



# Cancer-Specific Health Utilities: Evaluation of Core Measurement Properties of the EORTC QLU-C10D in Lung Cancer Patients—Data from Four Multicentre LUX-Lung Trials, Applying Six Country Tariffs

Micha J. Pilz<sup>1</sup> · Simone Seyringer<sup>2</sup> · Imad Al-Naesan<sup>1</sup> · Madeleine T. King<sup>3</sup> · Andrew Bottomley<sup>4</sup> · Richard Norman<sup>5</sup> · Lisa Schlosser<sup>6</sup> · Tobias Hell<sup>6</sup> · Eva Maria Gamper<sup>2</sup> on behalf of the EORTC Quality of Life Group

Accepted: 11 March 2024 / Published online: 2 May 2024

© The Author(s) 2024

## Abstract

**Background** Cost-utility analysis generally requires valid preference-based measures (PBMs) to assess the utility of patient health. While generic PBMs are widely used, disease-specific PBMs may capture additional aspects of health relevant for certain patient populations. This study investigates the construct and concurrent criterion validity of the cancer-specific European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Utility-Core 10 dimensions (QLU-C10D) in non-small-cell lung cancer patients.

**Methods** We retrospectively analysed data from four multicentre LUX-Lung trials, all of which had administered the EORTC Quality of Life Questionnaire (QLQ-C30) and the EQ-5D-3L. We applied six country-specific value sets (Australia, Canada, Italy, the Netherlands, Poland, and the United Kingdom) to both instruments. Criterion validity was assessed via correlations between the instruments' utility scores. Correlations of divergent and convergent domains and Bland-Altman plots investigated construct validity. Floor and ceiling effects were assessed.

**Results** The comparison of the EORTC QLU-C10D and EQ-5D-3L produced homogenous results for five of the six country tariffs. High correlations of utilities ( $r > 0.7$ ) were found for all country tariffs except for the Netherlands. Moderate to high correlations of converging domain pairs ( $r$  from 0.472 to 0.718) were found with few exceptions, such as the Social Functioning–Usual Activities domain pair (max.  $r = 0.376$ ). For all but the Dutch tariff, the EORTC QLU-C10D produced consistently lower utility values compared to the EQ-5D-3L ( $\bar{x}$  difference from  $-0.082$  to  $0.033$ ). Floor and ceiling effects were consistently lower for the EORTC QLU-C10D (max. 4.67% for utilities).

**Conclusions** The six country tariffs showed good psychometric properties for the EORTC QLU-C10D in lung cancer patients. Criterion and construct validity was established. The QLU-C10D showed superior measurement precision towards the upper and lower end of the scale compared to the EQ-5D-3L, which is important when cost-utility analysis seeks to measure health change across the severity spectrum.

## 1 Introduction

Lung cancer is the second most common malign tumour worldwide [1] and the most common cause of death from cancer [2]. Non-small-cell lung cancer (NSCLC) represents a massive burden, economically [3] as well as clinically [4] for healthcare systems. It accounts for up to 85% of all new lung cancer diagnoses [4]. The majority of patients with NSCLC are already in an advanced disease stage (IIIB or

IV) at diagnosis and in desperate need of treatment [4]. With less than 20% of newly diagnosed NSCLC patients alive after 5 years, the overall global 5-year survival rate for patients with NSCLC is very low [5]. Inoperable stage III NSCLC is typically treated with radiotherapy and standard chemotherapeutic agents like cisplatin, while the use of targeted treatments remains under investigation [6]. Patients with activating somatic mutations of the tyrosine kinase domain of epidermal growth factor receptor (EGFR) constitute up to one third of patients with NSCLC [7, 8].

Patients with EGFR mutations tend to have better clinical outcomes when treated with EGFR tyrosine kinase inhibitors (TKI) rather than chemotherapy [9, 10]. EGFR TKIs,

The research in this article was done during AB's tenure at EORTC. AB is no longer affiliated with EORTC.

Extended author information available on the last page of the article

### Key Points for Decision Makers

The cancer-specific preference-based measure European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Utility-Core 10 dimensions (QLU-C10D) is now a validated instrument for health economic evaluations in the lung cancer population.

Generally, there is a good concordance for the health state utility values derived by the generic measure EQ-5D-3L and the cancer-specific measure EORTC QLU-C10D.

