

Responsiveness and Convergent Validity of QLU-C10D and EQ-5D-3L in Assessing Short-Term Quality of Life Following Esophagectomy

Norma B. Bulamu (✉ norma.bulamu@flinders.edu.au)

Flinders University, College of Medicine and Public Health <https://orcid.org/0000-0003-4822-4858>

Ravi Vissapragada

Flinders University of South Australia: Flinders University

Gang Chen

Monash University Faculty of Economics: Monash Business School

Julie Ratcliffe

Flinders University of South Australia: Flinders University

Louise A. Mudge

The University of Adelaide School of Medicine: The University of Adelaide Adelaide Medical School

Mark B Smithers

The University of Queensland

Elizabeth A. Isenring

Bond University Faculty of Health Sciences and Medicine

Lorelle Smith

Royal Adelaide Hospital

Glyn G Jamieson

The University of Adelaide Discipline of Surgery

David I Watson

Flinders University of South Australia: Flinders University

Research

Keywords: Quality of life QLU-C10D, EQ-5D-3L, responsiveness, convergent validity, ceiling effects, esophagectomy

Posted Date: February 4th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-164925/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Aim: This study assessed the responsiveness and convergent validity of two preference-based measures; the newly developed cancer-specific EORTC Quality of Life Utility Measure-Core 10 dimensions (QLU-C10D) relative to the generic three-level version of the EuroQol 5 dimensions (EQ-5D-3L) in evaluating short-term quality-of-life outcomes/utilities after esophagectomy.

Methods: Participants were enrolled in a multicentre 2×2 factorial randomised controlled trial to determine the impact of preoperative and postoperative immunonutrition versus standard nutrition in patients with esophageal cancer. Quality-of-life was assessed seven days before and 42 days after esophagectomy. Standardized Response Mean and Effect Size were calculated to assess responsiveness. Ceiling effects for each dimension were calculated as the proportion of the best level responses for that dimension. Convergent validity was assessed using Spearman's correlation and the level of agreement was explored using Bland–Altman plots. Regression analysis was performed to identify which demographic and clinical factors influenced quality of life.

Results: Respondents were 164, predominantly male (81%) with mean age of 63 years. Quality-of-life significantly reduced on both measures with large effect sizes (>0.80), and greater mean difference on QLU-C10D. Ceiling effects were observed with social activities (86%), mobility (67%), anxiety (55%) and pain (19%) dimensions on EQ-5D-3L. For QLU-C10D ceiling effects were observed with emotional function (53%), physical function (16%), nausea (35%), sleep (31%), bowel problems (21%) and pain (20%). A strong correlation ($r=0.71$) was observed between EQ-5D-3L anxiety and QLU-C10D emotional function dimensions. Good agreement (3.7% observations outside the limits of agreement) was observed between the utility scores. Blood loss and blood transfusion predicted EQ-5D-3L utility while smoking and tumour length >3cm were predictive of QLU-C10D utility. Changes in QLQ-C30 dimensions of emotional function, role function and pain were predictive of changes in EQ-5D-3L utility while changes in physical, social and role function as well as all the symptom scales were predictive of change in QLU-C10D utility.

Conclusion: Although there is strong agreement between utility scores, QLU-C10D was more sensitive to short-term utility changes following esophagectomy. Cognisant of requirements by policy makers to apply generic utility measures in cost effectiveness studies, disease-specific measures should be used alongside the generic measures.

Trial registration

The trial was registered with the Australian New Zealand Clinical Trial Registry (ACTRN12611000178943) on the 15th February 2011.

1. Introduction

A plethora of instruments exist for the assessment of quality-of-life outcomes, including both generic and condition specific measures. As the name suggests, condition specific measures are designed with domains specific to a particular condition or disease and therefore used in assessing quality of life outcomes in that population. Generic measures on the other hand are not restricted to a particular disease or condition or population. Quality of life instruments can be further disaggregated into preference based and non-preference-

based instruments. Preference-based instruments also called multi-attribute utility instruments (MAUIs), combine a description of health and/or quality of life states and a scoring algorithm of the general population's "weighted" valuation of these states (based upon an aggregation of individuals' preferences for one state over another) [1, 2]. Preference-based instruments facilitate the calculation of quality adjusted life years (QALYs), and are therefore suitable for application in cost utility analysis (CUA), the most prevalent form of economic evaluation. Non-preference-based instruments on the other hand, may be applied in other types of economic evaluation including cost effectiveness analysis (CEA) where the calculation of QALYs is not required.

Decision making bodies including the National Institute for Health and Care Excellence (NICE) in the United Kingdom, the Pharmaceutical Benefits Advisory Committee (PBAC) and Medical Services Advisory Committee (MSAC) in Australia, require the use of generic MAUIs such as EuroQol 5 dimensions (EQ-5D-3L) in submissions of economic evaluations of cost effectiveness evidence when assessing applications of drugs or new health care technologies to be funded by their national governments [3, 4]. However, it has been argued that generic measures are not as sensitive to changes in quality of life in certain diseases as the condition specific measures and may not adequately reflect quality of life detriments or gains in these conditions which would ultimately affect any QALY estimations undertaken [5, 6]. This has led to the development of many condition specific MAUIs such as European Organization for Research and Treatment of Cancer Eight dimensions (EORTC-8D) and EORTC Quality of Life Utility Measure-Core 10 dimensions (QLU-C10D), which are both derived from the non-preference-based EORTC Quality of Life Questionnaire Core 30 item (EORTC QLQ-C30) [7, 8]. Although mapping algorithms exist to estimate the relationship between the disease-specific non-preference-based QLQ-C30 with several generic MAUIs and to ultimately generate QALYs, not many studies have compared its preference-based derivatives EORTC-8D and QLU-C10D to the generic measures [9, 10]. In their study comparing the generic EQ-5D-3L and disease specific EORTC-8D, Lorgelly et al [5] found that although both measures had similar discriminatory power, the generic QALYs were significantly lower than the disease-specific QALYs.

By far there has been a lack of longitudinal evidence on the comparisons of psychometric properties between the newly developed QLU-C10D and EQ-5D. The purpose of the data analysis reported in this paper was to contribute to this literature by assessing the responsiveness and convergent validity of the newly developed cancer specific QLU-C10D relative to the generic EQ-5D-3L in the context of short-term quality of life/utilities after esophagectomy for esophageal cancer.

2. Methods

2.1 Sample

This analysis was undertaken from a pooled sample of patients participating in a randomised control trial whose details are reported elsewhere [12]. Briefly, patients were randomised into four groups to receive an immunonutrition supplement or standard nutrition without added immunonutrients; 1) before but not after; 2) after but not before; 3) both before and after; and 4) none before or after esophagectomy. Quality of life was assessed seven days before and 42 days after esophagectomy. There were no significant differences in quality of life and clinical outcomes between the groups before and after esophagectomy. In this paper data was analysed as a pooled sample including all patients with both baseline (pre-operative) and follow-up (post-operative) quality of life scores on all instruments.

