of patients with stage IV cancer (30%) was slightly higher than that of stage 0/I, II, and III cancer (21%−22%). Most patients were categorized as ECOG grade 0 (78%) or grade 1 (21%). The full demographic and clinical characteristics of the sample are summarized in Table 1.

Both QLU-C10D and EQ-5D-5L scores were not normally distributed with 21.9% and 46.6% of participants having the highest scores of 1 with QLU-C10D and EQ-5D-5L, respectively (Fig S1). The mean (SD) QLU-C10D and EQ-5D-5L scores were 0.799 (0.224) and 0.916 (0.156) at baseline and 0.776 (0.232) and 0.920 (0.124) at follow-up, respectively. Mean scores for patient subgroups are provided in Table S1. Overall, the EQ-5D-5L scores were higher compared to the QLU-C10D scores. The agreement between the QLU-C10D and the EQ-5D-5L scores was moderate (ICC = 0.60; 95% CI:0.55−0.65). The Bland−Altman plot (Fig) shows that the difference

Table 1. Characteristics of included cancer patients

Characteristic	Baseline n (%)	Follow-up n (%)
Gender		
Male	263 (42)	122 (44)
Female	363 (58)	158 (56)
Age (y)		
20–39	33 (5)	13 (5)
40–59	256 (41)	121 (43)
≥60	337 (54)	146 (52)
Ethnic		
Chinese	527 (84)	231 (83)
Malay	50 (8)	26 (9)
Indian/others	49 (8)	23 (8)
Marital status		
Never married	123 (20)	58 (21)
Married	409 (65)	173 (62)
Divorced	50 (8)	26 (9)
Widowed	37 (6)	19 (7)
Refused to answer	7 (1)	4 (1)
Education		
Below secondary	124 (20)	54 (19)
Secondary	273 (44)	124 (44)
Postsecondary	47 (8)	18 (6)
Diploma and professional qualification	83 (13)	42 (15)
University and above	99 (16)	42 (15)
Individual monthly income (SGD)		
<2000	215 (34)	97 (35)
2000–3999	109 (17)	48 (17)
4000–5999	54 (9)	29 (10)
6000–9999	62 (10)	29 (10)
≥10,000	33 (5)	9 (3)
Do not know/decline to disclose	153 (24)	68 (24)
Cancer system		
Breast	191 (31)	79 (28)
Digestive/gastrointestinal	131 (21)	69 (25)
Head and neck	87 (14)	33 (12)
Respiratory/thoracic	38 (6)	21 (8)
Genitourinary	50 (8)	16 (6)
Others	129 (21)	62 (22)
Cancer stage		
0/I	129 (21)	46 (16)
II	132 (21)	44 (16)
III	139 (22)	65 (23)
IV	186 (30)	108 (39)
Unknown	40 (6)	17 (6)
ECOG		
Grade 0	394 (78)	163 (68)
Grade 1	105 (21)	74 (31)
Grade 2–4	7 (1)	3 (1)

SGD, Singapore dollar; ECOG, Eastern Cooperative Oncology Group.

Table 2. Proportion of reported health problems dimension at baseline (N = 626)

	%				
Health dimension	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**
EQ-5D-5L					
Mobility (MO)	85	12	3	1	0
Self-care (SC)	97	3	1	0	0
Usual activities (UAs)	88	9	2	1	0
Pain/discomfort (PD)	59	33	6	1	1
Anxiety/depression (AD)	77	19	3	1	1
QLU-C10D					
Physical functioning (PF)	58	33	8	1	
Role functioning (RF)	79	116	4	2	
Social functioning (SF)	73	19	6	2	
Emotional functioning (EF)	76	20	3	1	
Pain (PA)	56	34	7	3	
Fatigue (FA)	40	47	11	3	
Sleep disturbance (SL)	56	29	9	5	
Appetite loss (AP)	77	14	7	2	
Nausea (NA)	86	10	2	1	
Bowel problems (BO)	68	25	6	2	

QLU-C10D, Quality of Life Utility-Core 10 Dimensions.