The EORTC QLU-C10D shows improved measurement precision towards the upper and lower end of the scale compared to the generic measure EQ-5D-3L.

When utilising the EORTC Quality of Life Questionnaire (QLQ-C30) for clinical outcomes assessment in lung cancer trials, the QLU-C10D now complements its parent instrument with an algorithm useful for health economic decision making.

including afatinib, have been investigated as therapeutic agents for treating advanced NSCLC [11–13]. The EGFR is currently the most established molecular target in NSCLC. Afatinib is an irreversible epidermal growth factor family inhibitor. It improved first-line progression-free survival compared with chemotherapy in two large phase 3 trials in patients with EGFR mutation-positive advanced lung adenocarcinoma, as well as improving overall survival in patients with the EGFR del19 mutation [14–16].

Complementing the major challenge of improving survival, the assessment of patient-reported outcomes (PROs) and health-related quality of life (HRQoL) is an important aspect from the clinical [17] and economic perspective [3] in NSCLC patients. PRO data can be used to assess treatment options and be of particular interest if marginal differences are observed in overall survival or toxicity profile of different agents. PRO data are therefore valuable in health economic decision making [18]. Considering the significant burden NSCLC poses on both patients and healthcare systems [3], the evaluation of treatments and the assessment of the patients and societal preferences remains an essential task [19]. The assessment of societal preferences in allocating healthcare resources can be supported by assessing PROs using preference-based measures (PBMs) that allow the estimation of quality-adjusted life years (QALYs) for use in cost-utility analysis (CUA) [20], which is widely used by Health Technology Assessment (HTA) agencies in most industrialised countries.

PBMs are based on a health state classification system in conjunction with a value set consisting of utility decrements used to determine health state utility values [21, 22]. These health state utility values generally range from 0 (dead) to 1 (full or perfect health) [23]. Several generic PBMs such as the EQ-5D-3L [24] and others [21, 25] are used to facilitate health economic evaluations in various disease settings. Generic PBMs assess general and universally applicable health domains and are therefore used to estimate utility values across a variety of medical conditions. The comparability of results across disease groups and patient populations [26] currently makes generic PBMs the primary instruments to facilitate CUA assessments [27]. Complementing the generic PBMs, disease-specific PBMs have been developed to assess health state utility values in specific patient groups [26, 28–30]. Disease-specific PBMs are conceptualised to capture the most relevant aspects of health in certain disease and patient populations. Therefore, health state changes relevant to a certain population (e.g. changes in nausea, fatigue, or appetite loss for cancer patients) [22] can be accounted for when deriving health state utility values and performing CUAs [31].

For the estimation of health utility values in the cancer patient population, the European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Utility-Core 10 dimensions (QLU-C10D) was previously developed. For its development, the structure and content of the widely used HRQoL questionnaire EORTC Quality of Life Questionnaire (QLQ-C30) was utilised to identify the most relevant HRQoL domains to cancer patients. The EORTC QLU-C10D is therefore designed to provide a scoring algorithm for EORTC QLQ-C30, allowing the calculation of utility values from the respective PRO data [22].

In view of the different PBMs and their dis/advantages, there is an ongoing discussion regarding which PBM to use in a certain scenario (e.g. palliative care, elderly people) and condition (e.g. cancer, chronic disease) [28, 32–34]. Various aspects of generic versus disease-specific PBMs are relevant to this discussion; one of these is psychometric criteria. Validity aspects, such as criterion validity and construct validity, are considered important psychometric criteria for health status measures [35]. Therefore, the aim of this study was to assess the validity of a cancer-specific PBM, the EORTC QLU-C10D, in the NSCLC patient population using data from four LUX-Lung trials [36–39], and to compare it with the validity of the generic PBM, the EQ-5D-3L. In this article, we focus on three quality criteria for questionnaires assessing health status [35], namely floor and ceiling effects, criterion validity, and construct validity; while these concepts are widely used in psychology, they are similarly important when evaluating instruments for use underpinning estimation of QALYs for CUA. Additionally, we investigate the impact of applying different country tariff

utility decrements to the PRO data. This provides insight whether different country tariffs affect the validity parameters of the EORTC QLU-C10D. This article contributes to the body of knowledge concerning the use of the EORTC QLU-C10D as a PBM in different cancer populations and treatment contexts.