2.2 Quality Of Life Assessment

Quality of life was assessed using the self-administered and widely validated generic utility-based three level version of the EQ-5D-3L [13] and the cancer specific EORTC-QLQ-C30 [14] with the esophageal cancer add-on QLQ-OES18 [15]. The EQ-5D-3L descriptive system comprises five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression with 3 levels for each dimension: no problems, some problems, and extreme problems. The QLQ-C30 has one global HRQL scale, five functional scales (physical, role, emotional, cognitive, social), three symptom scales (fatigue, nausea or vomiting, pain) and six single items (sleeping disorders, appetite loss, dyspnoea, diarrhoea, constipation and financial problems). Each item has four alternative responses (1- not at all; 2- a little; 3- quite a bit; 4- very much). Responses to the QLQ-C30 were mapped onto the cancer specific preference based QLU-C10D to generate utility scores using the algorithm developed by King et al [7] for this purpose. QLU-C10D has four functional scales and six symptom scales obtained from 13 out of the 30 items of QLQ-C30. The functional scales include physical function, role function, social function and emotional functioning; the symptom scales include pain, fatigue, sleep, appetite, nausea and bowel problems. Being preference based, the scores range from 0 to 1. The QLQ-OES18 has four symptom scales (dysphagia, eating difficulties, reflux and pain) and six single items (trouble swallowing saliva, choking, dry mouth, taste, cough and speech); response alternatives are similar to those of the QLQ-C30.

2.3 Data Analysis

Data was analysed using Stata (StataCorp, College Station, TX, USA) [16]. Normally distributed data was analysed with one-way analysis of variance (ANOVA), and Kruskal-Wallis H test was used for analysis of non-normally distributed data. Regression analysis was performed in a stepwise manner. A p-value < 0.05 was considered statistically significant.

2.3.1 Mean difference

Utility scores for EQ-5D-3L and QLU-C10D were generated based on Australian general population scoring algorithms pertaining to the instrument [7, 17]. Basic descriptive statistics including means, medians and ranges were compared for each instrument at baseline and follow-up. The clinically important mean difference for EQ-5D-3L when used in populations with cancer varies from 0.07–0.12 [18]. A change of > 10 with the EORTC-QLQ C30 is considered clinically moderately relevant and > 20 as strongly relevant [19].

2.3.2 Responsiveness

To assess the ability of an instrument to detect changes in response to esophagectomy, or responsiveness, two statistical tests were applied, the Standardized Response Mean (SRM) and Effect Size (SES) [20]. These tests measure the change in index scores relative to the variation among the sample or the average change in scores normalized by a measure of deviation. Effect size demonstrates to what extent an instrument measures changes in quality of life relevant to a particular intervention or responsiveness. Internal responsiveness for each measure was assessed as the SES = ratio of the mean change to the standard deviation of scores at baseline statistic and the SRM = ratio of the mean change to the standard deviation of that change. Scores of < 0.20 = trivial effect, 0.20–0.50 = small effect, 0.50–0.80 = moderate effect, > 0.80 = large effect [21].

2.3.3 Ceiling effects

Ceiling effect is a “measurement limitation that occurs when the highest possible score or close to the highest score on a test or measurement instrument is reached, thereby decreasing the likelihood that the testing instrument has accurately measured the intended domain” [22]. Ceiling effects are present ‘if more than 15–20 % of respondents achieved the best possible score’ [23, 24]. Ceiling effect for EQ-5D-3L was calculated as the proportion of ‘no problem’ responses on each dimension and the proportion of ‘no problem’ in all dimensions. Similarly, for the QLU-C10D ceiling effect was calculated as the proportion of level 1 (highest level) on each dimension as well as on all dimensions. Ceiling effects were further explored by selecting those reporting full health in one instrument to see what they report in the other instrument. Lower ceiling effects suggest greater discriminant ability.

2.3.4 Convergent validity

Dimensions of the EQ-5D-3L provide a classification mechanism as opposed to measuring those constructs per se for example mobility is not a measure of physical function, however, studies assessing convergent validity of EQ-5D-3L with other instruments have utilised correlations within dimensions [25, 26]. Convergent validity was expected between dimensions measuring similar constructs on both measures such as mobility and physical function, pain/discomfort and pain. Correlations were classified as very weak ($r = 0-0.2$), weak ($r = 0.2-0.4$), moderate ($r = 0.4-0.7$), strong ($r = 0.7-0.9$) or very strong ($r = 0.9-1.0$) [21].

The limits of agreement between the instruments were explored using Bland–Altman plots [27, 28]. Utilities were power transformed to follow a normal distribution before plotting. Good agreement was demonstrated by less than 5% of points being outside of the limits of agreement (LOA).

2.3.5 Predictors of quality of life

Regression analysis was performed to identify which demographic and clinical factors (Table 1) influenced post-operative quality of life. To understand the disease specific domains that influenced change in quality of life between baseline and follow-up, a second regression analysis with the change in both functional and symptom domains of QLQ-C30 and OES-18 as independent predictors was run.

Table 1
Demographic and clinical characteristics

Characteristic	Statistics		
Demographic variables	n	Mean (sd)	Median (IQR)
Age	164	62.9 (7.9)	63.6 (58.4, 67.9)
Male Gender (n/%)			133 (81)
Alcohol (n/%)			106 (66)
Smoking (n/%)			22 (13)
Clinical variables	n	Mean (sd)	Median (IQR)
Hospital length of stay (days)	164	17 (15)	13 (11, 18)
ICU length of stay (days)	164	4 (4)	3 (1, 5)
Blood loss	163	268.1 (256.1)	200 (0, 400)
Blood transfusion (units)	164	0.3 (2.0)	0 (0,0)
Tumour length (cm)	114	3.9 (2.7)	3.3 (2, 5.5)
Total pack years	164	16 (18)	15 (0, 29)
Hospital length of stay > 10 days (n/%)			133 (81)
Tumour length > 3cm (n/%)			58 (51)
Co-morbidities			(n/%)
Hypertension			62 (36)
Diabetes			21 (13)
Respiratory			33 (20)
Cardiac			27 (17)
Treatment			(n/%)
Preoperative Radiotherapy			67 (41)
Preoperative Chemotherapy			129 (79)
ASA score			(n/%)
1 or 2			111 (68.1)
3			52 (31.9)
Pathological T staging			(n/%)
0/1/1a/1b/Tis			72 (44)
2/3			92 (56)

ICU = Intensive care unit; IQR = Interquartile range; SD = Standard deviation.

Characteristic	Statistics
TNM staging	(n/%)
0/IA/IB	75 (46)
II/IIB/IIIA/IIIB	82 (50)
IIIC/4	7 (4)
Procedure type	(n/%)
Open chest and abdominal approach	76 (46)
Hybrid (thoracoscopic)	88 (54)
<i>ICU = Intensive care unit; IQR = Interquartile range; SD = Standard deviation.</i>	

3. Results

3.1 Demographics

164 of the original cohort of 276 patients had complete before vs after surgery quality of life data available for analysis in this study. Table 1 summarises the demographic characteristics of the study sample. Participants were predominantly male (81%), mean age was 63 years, with a history of alcohol consumption (66%). The commonest surgical technique (54%) was a hybrid esophagectomy entailing an open abdominal phase, thoracoscopic chest phase and anastomosis in the left neck. Most patients had preoperative chemotherapy (79%), and length of the hospital stay was more than 10 days (81%) for most patients. There were no significant differences in demographic characteristics between patients included in this analysis and those excluded on all variables except for two variables. More patients in the excluded group had TNM stage IIIC/4 (14% compared to 4%) and underwent hybrid (thoracoscopic) esophagectomy (66% compared to 54% in those with complete data) – see Additional file table A1. However, these differences did not translate into differences in quality-of-life outcomes, see Additional file table A2.