Table 3. Known-group validity and sensitivity of EQ-5D-5L and QLU-10D at baseline

Instruments	Patient groups	N	Utility score mean (SD)	T statistics (P value)	ES (95% CIs)	RE (95% CIs)
Cancer stage						
EQ-5D-5L	I/II	261	0.944 (0.096)	3.79 (<0.01)	0.33 (0.18, 0.45)	reference
	III/IV	325	0.893 (0.188)			
QLU-C10D	I/II	261	0.835 (0.201)	3.48 (<0.01)	0.29 (0.13,0.44)	0.77 (0.22,1.76)
	III/IV	325	0.770 (0.238)			
Chronic disease						
EQ-5D-5L	With	465	0.906 (0.172)	2.61 (<0.01)	0.24 (0.12, 0.35)	reference
	Without	161	0.943 (0.091)			
QLU-C10D	With	465	0.785 (0.230)	2.58 (0.01)	0.24 (0.05, 0.38)	0.97 (0.17,3.25)
	Without	161	0.838 (0.200)			
General health						
EQ-5D-5L	Excellent/very good/good	412	0.944 (0.111)	6.60 (<0.01)	0.56 (0.37, 0.73)	reference
	Fair/poor	214	0.861 (0.207)			
QLU-C10D	Excellent/very good/good	412	0.840 (0.187)	6.66 (<0.01)	0.56 (0.38,0.73)	1.02 (0.56,1.94)
	Fair/poor	214	0.719 (0.263)			
ECOG						
EQ-5D-5L	Grade 0	394	0.924 (0.150)	2.43 (0.02)	0.26 (0.13,0.51)	reference
	Grade 1–4	112	0.882 (0.199)			
QLU-C10D	Grade 0	394	0.813 (0.220)	3.09 (<0.01)	0.33 (0.10,0.57)	1.62 (0.40, 18.3)
	Grade 1–4	112	0.738 (0.244)			

ECOG, Eastern Cooperative Oncology Group; ES, effect size; RE, relative efficiency; QLU-C10D, Quality of Life Utility-Core 10 Dimensions.

Table 4. Multitraits-multimethods matrix of EQ-5D-5L and QLU-C10D at baseline

	EQ-5D-5L					QLU-C10D								
Instruments and dimensions	**Mobility (MO)**	**Usual activity (UA)**	**Anxiety/ depression (AD)**	**Pain/ discomfort (PD)**	**Self-care (SC)**	**PF**	**RF**	**SF**	**EF**	**PA**	**FA**	**SL**	**AP**	**NA**
QLU-C10D														
Physical functioning (PF)	*0.40*	0.31	0.14	0.27	0.16									
Role functioning (RF)	0.30	*0.49*	0.28	0.33	0.28	0.33								
Social functioning (SF)	0.22	*0.38*	0.34	0.35	0.23	0.24	0.41							
Emotional functioning (EF)	0.15	0.17	*0.46*	0.25	0.16	0.17	0.25	0.31						
Pain (PA)	0.25	0.26	0.15	***0.64***	0.13	0.29	0.33	0.34	0.21					
Fatigue (FA)	0.21	0.31	0.27	0.41	0.17	0.37	0.35	0.45	0.33	0.37				
Sleep disturbance (SL)	0.16	0.16	0.17	0.27	0.11	0.22	0.18	0.25	0.25	0.32	0.31			
Appetite loss (AP)	0.16	0.19	0.2	0.23	0.05	0.28	0.30	0.32	0.17	0.19	0.33	0.28		
Nausea (NA)	0.09	0.14	0.19	0.24	0.04	0.17	0.22	0.27	0.20	0.21	0.27	0.21	0.40	
Bowel problems (BO)	0.18	0.18	0.07	0.14	0.14	0.23	0.16	0.19	0.13	0.17	0.23	0.20	0.16	0.18
EQ-5D-5L														
UA	0.40													
AD	0.19	0.21												
PD	0.36	0.33	0.28											
SC	0.35	0.41	0.17	0.20										

QLU-C10D, Quality of Life Utility-Core 10 Dimensions.
The validity diagonal is marked with italic. Spearman correlation coefficients were used and correlation coefficient greater than 0.5 was regarded as strong correlation and bolded.