## 2 Methods

### 2.1 The EORTC QLQ-C30

The EORTC QLQ-C30 version 3.0 [40] consists of 30 questions that form 15 scales, five of which are functioning scales (Physical, Role, Emotional, Social, and Cognitive), nine are symptom scales (Fatigue, Nausea and Vomiting, Pain, Dyspnoea, Insomnia, Appetite Loss, Constipation, Diarrhoea, and Financial Difficulties), and one is a global health status/quality of life scale. Responses are provided on a 4-point Likert scale (“not at all”, “a little”, “moderate”, “very much”) for all questions except for the global health status and quality of life, which are rated from 1 “very poor” to 7 “excellent”.

### 2.2 EORTC QLU-C10D

The EORTC QLU-C10D [22] comprises two components. The first is a health state classification system based on 13 of 30 items of the EORTC QLQ-C30 [22] that form ten dimensions (Role Functioning, Sleep, Appetite, Social Functioning, Bowel Problems, Emotional Functioning, Pain, Fatigue, Physical Functioning, and Nausea) with the same response options as the EORTC QLQ-C30 (“not at all”, “a little”, “moderate”, “very much”). The second is a series of country-specific preference-based scoring algorithms that allows derivation of QLU-C10D health utilities from QLQ-C30 data. These are derived from valuation studies that use a standardised discrete-choice experiment to elicit the health state preferences of the general population of each country [41]. To date, country-specific tariffs have been developed for six European countries and three English-speaking countries in addition to the UK [42–50], with further tariffs in development in Europe as well as three Asian countries.

### 2.3 EQ-5D-3L

The EQ-5D-3L is an established generic PBM [24, 51] that is frequently used in health economic studies. It comprises five items on generic health issues (Mobility, Self-Care, Usual Activities, Pain/Discomfort, Anxiety/Depression) with three response options each (“no problems”, “some problems”, “severe problems/unable to”) as well as a visual analogue scale (VAS) (0–100) assessing self-rated overall

health. For its use as a PBM, the EuroQol measurement system typically relies on the descriptive system, not on the VAS. Valuation studies for the EQ-5D-3L rely on time trade-off, VASs, or discrete-choice experiments, in general populations [52], guided by the EuroQol Valuation Technique manual. National tariffs are available for a large range of countries and are available at the EuroQol website [52].

### 2.4 Data Sources

For this analysis, we used data from four LUX-Lung studies, multicentre randomised controlled trials (RCTs) that assessed the treatment benefit of afatinib in patients with EGFR-mutated adenocarcinoma (LUX-Lung 1 [LL1], LL3, LL5) [36–38] or patients with squamous cell carcinoma (LL8) [39]. Only adult patients with pathologically diagnosed NSCLC stage IIb or IV were included in these trials. Patients were positive for the EGFR mutation or had squamous cell carcinoma. Most patients were either pre-treated with at least one cycle of chemotherapy and/or a treatment of erlotinib/gefitinib. The intervention arms all included treatments with afatinib, the control arms were either placebo (LUX-Lung 1), cisplatin/pemetrexed (LUX-Lung 3), investigators choice of chemotherapy (LUX-Lung 5), or erlotinib for patients with squamous cell carcinoma (LUX-Lung 8). All trials collected HRQoL data using both the EQ-5D-3L and the EORTC QLQ-C30. The EQ-5D-3L and the EORTC QLQ-C30 were administered at the same time point. Baseline data from patients who completed both EORTC QLU-C10D and EQ-5D-3L measures are included in the current analyses. Further details regarding the original trials can be found elsewhere [36–39].

### 2.5 Statistical Analysis

#### 2.5.1 Data Selection

Data from these four Lux-Lung studies were pooled, giving a total sample of 1736 patients. Patients were included in the current analyses if they had complete data for both the EORTC QLU-C10D and the EQ-5D-3L. For the assessment of core validity parameters, the analysis relied on baseline data only. In a subsequent publication, further psychometric criteria of the EORTC QLU-C10D will also be evaluated using the longitudinal data of the LUX-Lung studies. All analysis were done using the software R [53].

#### 2.5.2 Descriptive Statistics

Sociodemographic characteristics of the pooled sample are descriptively presented as absolute frequencies, means and standard deviations (SDs). EORTC QLU-C10D and

EQ-5D-3L scores were calculated in accordance with the respective valuation studies. Here, we used the utility decrements of Australia [46, 54], Canada [47, 55], Italy [43, 56], the Netherlands [44, 57], Poland [43, 58], and the United Kingdom [49, 59].