3.2 Quality Of Life Mean Difference And Effect Size

Quality of life reduced between baseline (7 days before) and follow-up (42 days) after esophagectomy on all measures, 0.85 to 0.69 on EQ-5D-3L, 0.81 to 0.52 with QLU-C10D, and 84.3 to 62.5 with QLQ-C30 and these changes were statistically significant (Table 2). The mean score differences for EQ-5D-3L and QLQ-C30 were within the range of their respective minimum clinically important differences at 0.16 and 21.6 respectively.

The mean difference for QLU-C10D was 0.29 but the minimum clinically important difference for this measure has not yet been established. Large effects were detected for all measures using both SES and SRM (> 80). QLU-C10D was more responsive with the largest SES (1.53) and SRM (1.37).

Table 2
Descriptive statistics including mean difference and effect size

	EQ-5D-3L	QLU-C10D	EORTC summary score
Pre-operative mean (sd)	0.85 (0.15)	0.81 (0.16)	84.3 (13.1)
Post-operative mean(sd)	0.69 (0.16)	0.52 (0.22)	62.5 (17.9)
Mean difference (95% CI)	0.16 (0.13, 0.19)	0.29 (0.26, 0.32)	21.6 (18.9, 24.4)
p-value ¹	0.000*	0.000*	0.000*
Standard Effect size ²	1.08	1.53	1.38
Standard response mean ³	0.84	1.37	1.22
<i>¹ttest difference in mean between baseline (7 days before esophagectomy) and follow-up (42 days after esophagectomy); ²Cohen's d effect size calculation</i>			
<i>³Standard Response Measure (ratio of the mean change to the standard deviation of that change), < 0.20 = trivial effect, 0.20–0.50 = small effect, 0.50–0.80 = moderate effect, > 0.80 = large effect; *Statistically significant values at < 0.05</i>			

3.3 Ceiling Effects

Distribution of the scores within the sample was similar for both measures at baseline and follow-up. However, a clustering of EQ-5D indices at the upper level with a gap between 1 and the lower levels was observed at baseline. Following esophagectomy, ceiling effects (> 15%) on the EQ-5D-3L were observed with social activities (86%), mobility (67%), anxiety (55%) and pain (19%) dimensions and with all dimensions at baseline (Fig. 1). Social activities dimension showed the greatest ceiling effect at both baseline (99%) and follow-up (86%). Usual activities and pain showed the greatest (> 50%) reduction in ceiling effects between baseline and follow-up.

For QLU-C10D ceiling effects were observed with emotional function (53%), physical function (16%), nausea (35%), sleep (31%), bowel problems (21%) and pain (20%) following esophagectomy and all dimensions pre-operatively (Fig. 2). Nausea (70%) and emotional function (53%) had the highest ceiling effects at baseline and follow-up respectively. Role function and appetite showed the greatest (> 50%) reduction in ceiling effect between baseline and follow-up.

At baseline 13 respondents reported full health on both measures, while 55 reported full health on EQ-5D-3L but not QLU-C10D. Nine respondents reported full health on EQ-5D-3L but none reported full health on QLU-C10D at follow-up. Table 3 summarises QLU-C10D responses for the respondents who reported full health on EQ-5D-3L but not QLU-C10D at baseline (55) and at follow-up (9). Patients reporting full health with EQ-5D-3L still had problems when the disease specific measure was used, particularly with social function (51%) and fatigue (67%) at baseline and fatigue (78%) at follow-up where majority had less than the highest domain score.

Table 3

Distribution of QLU-C10D responses for participants reporting full health on EQ-5D but not on QLU-C10D

Level	Functional domains				Symptoms						
	Physical	Role	Social	Emotional	Pain	Fatigue	Sleep	Appetite	Nausea	Bowel	
Baseline n = 55 ¹											
1 (best)	69	80	36	80	87	25	65	73	84	65	
2	29	18	51	20	13	67	35	24	16	33	
3	2	2	9	0	0	7	0	4	0	0	
4 (worst)	0	0	4	0	0	0	0	0	0	2	
Follow-up n = 9 ¹											
1 (best)	44	56	67	89	78	0	56	33	56	67	
2	33	33	22	11	22	78	22	22	22	33	
3	22	11	11	0	0	11	22	22	22	0	
4 (worst)	0	0	0	0	0	11	0	22	0	0	
¹ 55 respondents reported full health on EQ-5D but not QLU-C10D at baseline and 9 respondents reported full health on EQ-5D but not on QLU-C10D at follow-up											

3.4 Convergent Validity

Results of the correlation between measures for both utility and dimensions scores are reported in Table 4. QLU-C10D and EQ-5D-3L utility score were strongly correlated ($r = 0.71$). Correlation between the EQ-5D-3L utility score and QLU-C10D functional domains were moderate ($r > 0.4$) but weak correlations ($r < 0.4$) were observed with the symptom domains. At the dimension level, anxiety/depression was strongly correlated ($r = 0.71$) with emotional function on the QLU-C10D while moderate correlations were observed for mobility and physical function ($r = 0.6$), usual activities with role function ($r = 0.68$), social function ($r = 0.54$), and fatigue ($r = 0.41$) as well as pain/discomfort with pain ($r = 0.55$). Very weak correlations were observed for pain/discomfort with role function ($r = 0.18$), nausea ($r = 0.17$) and bowel problems ($r = 0.2$), mobility with pain ($r = 0.19$) and nausea ($r = 0.18$), as well as usual activities with pain ($r = 0.2$) and personal care with fatigue ($r = 0.18$). Results of the correlation between QLU-C30 and EQ-5D-3L domains are presented in the appendix (Additional file Table A1).

Table 4
Correlation between QLU-C10D and EQ-5D domains at follow-up¹

	EQ-5D utility score	Mobility	Personal care	Usual activities	Pain/discomfort	Anxiety/depression
QLU-C10D utility score	0.7138	-0.4804	-0.2754	-0.5857	-0.2402	-0.4608
Function domains						
Physical function	-0.5532	0.6022	0.3468	0.4594		0.2783
Role function	-0.6132	0.4161	0.2535	0.6758	0.1767	0.2668
Social function	-0.5715	0.3075	0.2950	0.5387		0.3766
Emotional function	-0.6009	0.2474		0.3995		0.7113
Symptom domains						
Pain	-0.3190	0.1886		0.1966	0.5517	
Fatigue	-0.4716	0.3265	0.1755	0.4095		0.3155
Sleep	-0.3385	0.2567	0.1926	0.2763		0.3241
Appetite	-0.3470	0.2106		0.2769		0.3235
Nausea	-0.3790	0.1829		0.2721	0.1746	0.3561
Bowel problems	-0.2231				0.1990	0.2780
¹ Correlations between dimensions were explored using patient responses at follow-up (42 days after esophagectomy). Only significant correlations (at < 0.05) are reported. Correlations were classified as very weak ($r = 0-0.2$), weak ($r = 0.2-0.4$), moderate ($r = 0.4-0.7$), strong ($r = 0.7-0.9$) or very strong ($r = 0.9-1.0$).						