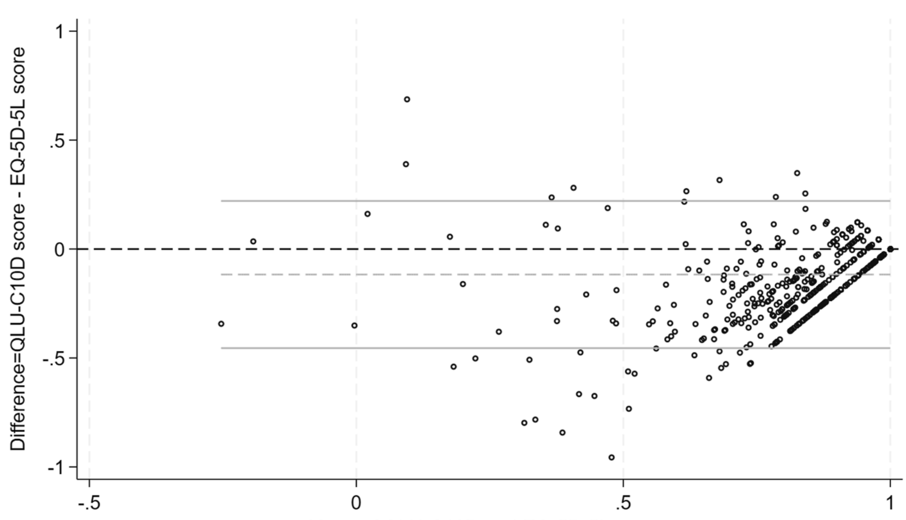

Figure 1. Bland−Altman plot of QLU-C10D and EQ-5D-5L scores. The two solid lines represent 95% CI of the difference between the two scores whereas the dashed line represents the mean of the difference. QLU-C10D, Quality of Life Utility-Core 10 Dimensions.

between the two scores became smaller as the scores approached 1.

The most prevalent health problems reported with the QLU-C10D were fatigue (60%) and pain (44%), whereas pain/discomfort (42%) and anxiety/depression (23%) were the most frequently reported health problems with the EQ-5D-5L; the least reported health problems with QLU-C10D and EQ-5D-5L were in the dimension of nausea (16%) and self-care (3%), respectively (Table 2).

As shown in Table 3, both QLU-C10D and EQ-5D-5L scores differed as expected between subgroups of patients defined by cancer stage, chronic disease, general health, and ECOG. The ES of the EQ-5D-5L was slightly higher than that of the QLU-C10D for cancer stage (0.33 vs 0.29), but lower for ECOG subgroups (0.26 vs 0.33). The ESs of the two instruments were nearly identical for chronic disease and general health subgroups. RE estimates derived from t-statistic values reflected these results, but as all RE CIs included one, we inferred that the differences in sensitivity were not statistically significant.

The multitrait-multimethod correlation matrix is presented in Table 4. The correlation coefficients between similar dimensions ranged from 0.38 to 0.64, suggesting moderate to strong correlation. The correlation coefficients between other dimensions of QLU-C10D and EQ-5D-5L ranged from very weak to moderate (0.04−0.41). The correlation coefficients between dimensions in the same instrument ranged from weak to moderate (0.17−0.41). The correlation coefficient between QLU-C10D and EQ-5D-5L scores was strong (0.64).