### 2.5.3 Criterion Validity

Criterion validity is defined as the association of a health status measure with a gold-standard [60]. It was assessed by correlating the utility scores of the EORTC QLU-C10D with those of the EQ-5D-3L using Pearson correlations. The EQ-5D-3L serves here as a comparator measure as it has been considered a standard in performing health economic evaluations [61]. Criterion validity was considered established by a correlation coefficient of at least 0.7 [35].

### 2.5.4 Construct Validity

Construct validity is defined as the extent to which the results of an instrument are consistent with prespecified hypotheses, such as, but not limited to, the relationships with the results of other instruments [62]. It was evaluated by investigating hypothesised high correlations (convergent validity) and low correlations (divergent validity) of theoretically converging domains (Physical Functioning–Mobility, Role Functioning–Usual Activities, Social Functioning–Usual Activities, Emotional Functioning–Anxiety/Depression, Pain–Pain/Discomfort) or diverging domains (all other domain pairs) of the EORTC QLU-C10D and the EQ-5D-3L using Spearman correlations. The hypothesised theoretically converging and diverging pairs of domains can be found in Table 1. At least 75% of all the hypothesised directions of correlations should be observable in order to establish a good construct validity [35]. The standard classification of weak ( $r = 0.30–0.49$ ), moderate ( $r = 0.50–0.69$ ), and strong ( $r \geq 0.70$ ) [63] was used to categorise the correlations between scores. Furthermore, we plotted Bland-Altman plots that graphically illustrate the scattering of scores across the measurement continuum, as well as the level of agreements (defined as  $1.96 \times$  the SD of the mean scores) of the two measures.

### 2.5.5 Floor and Ceiling Effects

Floor and ceiling effects were estimated as the frequencies of the highest/lowest possible score, once for the utilities and once for each of the domains of the EORTC QLU-C10D and the EQ-5D-3L. Floor and ceiling effects are present if  $> 15\%$  of patients achieve the lowest/highest possible score [35]. The presence of floor or ceiling effects would indicate a reduced reliability of the measures towards the lower/upper end of the scale [35]. When performing valuation studies, floor and ceiling effects potentially limit the correlation and

agreement of scores between the two measures, as the measurement sensitivity in the lower/upper ranges of the scale may be limited for one instrument while the other instrument has sensitivity to health state differences at those extremes.

## 3 Results

### 3.1 Sample Characteristics

A total of 1736 patients were included in this analysis (see Table 2). The different Lux-Lung studies contributed 522 (LL1), 291 (LL3), 194 (LL5), and 729 (LL8) patients to the composition of the sample. The median age was 62 years, and 41% of the patients were female. Out of the 1736 patients, 1040 were assigned to the experimental arms of the RCTs, and 696 patients were in the control arm (see Table 2).

### 3.2 Criterion Validity

The correlation of utilities between the two PBMs ranged from 0.649 for the Dutch country tariff to 0.718 for the Polish country tariff. For all except for one (the Netherlands) investigated countries, correlations between utilities exceeded the predefined threshold of  $r = 0.7$ , constituting a good criterion validity for most country tariffs when calculating EORTC QLU-C10D utilities.

### 3.3 Construct Validity

Correlations between pairs of domains were clearly higher in corresponding domains than in non-corresponding domains but were somewhat lower than expected for most corresponding domains. A pattern different from the other countries was seen using the Dutch weights where the expected high correlation between Emotional Functioning and Anxiety/Depression was low ( $r = 0.147$ ) and the correlation between Pain and Pain/Discomfort was lower than expected and lower than in other countries ( $r = 0.472$ ). For details, see Table 3 and supplementary tables s1–S6 in the electronic supplementary material (ESM).