The Bland Altman plot (Fig. 3) showed a small mean difference and good agreement between QLU-C10D and EQ-5D-3L utility scores as only < 5% observations were outside the limits of agreement.

3.5 Factors Influencing Quality Of Life

Tables 5 shows the factors influencing post-operative quality of life. After controlling for pre-operative quality of life, reduction in EPA between day 1 and 7 was associated with high post-operative QoL on both measures, blood loss and blood transfusion predicted post-operative EQ-5D-3L utility while smoking and tumour length > 3cm were predictive of QLU-C10D utility.

Table 5
Factors influencing post-operative utility score

Clinical predictor	EQ-5D-3L utility score		QLU-C10D utility score	
	Coefficient (SE)	Beta	Coefficient (SE)	Beta
QLU-C10D pre-op			0.472 (0.114)**	0.360
Smoking			-0.128 (0.057)*	-0.198
Tumour length > 3cm			0.102 (0.039)**	0.236
Blood loss of 1 or more units	-0.062 (0.031)*	-0.179		
Blood transfusion	-0.017 (0.006)**	-0.267		
Change in Bloods Dminus EPA (day 1 – day 7)	-0.015 (0.006)*	-0.223	-0.022 (0.009)*	-0.225
Constant	0.760 (0.028)		0.145 (0.099)	
N	110		110	
R-squared	0.169		0.251	
<i>Stepwise regression results reported. Standard errors in parentheses, *p < 0.05, **p < 0.01</i>				

Table 6 shows the dimensions of QLQ-C30 and OES-18 that are predictive of the change in quality of life/utility. Changes in QLQ-C30 functional domains were the strongest (highest beta) predictors of change in utility scores for both measures.

Table 6
Function and symptom domains predicting change in quality of life

Domain	Change in EQ-5D-3L utility		Change in QLU-C10D utility	
	Coefficient (SE)	Beta	Coefficient (SE)	Beta
Change in QLQ-C30 functional scores[@]				
Physical function			0.003 (0.0004)**	0.273
Emotional function	0.002 (0.0005)**	0.261		
Social function			0.001 (0.0003)**	0.202
Role function	0.002 (0.0004)**	0.293	0.001 (0.0003)*	0.090
Cognitive function				
Change in QLQ-C30 symptom scores[#]				
Nausea and vomiting			-0.002 (0.0002)**	-0.217
Pain	-0.002 (0.0004)**	-0.296	-0.002 (0.0002)**	-0.242
Appetite loss			-0.001 (0.0002)**	-0.17
Diarrhoea			-0.001 (0.0002)**	-0.078
Insomnia			-0.001 (0.0002)**	-0.13
Fatigue			-0.001 (0.0003)*	-0.087
Change in OES-18 symptoms scores[#]				
OES Swallowing saliva	0.001 (0.0004)*	0.155		
Constant	0.044 (0.018)		0.017 (0.011)	
N	155		155	
R-squared	0.4158		0.8898	
<i>Stepwise regression results reported. Standard errors in parentheses, *p < 0.05, **p < 0.01; @A higher score indicates better QoL. #A higher score indicates worse QoL</i>				

Changes in QLQ-C30 dimensions of emotional function ($\beta = 0.26$), role function ($\beta = 0.29$) and pain ($\beta = -0.296$) were predictive of changes in EQ-5D-3L utility while changes in physical ($\beta = 0.27$), social ($\beta = 0.2$) and role function (0.09) as well as all the symptom scales were predictive of change in QLU-C10D utility. The strongest predictors were pain for EQ-5D-3L and physical function for QLU-C10D.

4. Discussion

The clinical trial [12] underpinning the current study compared a homogenous group of patients with esophageal cancer (Table 1) who underwent surgical resection and received different regimens of nutritional support. As the outcomes from the trial showed no differences in clinical and quality of life outcomes between the different nutritional support regimens the data was analysed as a pooled sample in this current paper. This analysis

assessed the responsiveness and convergent validity of the cancer specific QLU-C10D and generic EQ-5D-3L for measurement of short-term quality of life outcomes following esophagectomy.

As expected, both measures showed statistically significant reductions in scores following surgery (Table 2). The mean score difference for the EQ-5D-3L was clinically significant at 0.16 [18]. The mean score difference for the QLQ-C30 was strongly relevant clinically at 21.6 [19]. As the QLU-C10D is a relatively new instrument, there has been no minimally important difference reported in the literature as yet. On comparing the EQ-5D and QLU-C10D, the EQ-5D mean utility score was greater than QLU-C10D by 0.04 at baseline and 0.17 at follow-up. The small difference at baseline but not at follow-up suggests a high degree of convergence for mild health states but not for severe states such as after esophagectomy, where limitations of a generic measure become apparent, and a more sensitive disease specific measure is preferred.

All instruments reported a large effect size with QLU-C10D and EORTC QLQ-C30 being larger than the EQ-5D-3L. Most studies with EQ-5D-3L have reported low to moderate effect size [29]. However, studies with large expected changes in health status such as after surgery have reported large effect size [30–32]. Although both EQ-5D-3L and QLU-C10D are on the 0–1 scale, a much larger effect size was observed with QLU-C10D. This observation is similar to other studies comparing disease specific and generic measures where disease specific measures have much larger effect sizes or show greater responsiveness [33]. This is because the disease specific measures, unlike the generic, assess domains of quality of life that are of greatest importance to the particular condition.

For both EQ-5D-3L and QLU-C10D, ceiling effects were observed for all dimensions at both pre- and post-operative, with higher ceiling effects observed with EQ-5D-3L (Figs. 2 and 3). The highest ceiling effect for EQ-5D-3L was with social care at both timepoints and the highest for QLU-C10D was nausea symptoms pre-operative and emotional function after esophagectomy. Ceiling effects of the EQ-5D-3L have been reported in several patient populations but the levels observed here were higher than reported in other studies among similar populations [34, 35]. In addition, the proportion reporting the best score for all dimensions (or full health) was much lower with the QLU-C10D at both timepoints. Yet patients reporting full health with EQ-5D-3L still reported problems when the disease-specific measure was used, particularly with social function and fatigue at pre-operative and fatigue at after esophagectomy (Table 4). This is similar to findings by Lorgelly et al [5] who showed that EQ-5D-3L was less sensitive to fatigue impairment when compared to the disease specific EORTC-8D. Although EQ-5D-3L had a large effect size and demonstrated a clinically meaningful change in utility scores, the high ceiling effects suggest low discriminant ability. This means that the EQ-5D-3L is not as capable of identifying all or most of the change in quality of life and distinguishing health states in this population. To address such ceiling effects experienced with EQ-5D-3L, a 5-level version of the instrument, EQ-5D-5L has been developed [36], however, no studies are available comparing it to the QLU-C10D yet.