A total of 280 patients completed the follow-up survey. Compared to the baseline sample, the follow-up sample comprised higher proportion of patients with stage IV cancer (39%) and patients with ECOG grade 1 (31%). Among those patients, 89 (32%) reported improved general health, 128 (46%) felt the same, and 63 (22%) reported worse general health. The results of the responsiveness analysis are listed in Table 5. The mean change (SD) in the QLU-C10D score from baseline was 0.070 (0.208) for the improved group and −0.119 (0.279) for the deteriorated group, whereas the mean change (SD) of the EQ-5D-5L score was 0.055 (0.171) for the improved and −0.003 (0.170) for the deteriorated group. The SES (0.28 vs 0.31) and SRM (0.34 vs 0.32) values of QLU-C10D and EQ-5D-5L were similar, as were their SRM CIs. However, SES (−0.53 vs −0.01) and SRM (−0.43 vs −0.02) values of QLU-C10D were considerably higher than those of EQ-5D-5L for the deteriorated group (Table 5). The CI of the QLU-C10D SRM did not include zero, whereas the CI of the EQ-5D-5L SRM did include zero. We therefore inferred that two measures were similarly responsive to improvement in health, whereas for deterioration in health, the QLU-C10D was responsive but the EQ-5D-5L was not.

4. Discussion

To our knowledge, this study is the first head-to-head comparison of the psychometric properties of QLU-C10D and EQ-5D-5L scored using value sets developed with recommended valuation protocols and health preferences from the same general population. Such a design isolated the possible confounding effects of valuation impression and culture on the performance of these two preference-based instruments.

Overall, the correlations between EQ-5D-5L and QLU-C10D dimensions supported their convergent and discriminant validity. However, there were some unexpected results. For instance, the correlation between QLU-C10D physical functioning and EQ-5D-5L mobility (0.40) was lower than correlations reported in other studies (0.58−0.68) [13,15,20,26]. This discrepancy may be due to the relatively healthier status of our patient sample which led to reduced data variability on the dimension scales. For instance, patients in the current study reported higher health utility scores (EQ-5D-5L: 0.92 vs 0.85 in the study by Bulamu et al) and a greater proportion were classified as ECOG 0 (78% vs 46% in the study by Gamper et al). The aforementioned correlation was not higher than the correlation between EQ-5D-5L mobility and usual activities (0.40), violating the expectation that correlations between similar dimensions across instruments were not consistently higher than correlations between different dimensions within the same instrument. This unexpected result could also be due to the overlap in constructs targeted by mobility and usual activities dimensions in EQ-5D-5L.

The known-group comparison results provided clear evidence for the construct validity of both instruments. Compared to the QLU-C10D, the EQ-5D-5L showed slightly higher sensitivity to cross-sectional difference in subgroups of patients known to differ in cancer stage. Conversely, QLU-C10D performed better in discriminating between ECOG subgroups. With the caveat that CIs indicated these slight differences were not statistically significant, we note that other studies have revealed similar patterns. Gamper et al showed that QLU-C10D had higher discriminatory power for ECOG groups compared to EQ-5D-5L [20]; however, the CIs for the ES overlapped, and CIs for the RE were not reported. Pan et al also found that EQ-5D-5L was more sensitive to cancer stages compared to the QLU-C10D, but they did not report the CIs for either ES or RE [13,14]. To compare the sensitivity of the two measurement tools, we recommend that future studies report CIs alongside point estimates of the relative efficiency measures used, and that a meta-analysis is conducted to provide a robust conclusion when the amount of evidence is sufficient.

A key finding of this study is that QLU-C10D showed significantly better responsiveness to patient-reported health deterioration compared to EQ-5D-5L. Bulamu et al