Assessment of overall agreement between utilities depicted in Bland-Altman plots showed a similar pattern across countries with regards to a certain proportional bias (i.e. score differences are different across the measurement continuum). For all countries, with the exception of the Netherlands, the EORTC QLU-C10D resulted in lower mean utility scores than the EQ-5D-3L (Fig. 1). The mean differences (= blue bias line in the plots) lie between  $-0.08$  (95% confidence interval [CI]  $-0.090$  to  $-0.074$ ) (Canadian

**Table 1** QLU-C10D and EQ-5D domain and health state description analogies

Counterpart domains			
QLU-C10D		EQ-5D-3L	
Domain	Health state descriptions	Domain	Health state descriptions
Physical Functioning	1 No trouble taking a long walk outside of the house 2 No trouble taking a short walk outside of the house, but at least a little trouble taking a long walk 3 At least a little trouble taking a short walk outside of the house, and at least a little trouble taking a long walk 4 Quite a bit or very much trouble taking a short walk outside the house	Mobility	1 No problems in walking about 2 Some problems in walking about 3 Confined to bed
Role Functioning	1 Not at all limited in pursuing work or other daily activities 2 A little limited in pursuing work or other daily activities 3 Quite a bit limited in pursuing work or other daily activities 4 Very much limited in pursuing work or other daily activities	Usual Activities (e.g. work, study, house-work, family, or leisure activities)	1 No problems with performing usual activities 2 Some problems with performing usual activities 3 Unable to perform usual activities
Social Functioning	1 Physical condition or medical treatment interferes not at all with social or family life 2 Physical condition or medical treatment interferes a little with social or family life 3 Physical condition or medical treatment interferes quite a bit with social or family life 4 Physical condition or medical treatment interferes very much with social or family life	Usual Activities (e.g. work, study, house-work, family, or leisure activities)	1 No problems with performing usual activities 2 Some problems with performing usual activities 3 Unable to perform usual activities
Emotional Functioning	1 Not at all feeling depressed 2 Feeling a little depressed 3 Feeling quite a bit depressed 4 Feeling very much depressed	Anxiety/Depression	1 Not anxious or depressed 2 Moderately anxious or depressed 3 Extremely anxious or depressed
Pain	1 No pain 2 A little pain 3 Quite a bit pain 4 Very much pain	Pain/Discomfort	1 No pain or discomfort 2 Moderate pain or discomfort 3 Extreme pain or discomfort
Domains without counterpart			
QLU-C10D		EQ-5D-3L	
Fatigue	1 Not at all tired 2 A little pain 3 Quite a bit pain 4 Very much pain	Self-Care	1 No problems with self-care 2 Some problems washing or dressing 3 Unable to wash or dress
Sleep Disturbance	1 No trouble sleeping 2 A little trouble sleeping 3 Quite a bit trouble sleeping 4 Very much trouble sleeping		
Appetite Loss	1 Not at all lacking appetite 2 A little lacking appetite 3 Quite a bit lacking appetite 4 Very much lacking appetite		
Nausea	1 Not at all feeling nauseated 2 A little feeling nauseated 3 Quite a bit feeling nauseated 4 Very much feeling nauseated		

**Table 1** (continued)

Domains without counterpart	
QLU-C10D	EQ-5D-3L
Bowel Problems	1 No constipation or diarrhoea 2 A little constipation or diarrhoea 3 Quite a bit constipation or diarrhoea 4 Very much constipation or diarrhoea

*QLU-C10D* Quality of Life Utility-Core 10 Dimensions

**Table 2** Data selection and sample characteristics

	Original study sample	Baseline data included in Current analyses	Excluded <sup>a</sup>
LUX-Lung 1	585	522	63
LUX-Lung 3	345	291	54
LUX-Lung 5	202	194	8
LUX-Lung 8	795	729	66
Pooled: HRQoL data at baseline		1736	
Age (median)		62	
Sex (male/female)		1024 (59%)/712 (41%)	
Experimental vs. control <sup>b</sup>		1040/696	

*HRQoL* health-related quality of life, *QLQ* Quality of Life Questionnaire

<sup>a</sup>HRQoL for either *QLQ-C30* and/or *EQ-5D* missing at baseline

<sup>b</sup>Experimental: Afatinib (Lux-Lung 1; Lux-Lung 3; Lux-Lung 8), afatinib + paclitaxel (Lux-Lung 5). Control: Placebo (Lux-Lung 1), cisplatin/pemetrexed (Lux-Lung 3), investigators' choice chemotherapy (Lux-Lung 5), erlotinib (Lux-Lung 8)

tariff) and 0.03 (95% CI 0.0243–0.0417) (Dutch tariff), with SDs between 0.17 and 0.18 (see Table 4). Scrutinising the patterns in the Bland-Altman plots (Figure 1) suggests that high ranging health states show less variability around the bias line than low ranging health states. Hereby, the discrepancy between the utility values varies over the measurement continuum. The differences tend to become larger as the average of the two PBM scores decreases, and sporadic outliers outside the upper and lower level of agreement appear at the lower end of the health state average. Histograms displaying the distribution of mean difference of utility values are provided in the supplementary Figure S1 in the ESM.