Strong correlations were observed between QLU-C10D and EQ-5D-3L utility scores. In addition, the EQ-5D utility score was moderately correlated with all functional domains but weakly correlated with the symptom domains of the QLU-C10D. It's important to note that domains that assessed similar constructs such as emotional function on the QLU-C10D and anxiety/depression on the EQ-5D showed a strong correlation or moderate correlation such as physical function with mobility, role function, social function and fatigue with usual activities, as well as pain with pain/discomfort. The Bland-Altman plot showed a small mean difference and good agreement between the measures.

Factors Influencing Change In Quality Of Life

It is important to establish whether certain risk factors or particular groups of patients are vulnerable to a poorer outcome. We investigated this through examining the overall postoperative utility using EQ-5D-3L and QLU-C10D to see which was more sensitive to differentiating between the risk factors. It was important to choose postoperative QoL as the outcome (versus change in QoL) because a change in QoL could over or underestimate the QoL depending on the pre-operative score. This is especially true in an acute setting, where patients experience a large drop in quality of life, but the postoperative score can be used independently to score the impact of the disease state. Smoking was a strong predictor of poor QLU-C10D utility. It is likely that current smoking status is a stronger predictor of poor post-operative QoL than overall history of smoking, which has been seen previously in the literature [37]. In addition, tumour length was seen to be predictive of QoL, but surprisingly as a positive correlation. Approximately 51% of the patients in our sample had tumours that were larger than 3 cm, it is possible that this result is due to confounders. Larger tumours often require larger radiotherapy fields during the neoadjuvant treatment phase, and a larger tumour can require more extensive dissection, potentially increasing the risk of early postoperative complications that would be expected to adversely impact QoL outcomes. In general, however, previously published literature suggests that tumour length is not associated with worsening QoL scores [38]. Blood loss and blood transfusion requirement were weakly correlated with EQ-5D-3L utility, and this is likely because transfusion requirements are indicators of more difficult procedures, which might increase complication risk and adversely impact early quality of life outcomes. However, transfusion was not correlated with QLU-C10D utility, suggesting that EQ-5D-3L might be a better instrument for investigating generic changes in QoL rather than those specific to disease.

A small change in the functional scores was predictive of higher follow-up quality of life. For the symptom scores, a large change in scores was predictive of greater change in utility between baseline and follow-up. Changes in symptom scores were stronger predictors of change in quality of life than changes in functional scores. This implies that in assessing short-term quality of life following esophagectomy, an instrument that is more sensitive to changes in symptoms is preferred, in this case the QLU-C10D is preferred to the EQ-5D-3L.

5. Conclusion

The findings from this study suggest that generic utility instruments such as the EQ-5D-3L may not be sufficient in isolation when assessing short-term quality of life following esophagectomy and should be combined with a condition specific measure. Cognisant of requirements by policy makers such as NICE in the UK and PBAC and MSAC in Australia to apply generic utility instruments in cost effectiveness studies, we recommend the application of a condition specific utility instrument alongside the generic instrument. In this way the sensitivity of a cost-effectiveness assessment decision can be determined by taking into account utility estimates generated from both a condition specific and a generic perspective.

Abbreviations

CEA
Cost Effectiveness Analysis
CUA
Cost Utility Analysis

EORTC
European Organization for Research and Treatment of Cancer
EORTC QLQ-C30
EORTC Quality of Life Questionnaire Core 30 item
EORTC QLU-C10D
EORTC Quality of Life Utility Measure-Core 10 dimensions
EORTC-8D
EORTC Eight dimensions
EQ-5D
EuroQol 5 dimensions
LOA
Limits of Agreement
MAUIs
Multi-Attribute Utility Instruments
MSAC
Medical Services Advisory Committee
NICE
National Institute for Health and Care Excellence
PBAC
Pharmaceutical Benefits Advisory Committee
QALYs
Quality Adjusted Life Years
SES
Standard Effect Size
SRM
Standardized Response Mean

Declarations

Ethics approval and consent to participate: Ethical approval for the study was given by the Royal Adelaide Hospital Clinical Research Ethics Committee, and the ethics committees at each participating site. Patients were recruited from 11 major teaching hospitals undertaking

oesophageal cancer surgery across Australia. Informed consent was obtained from all participants.

Consent for publication: Not applicable

Availability of data and materials: The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests: The authors declare that they have no competing interests

Funding: This study was funded by a National Health & Medical Research Council of Australia Competing Project Grant (627110). The trial was registered with the Australian New Zealand Clinical Trial Registry (ACTRN12611000178943) on the 15th February 2011.

Professor Watson and Professor Ratcliffe received a Beat Cancer Hospital Research Package Grant which was funded by the Cancer Council of South Australia's Beat Cancer Project on behalf of its donors and the State Government of South Australia Department of Health, together with the support of the Flinders Medical Centre Foundation, its donors and partners. This Grant funded Dr. Norma Bulamu's and Dr. Gang Chen's salary.

Authors' contributions: NBB, RV, GC, JR and DIW were responsible for study conception, data analysis and interpretation of results. LAM, MBS, EAI, LS, GGJ and DIW were part of the Australian Immunonutrition Study Group, which was responsible for running the randomised control trial, data collection and reporting of the first study. NBB, RV and GC wrote the first draft and all co-authors contributed towards revising and editing the manuscript. All authors read and approved the final manuscript.

Acknowledgements:**The Australian Immunonutrition Study Group:** Louise A. Mudge¹, Dr Justin Bessell², A/Prof Andrew Barbour³, Elizabeth A. Isenring⁴, Prof Ahmad Aly⁵, Dr Stephen Archer⁶, A/Prof Mohammed Ballal⁷, Jessica Barbon⁸, Katie Benton⁹, Melissa Bond¹⁰, Melissa Berryman¹¹, Dr Tim Bright², A/Prof Richard Cade¹², Anna Cardamis¹¹, Rosemary Carroll¹³, Kara Cashman¹⁴, Larissa Chan¹⁵, Brooke Chapman¹⁶, Sally Chapman¹⁷, Dr Derek Chen⁶, Dr Jacob Chisholm², Wendy Davidson⁹, A/Prof Peter Devitt¹, Dr Cuong Dong¹⁸, Ra'eesa Doola¹⁷, Suzanne Edwards¹⁴, Dr Krishna Epari⁷, Maree Farley⁵, Julie Farrow¹⁹, Dr Maree Ferguson⁹, Prof David Fletcher⁷, Kerry Forbes²⁰, Katherine Fullerton¹⁹, Dr Philip Game¹, Dr Susan Gan², Prof David Gotley³, Belinda Gout¹⁶, Jodi Gray²¹, Susan Heaney¹⁹, Glyn G. Jamieson¹, Dr Mary Anne Johnson¹², Megan Johnstone¹, Dr Sanjeeva Kariyawasam⁷, Prof Jonathan Karnon²¹, Amber Kelaart²², Liz Kellett⁸, Erin Kennedy¹⁶, Rhiannon Krane⁸, Sylvia Lemass⁹, Jenelle Loeliger²², Dr Andrew Lord²³, Prof John Ludbrook²⁴, Catherine McFarlane¹⁷, Michelle McPhee¹¹, Selena Yue-Xian Ooi¹⁰, Leonie Pearce¹⁶, Kate Pettigrew⁸, Emma Putrus¹⁰, Georgina Rassias¹⁰, Alison Shanks¹⁰, Dr Jon Shenfine¹, Emma Louise Smith⁸, Lorelle Smith¹, Justin Singleton¹⁹, B. Mark Smithers³, Rebecca Lindstrom-Sowman¹⁹, Dr John Spillane¹⁸, Liliana Sputore²⁰, Belinda Steer²², Tom Sullivan¹⁴, Laisa Teleni⁹, Deb Tolcher¹⁷, Janine Thomas³, A/Prof Sarah Thompson¹, Tina Thorpe¹⁸, David I. Watson², Cheryl Watterson¹⁹, Dr Vanessa Wills¹³, Anita Wilton¹¹, Kellie Wright¹⁶, Dr Tim Wright¹³