Table 5. Responsiveness to self-reported general health of EQ-5D-5L and QLU-C10D

		Reported general health states			
		Improved		Deteriorated	
		$N = 89$		$N = 63$	
Measure	Instrument	Mean	SD	Mean	SD
Baseline index score mean (SD)	EQ-5D-5L	0.902	0.179	0.869	0.218
	QLU-C10D	0.756	0.246	0.778	0.226
Follow-up index score mean (SD)	EQ-5D-5L	0.957	0.061	0.866	0.176
	QLU-C10D	0.826	0.162	0.659	0.317
Mean change of index score from baseline (SD)	EQ-5D-5L	0.055	0.171	−0.003	0.170
	QLU-C10D	0.070	0.208	−0.119	0.279
		Mean	95% Bootstrap CI	Mean	95% Bootstrap CI
Standardized effect size (SES)	EQ-5D-5L	0.31	(0.18,0.41)	−0.01	(−0.34,0.15)
	QLU-C10D	0.28	(0.13,0.43)	−0.53	(−0.88,−0.25)
Standardized response mean (SRM)	EQ-5D-5L	0.32	(0.18,0.44)	−0.02	(−0.26,0.24)
	QLU-C10D	0.34	(0.16,0.49)	−0.43	(−0.61,−0.24)
Relative efficiency (RE)	EQ-5D-5L	Reference		Reference	
	QLU-C10D	1.10	(0.38,2.54)	535.00	(3.20, 9.9×10^5)

[26] and Shaw et al [27] found similar results, although Bulamu used EQ-5D-3L and Shaw mapped EQ-5D-5L responses to EQ-5D-3L and used EQ-5D-3L value sets to calculate EQ-5D-5L scores. To investigate possible reasons for the differential responsiveness of EQ-5D-5L and QLU-C10D, as a post hoc analysis, we calculated the proportion of deterioration in each health dimension among the patients who reported deteriorated health status. The QLU-C10D health dimensions with the highest proportions deteriorating were social functioning (24%), bowel problems (19%), role functioning (18%), and sleep disturbance (16%), whereas the EQ-5D-5L dimensions with the highest proportions were pain/discomfort (22%), anxiety/depression (18%), and mobility and usual activities (13% for both). Therefore, the better responsiveness of QLU-C10D is likely to be driven by its social and treatment-related dimensions which are absent from EQ-5D-5L. This observation supports the "bolt-on" strategy for EQ-5D-5L [28], which means that adding supplementary dimensions to the instrument to increase its sensitivity to the impact of health conditions and treatments on HRQL to which the standard EQ-5D-5L is not sufficiently sensitive. Although bolt-on has been proven to increase the sensitivity and responsiveness of EQ-5D in some populations [29,30], its usefulness in cancer outcomes research has not been explored. Therefore, it would be interesting to test whether existing bolt-ons such as sleep and fatigue, and new bolt-ons such as bowel problems, can improve the psychometric performance of EQ-5D in patients with cancer.

We found that QLU-C10D and EQ-5D-5L utility scores were only in moderate agreement, with QLU-C10D scores being lower than EQ-5D-5L scores, consistent with the results reported by Döhmen et al [15]. The lower QLU-C10D scores are likely due to its numerous symptom dimensions which are able to capture common symptoms of patients with cancer who are undergoing active treatment, such as fatigue, sleep disturbance, and appetite loss, each of which adds a decrement in utility to the QLU-C10D score. Given the divergence in utility scores, cost-effectiveness analysis (CEA) based on QALYs derived from QLU-C10D and EQ-5D-5L may lead to different conclusions. Although it is hard to predict how CEA results could be affected when EQ-5D-5L or QLU-C10D utility scores are used to calculate QALYs, Wailoo et al illustrated that replacing EQ-5D-3L scores with EQ-5D-5L scores would result in substantial, yet unpredictable changes in CEA results [31]. Therefore, future studies are needed to study how the use of the two instruments may affect CEAs of anticancer treatment or interventions.

This study had several strengths and some limitations. Strengths included a comprehensive range of psychometric analyses, large sample sizes, and estimates of CIs for sensitivity and responsiveness. Limitations included that we recruited outpatients in Singapore, which may limit the findings' generalizability, and that EQ-5D-5L and EORTC QLQ-C30 data were collected via face-to-face interviews rather than self-completion, which may have induced some response bias due to patients being embarrassed or inhibited to report how they really felt. It is generally recommended that patients should self-complete patient-reported questionnaires for this reason. Notably, the utility scores and functional status (assessed using ECOG) were relatively high among the recruited patients with cancer, resulting in reduced sample variability. Such a relatively healthy sample is not optimal for psychometric testing as it may lead to underestimated psychometric properties, such as sensitivity to health improvements.