### 3.4 Floor and Ceiling Effects

For the EORTC QLU-C10D utilities, between 2.94% and 4.67% reached the highest possible utility value, whereby the Dutch tariff showed the highest proportion of responders (4.67%). Larger ceiling effects were observed for the EQ-5D-3L utilities, which were 22.93% across all country tariffs. For both the EORTC QLU-C10D and the EQ-5D-3L, no floor effect was observed, as the proportion of responders reaching the lowest possible utility value remained below 1 percentage point across all country tariffs (Table 5).

Ceiling and floor effects for specific domains were scrutinised using the EORTC QLU-C10D and the EQ-5D-3L raw data (e.g. without a country tariff applied to the data). The highest ceiling effect in a single domain of the EORTC QLU-C10D raw data was 75.98%, for Nausea. Ceiling effects for the other symptom domains ranged from 25.0% for Fatigue to 55.8% for Bowel Problems. For the functioning domains, the domain with the highest ceiling effect was Emotional Functioning, with 54.6%, with Physical Functioning (25.6%) showing the lowest ceiling effect. Floor effects were not present for the EORTC QLU-C10D, while Physical Functioning (10.0%) showed the highest proportion of responders reaching the lowest possible score. All the other domains remained below 10 percentage points. Note that high response levels across all domains indicate a high level of functioning and a low level of symptom burden, respectively, in this patient population [22].

For the EQ-5D-3L, ceiling effects for Self-Care reached the highest value of all single domains, with 85.8%. For Mobility, the EQ-5D-3L domain which theoretically corresponds to the EORTC QLU-C10D Physical Functioning, the ceiling effect was 61.12%. For the other domains, ceiling effects were 61.1% for Usual Activities, 58.7% for Anxiety/Depression, and 38.3% for Pain/Discomfort, respectively.

**Table 3** Correlations between the EORTC QLU-C10D and EQ-5D-3L utilities and domain scores at baseline

	QLU-C10D	EQ-5D-3L	Expected correlation	AUS	CAN	IT	NL	PL	UK
Pearson's <i>r</i>	Utilities	Utilities	Strong	0.712	0.709	0.706	0.649	0.718	0.706
Spearman's $\rho$	Physical Functioning	Mobility	Strong, moderate	0.543	0.540	0.543	0.540	0.543	0.540
	Role Functioning	Usual Activities	Strong, moderate	0.612	0.612	0.612	0.612	0.612	0.612
	Social Functioning	Usual Activities	Moderate	0.364	0.376	0.376	0.376	0.364	0.376
	Emotional Functioning	Anxiety/Depression	Strong	0.547	0.547	0.359	0.147	0.547	0.547
	Pain	Pain/Discomfort	Strong	0.679	0.679	0.679	0.472	0.679	0.679
	Fatigue <sup>a</sup>	All	Weak	0.391	0.387	0.391	0.336	0.387	0.391
	Sleep Disturbances <sup>a</sup>	All	Weak	0.315	0.315	0.315	0.315	0.318	0.318
	Appetite Loss <sup>a</sup>	All	Weak	0.345	0.348	0.324	0.345	0.348	0.324
	Nausea <sup>a</sup>	All	Weak	0.246	0.246	0.246	0.246	0.246	0.246
	Bowel Problems <sup>a</sup>	All	Weak	0.260	0.260	0.260	0.260	0.260	0.260

AUS Australia, CAN Canada, EORTC European Organisation for Research and Treatment of Cancer, IT Italy, NL the Netherlands, PL Poland, QLU-C10D Quality of Life Utility-Core 10 Dimensions, UK United Kingdom

<sup>a</sup>For EORTC QLU-C10D domains without an EQ-5D counterpart, the highest observed correlation is displayed (divergent validity)

No floor effect was found for the EQ-5D-3L, whereby Pain/Discomfort (5.5%) showed the highest proportion of responders reaching the lowest possible value. All other domains remained below 5 percentage points. For further details see Fig. 2. Floor and ceiling effects on the domain level correspond to the lowest and highest item scores for single-item domains and to the lowest and highest domain scores for multi-item domains (EORTC QLU-C10D: Physical Functioning, Social Functioning, and Bowel Problems), respectively.