1. Discipline of Surgery, The University of Adelaide,
2. Discipline of Surgery, College of Medicine and Public Health, Flinders University, Adelaide
3. Discipline of Surgery, University of Queensland, Upper GI and Soft Tissue Unit, Princess Alexandra Hospital, Brisbane
4. Faculty of Health Sciences & Medicine, Bond University, Brisbane
5. Upper GI Surgery, Austin Health, Melbourne
6. Discipline of Surgery, Royal Perth Hospital
7. Department of Surgery, Freemantle Hospital and Health Services, Perth
8. Dietetics and Nutrition, Flinders Medical Centre, Adelaide
9. Nutrition and Dietetics, Princess Alexandra Hospital, Metro South Health, Brisbane
10. Department of Clinical Dietetics, Royal Adelaide Hospital
11. Dietetics, Box Hill Hospital, Eastern Health, Melbourne
12. Upper GI Surgery, Box Hill Hospital, Eastern Health, Melbourne
13. Surgical Services, John Hunter Hospital, Hunter New England Health, Newcastle

14. Data Management & Analysis Centre, The University of Adelaide
15. Dietetics and Nutrition, Royal Perth Hospital
16. Nutrition and Dietetics, Austin Health, Melbourne
17. Nutrition and Dietetics, Mater Private Hospital, Brisbane
18. Upper Gastrointestinal Service, Peter MacCallum Cancer Centre, Melbourne
19. Department of Nutrition and Dietetics, John Hunter Hospital, Hunter New England Health, Newcastle
20. Dietetic and Nutrition Department, Freemantle Hospital and Health Services, Perth
21. Public Health, The University of Adelaide
22. Nutrition and Dietetics, Peter MacCallum Cancer Centre, Melbourne
23. St Andrew's Hospital, Adelaide
24. Deceased

References

1. Brazier J, Ratcliffe J, Salomon AJ, Tsuchiya A. Methods for obtaining health state values: generic preference based measures of health and the alternatives. In: *Measuring and Valuing Health Benefits for Economic Evaluation*. New York: Oxford University Press; 2007c. pp. 175–256.
2. Drummond FM, Sculpher JM, Torrance WG, O'Brien JB, Stoddart LG. Multi-attribute health status classification systems with preference scores *Methods for the Economic Evaluation of Health Care Programs*. Third Edition ed. New York: Oxford University Press; 2005d. p. 154 – 72.
3. Harris A, Bulfone L. Getting value for money: The Australian experience. In: Jost TS, editor. *Health Care Coverage Determinations: An International Comparative Study*. Maidenhead: Open University Press, McGraw-Hill International; 2004.
4. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. In: NICE, editor. 2013.
5. Lorgelly PK, Doble B, Rowen D, Brazier J. Condition-specific or generic preference-based measures in oncology? A comparison of the EORTC-8D and the EQ-5D-3L. *Qual Life Res*. 2017;26(5):1163–76. doi:10.1007/s11136-016-1443-y.
6. 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-. doi:10.1186/1477-7525-9-106.
7. 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. doi:10.1007/s11136-015-1217-y.
8. Rowen D, Brazier J, Young T, Gaugris S, Craig BM, King MT, et al. Deriving a preference-based measure for cancer using the EORTC QLQ-C30. *Value Health*. 2011;14(5):721–31. doi:10.1016/j.jval.2011.01.004.
9. Doble B, Lorgelly P. Mapping the EORTC QLQ-C30 onto the EQ-5D-3L: assessing the external validity of existing mapping algorithms. *Qual Life Res*. 2016;25(4):891–911. doi:10.1007/s11136-015-1116-2.
10. Mukuria C, Rowen D, Harnan S, Rawdin A, Wong R, Ara R, et al. An Updated Systematic Review of Studies Mapping (or Cross-Walking) Measures of Health-Related Quality of Life to Generic Preference-Based

- Measures to Generate Utility Values. *Appl Health Econ Health Policy*. 2019;17(3):295–313. doi:10.1007/s40258-019-00467-6.
11. Kim H, Cook G, Goodall S, Liew D. Comparison of Disease-Specific (QLU-C10D) and Generic (EQ-5D) Utilities in Cost-Effectiveness Analysis. *Value in Health Regional Issues*. 2020;22:23. doi:https://doi.org/10.1016/j.vhri.2020.07.147.
 12. Mudge LA, Watson DI, Smithers BM, Isenring EA, Smith L, Jamieson GG. Multicentre factorial randomized clinical trial of perioperative immunonutrition versus standard nutrition for patients undergoing surgical resection of oesophageal cancer. *Br J Surg*. 2018;105(10):1262–72. doi:10.1002/bjs.10923.
 13. 2014. <http://www.euroqol.org/>. Accessed 24/01/2014.
 14. 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. doi:10.1093/jnci/85.5.365.
 15. Blazeby JM, Conroy T, Hammerlid E, Fayers P, Sezer O, Koller M, et al. Clinical and psychometric validation of an EORTC questionnaire module, the EORTC QLQ-OES18, to assess quality of life in patients with oesophageal cancer. *Eur J Cancer*. 2003;39(10):1384–94. doi:10.1016/s0959-8049(03)00270-3.
 16. StataCorp. *Stata Statistical Software: Release 15*. College Station. TX: StataCorp LLC; 2017.
 17. Viney R, Norman R, King MT, Cronin P, Street DJ, Knox S, et al. Time trade-off derived EQ-5D weights for Australia. *Value Health*. 2011;14(6):928–36. doi:10.1016/j.jval.2011.04.009.
 18. 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. doi:10.1186/1477-7525-5-70.
 19. Kvam AK, Fayers PM, Wisloff F. Responsiveness and minimal important score differences in quality-of-life questionnaires: a comparison of the EORTC QLQ-C30 cancer-specific questionnaire to the generic utility questionnaires EQ-5D and 15D in patients with multiple myeloma. *Eur J Haematol*. 2011;87(4):330–7. doi:10.1111/j.1600-0609.2011.01665.x.
 20. Beaton DE, Bombardier C, Katz JN, Wright JG. A taxonomy for responsiveness. *J Clin Epidemiol*. 2001;54(12):1204–17. doi:10.1016/s0895-4356(01)00407-3.
 21. Cohen J. *Statistical power analysis for the behavioral sciences*. Academic press; 2013.
 22. Taylor TH. Ceiling Effect. In: Salkind NJ, editor. *Encyclopedia of Research Design*. Thousand Oaks: SAGE Publications, Inc.; 2010. pp. 133–4.
 23. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Quality of life research*. 1995;4(4):293–307.
 24. Garin O. Ceiling Effect. In: Michalos AC, editor. *Encyclopedia of Quality of Life and Well-Being Research*. Dordrecht: Springer Netherlands; 2014. pp. 631–3.
 25. Obradovic M, Lal A, Liedgens H. Validity and responsiveness of EuroQol-5 dimension (EQ-5D) versus Short Form-6 dimension (SF-6D) questionnaire in chronic pain. *Health Qual Life Outcomes*. 2013;11:110. doi:10.1186/1477-7525-11-110.
 26. Ravens-Sieberer U, Wille N, Badia X, Bonsel G, Burström K, Cavrini G, et al. Feasibility, reliability, and validity of the EQ-5D-Y: results from a multinational study. *Qual Life Res*. 2010;19(6):887–97. doi:10.1007/s11136-010-9649-x.

27. Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet*. 1986;327(8476):307–10.
28. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *International journal of nursing studies*. 2010;47(8):931–6.
29. Tordrup D, Mossman J, Kanavos P. Responsiveness of the EQ-5D to clinical change: is the patient experience adequately represented? *Int J Technol Assess Health Care*. 2014;30(1):10–9. doi:10.1017/s0266462313000640.
30. Brazier JE, Harper R, Munro J, Walters SJ, Snaith ML. Generic and condition-specific outcome measures for people with osteoarthritis of the knee. *Rheumatology*. 1999;38(9):870–7. doi:10.1093/rheumatology/38.9.870.
31. Krabbe PFM, Peerenboom L, Langenhoff BS, Ruers TJM. Responsiveness of the generic EQ-5D summary measure compared to the disease-specific EORTC QLQ C-30. *Qual Life Res*. 2004;13(7):1247–53. doi:10.1023/B:QURE.0000037498.00754.b8.
32. Solberg TK, Olsen J-A, Ingebrigtsen T, Hofoss D, Nygaard ØP. Health-related quality of life assessment by the EuroQol-5D can provide cost-utility data in the field of low-back surgery. *Eur Spine J*. 2005;14(10):1000–7. doi:10.1007/s00586-005-0898-2.
33. Krahn M, Bremner KE, Tomlinson G, Ritvo P, Irvine J, Naglie G. Responsiveness of disease-specific and generic utility instruments in prostate cancer patients. *Qual Life Res*. 2006;16(3):509. doi:10.1007/s11136-006-9132-x.
34. Kaarlola A, Pettila V, Kekki P. Performance of two measures of general health-related quality of life, the EQ-5D and the RAND-36 among critically ill patients. *Intensive Care Med*. 2004;30(12):2245–52. doi:10.1007/s00134-004-2471-6.
35. 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. doi:10.1007/s11136-011-0018-1.
36. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of life research: an international journal of quality of life aspects of treatment care rehabilitation*. 2011;20(10):1727–36. doi:10.1007/s11136-011-9903-x.
37. Balduyck B, Sardari Nia P, Cogen A, Dockx Y, Lauwers P, Hendriks J, et al. The effect of smoking cessation on quality of life after lung cancer surgery. *Eur J Cardiothorac Surg*. 2011;40(6):1432-7; discussion 7–8. doi:10.1016/j.ejcts.2011.03.004.
38. Earlam S, Glover C, Fordy C, Burke D, Allen-Mersh TG. Relation between tumor size, quality of life, and survival in patients with colorectal liver metastases. *J Clin Oncol*. 1996;14(1):171–5. doi:10.1200/jco.1996.14.1.171.