5. Conclusion

Although QLU-C10D and EQ-5D-5L are both valid and sensitive PBMs for assessing the HRQL of patients with cancer, QLU-C10D appears to be more responsive to self-perceived health deterioration. Considering this and the systematic differences in scores between these two instruments, they may not be used interchangeably and should be selected with discretion.

CRediT authorship contribution statement

Luying Wang: Writing – review & editing, Writing – original draft, Visualization, Validation, Formal analysis, Data curation. **Mihir Gandhi:** Writing – review & editing, Resources, Investigation, Formal analysis, Data curation. **Ravindran Kanesvaran:** Writing – review & editing, Supervision, Resources, Investigation. **Mohamad Farid Bin Harunal Rashid:** Supervision, Resources. **Dawn Qingqing Chong:** Supervision, Resources. **Wen-Yee Chay:** Supervision, Resources. **Richard Norman:** Writing – review & editing, Methodology. **Madeleine T. King:** Writing – review & editing, Visualization, Validation, Methodology. **Nan Luo:** Writing – review & editing, Visualization, Validation, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

N.L. reports financial support was provided by Singapore Cancer Society and the National Medical Research Council Singapore. There are no competing interests for any other author.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jclinepi.2025.111965.

Data availability

Data will be made available on request.

References

[1] 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. https://doi.org/10.1007/s11136-015-1217-y.

[2] 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. https://doi.org/10.1007/s11136-015-1115-3.

[3] 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. https://doi.org/10.1093/jnci/85.5.365.

[4] Cao Y, Xu J, Norman R, King MT, Kemmler G, Huang W, et al. Chinese utility weights for the EORTC cancer-specific utility instrument QLU-c10d. Qual Life Res 2024;33(12):3335–49. https://doi.org/10.1007/s11136-024-03776-z.

[5] Rohde G, Lehmann J, Pilz MJ, Rojas-Concha L, Holzner B, King MT, et al. Norwegian and swedish value sets for the EORTC QLU-c10d utility instrument. Qual Life Res 2025;34(2):429–43. https://doi.org/10.1007/s11136-024-03824-8.

[6] Gandhi M, Kanesvaran R, Rashid M, Chong DQ, Chay WY, Tan RL, et al. Valuation of the EORTC quality of life utility core 10 dimensions (QLU-c10d) in a multi-ethnic asian setting: how does having cancer matter? Pharmacoeconomics 2024;42(12):1413–25. https://doi.org/10.1007/s40273-024-01432-5.

[7] Seyringer S, Pilz MJ, Jansen F, Buttner M, King MT, Norman R, et al. Cancer-specific utility instrument for health economic evaluations: a synopsis of the EORTC QLU-c10d user manual and current validity evidence. Eur J Cancer 2025;217:115235. https://doi.org/10.1016/j.ejca.2025.115235.

[8] Nhci. Guideline for economic evaluations in healthcare (2024 version). 2024. Available at: https://english.zorginstituutnederland.nl/about-us/publications/reports/2024/01/16/guideline-for-economic-evaluations-in-healthcare. Accessed June 7, 2024.

[9] Cadth. Guidelines for the economic evaluation of health technologies: Canada - 4th edition. 2021. Available at: https://www.cadth.ca/guidelines-economic-evaluation-health-technologies-canada-4th-edition. Accessed June 7, 2024.

[10] Nice. NICE health technology evaluations: the manual. 2023. Available at: https://www.nice.org.uk/process/pmg36. Accessed June 7, 2024.

[11] Gibson A, Longworth L, Bennett B, Pickard AS, Shaw JW. Assessing the content validity of preference-based measures in cancer. Value Health 2024;27(1):70–8. https://doi.org/10.1016/j.jval.2023.10.006.

[12] Shah KK, Bennett B, Lenny A, Longworth L, Brazier JE, Oppe M, et al. Adapting preference-based utility measures to capture the impact of cancer treatment-related symptoms. Eur J Health Econ 2021;22(8):1301–9. https://doi.org/10.1007/s10198-021-01337-6.

[13] Pilz MJ, Seyringer S, Hallsson LR, Bottomley A, Jansen F, King MT, et al. The EORTC QLU-c10d is a valid cancer-specific preference-based measure for cost-utility and health technology assessment in the netherlands. Eur J Health Econ 2024;25(9):1539–55. https://doi.org/10.1007/s10198-024-01670-6.

[14] 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. https://doi.org/10.1007/s10198-022-01523-0.

[15] Dohmen A, Obbarius A, Kock M, Nolte S, Sidey-Gibbons CJ, Valderas JM, et al. The EORTC QLU-c10d distinguished better between cancer patients and the general population than PROPr and EQ-5d-5l in a cross-sectional study. J Clin Epidemiol 2025;177:111592. https://doi.org/10.1016/j.jclinepi.2024.111592.

[16] 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. https://doi.org/10.1007/s40271-020-00466-z.

[17] 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. https://doi.org/10.1007/s11136-020-02644-w.

[18] Kim SH, Kim HJ, Lee SI, Jo MW. Comparing the psychometric properties of the EQ-5d-3l and EQ-5d-5l in cancer patients in korea. Qual Life Res 2012;21(6):1065−73. https://doi.org/10.1007/s11136-011-0018-1.

[19] Luo N, Thakumar AV, Cheng LJ, Yang Z, Rand K, Cheung YB, et al. Developing an EQ-5d-5l value set for Singapore. Pharmacoeconomics 2025. https://doi.org/10.1007/s40273-025-01519-7.

[20] 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. https://doi.org/10.1016/j.jclinepi.2021.03.015.

[21] Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychol Bull 1959;56(2):81−105.

[22] Cohen J. Statistical power analysis for the behavioral sciences. New York: Routledge; 1988.

[23] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15(2):155−63. https://doi.org/10.1016/j.jcm.2016.02.012.

[24] Snecdecor GW, Cochran WG. Statistical methods. 8th ed. Ames, Lowa: Wiley-Blackwell; 1989.

[25] Gandhi M, Ang M, Teo K, Wong CW, Wei YC, Tan RL, et al. EQ-5d-5l is more responsive than EQ-5d-3l to treatment benefit of cataract surgery. Patient 2019;12(4):383−92. https://doi.org/10.1007/s40271-018-00354-7.

[26] 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. https://doi.org/10.1186/s12955-021-01867-w.

[27] Shaw JW, Bennett B, Trigg A, Derosa M, Taylor F, Kiff C, et al. A comparison of generic and condition-specific preference-based measures using data from nivolumab trials: 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. https://doi.org/10.1016/j.jval.2021.05.022.

[28] Longworth L, Yang Y, Young T, Mulhern B, Hernandez AM, Mukuria C, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. Health Technol Assess 2014;18(9):1−224. https://doi.org/10.3310/hta18090.

[29] Finch AP, Brazier J, Mukuria C. Selecting bolt-on dimensions for the EQ-5d: testing the impact of hearing, sleep, cognition, energy, and relationships on preferences using pairwise choices. Med Decis Making 2021;41(1):89−99. https://doi.org/10.1177/0272989X20969686.

[30] Angyal MM, Janssen MF, Lakatos PL, Brodszky V, Rencz F. The added value of the cognition, dining, gastrointestinal problems, sleep and tiredness bolt-on dimensions to the EQ-5d-5l in patients with coeliac disease. Eur J Health Econ 2024;26:473−85. https://doi.org/10.1007/s10198-024-01719-6.

[31] Wailoo A, Alava MH, Pudney S, Barton G, O'Dwyer J, Gomes M, et al. An international comparison of EQ-5d-5l and EQ-5d-3l for use in cost-effectiveness analysis. Value Health 2021;24(4):568−74. https://doi.org/10.1016/j.jval.2020.11.012.