Figures

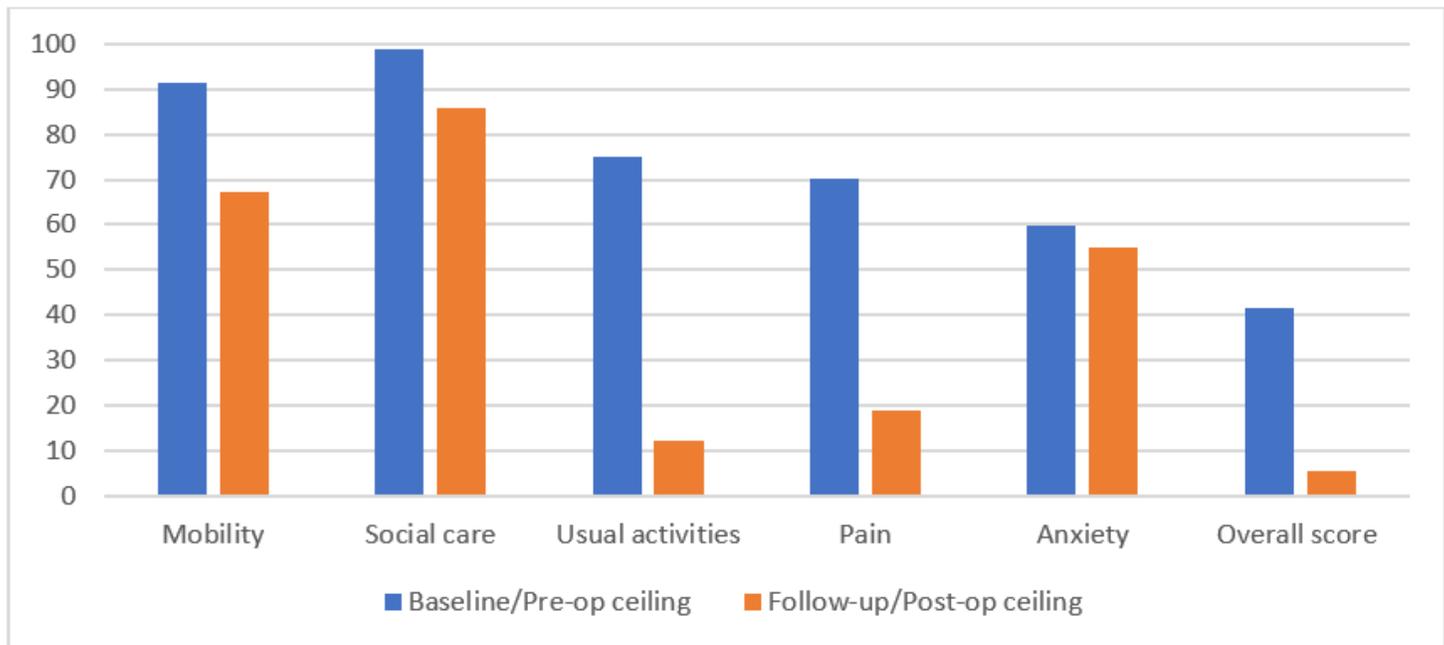


Figure 1

EQ-5D-3L ceiling effect at baseline and follow-up

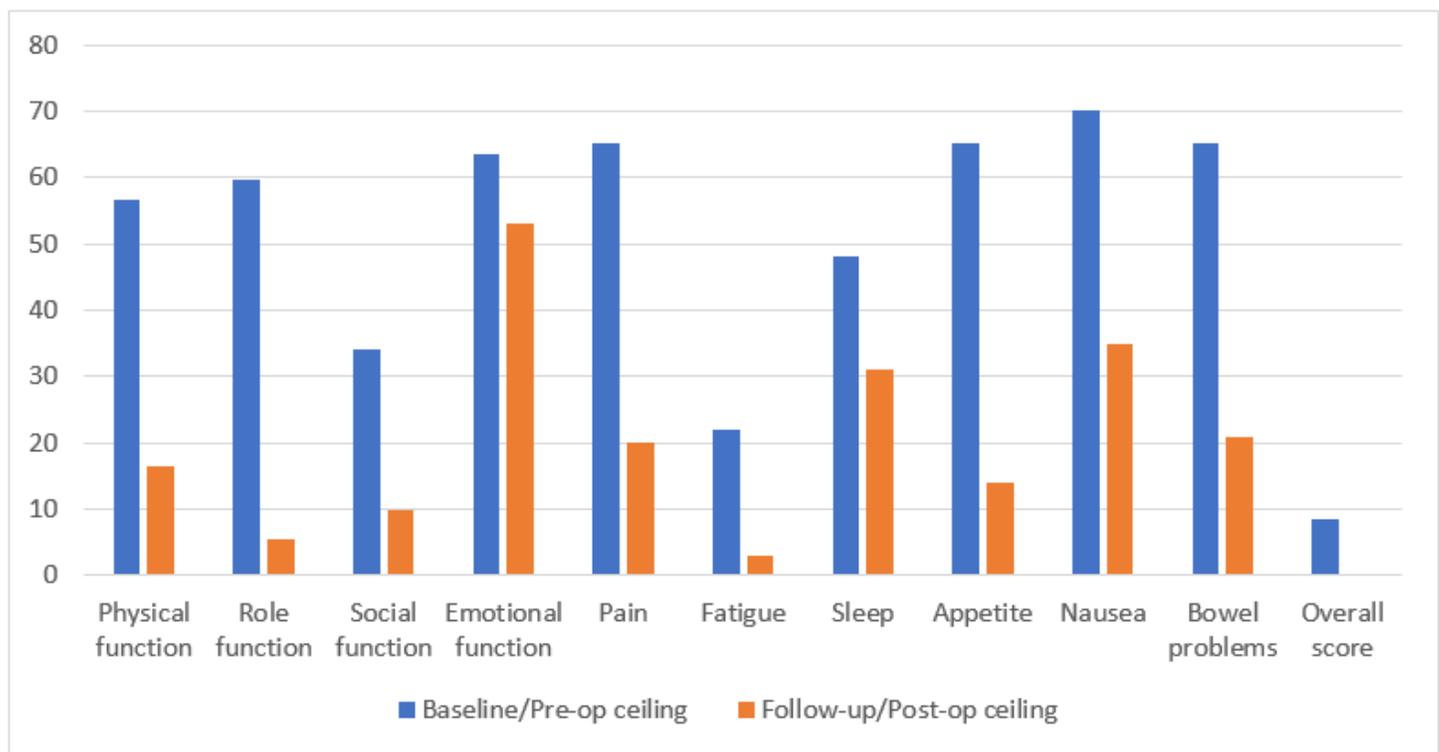


Figure 2

QLU-C10D ceiling effects at baseline and follow-up

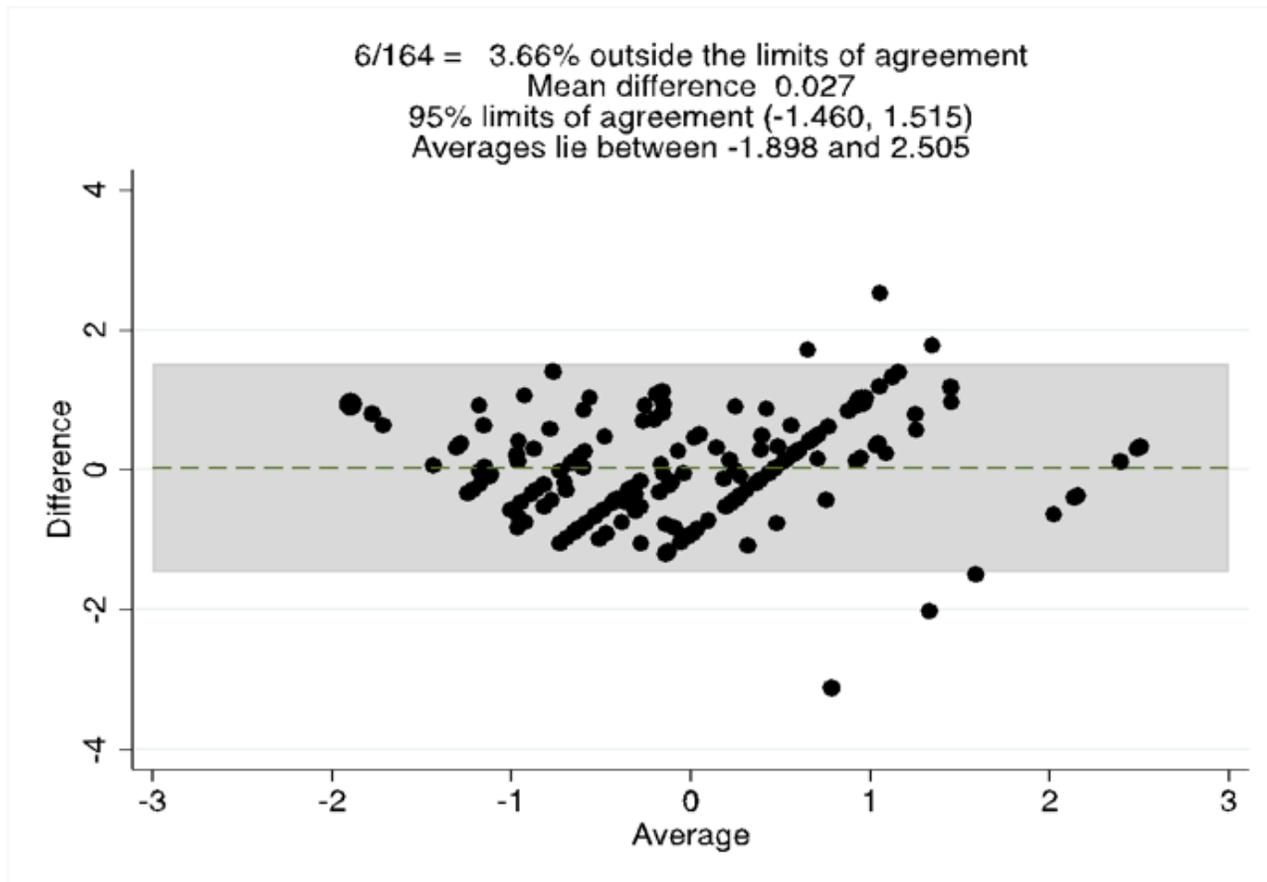


Figure 3

Bland–Altman plot of QLU-C10D and EQ-5D-3L at follow-up

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile.docx](#)