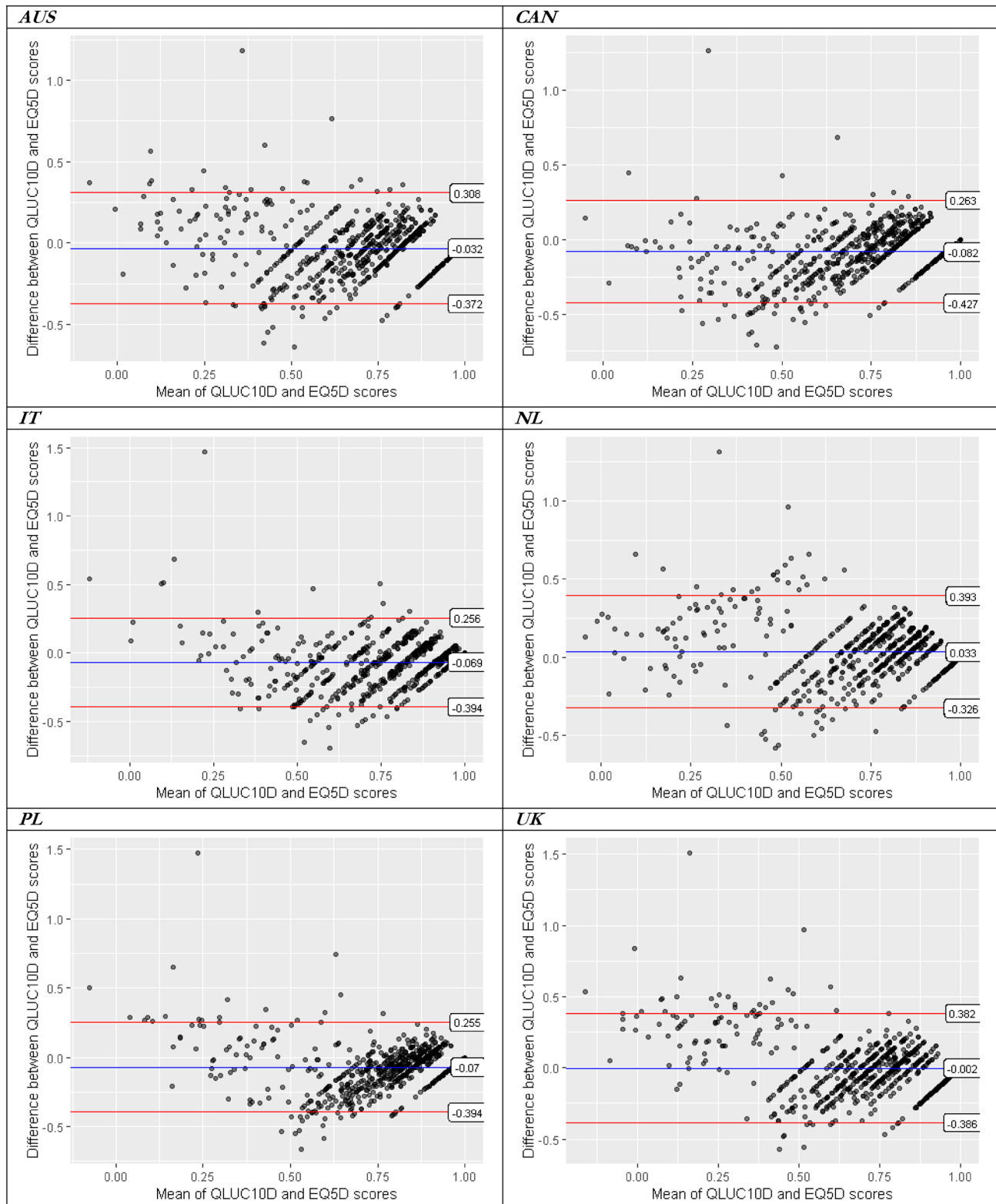
## 4 Discussion

The EORTC QLU-C10D is a recently developed preference-based scoring algorithm for the widely used EORTC QLQ-C30, designed to derive utilities that could inform health economic evaluations in cancer patient populations [22, 40]. In this article, we aimed to report on three psychometric properties, criterion validity, construct validity, and ceiling and floor effects in a lung cancer population; these concepts are widely used in psychology, but are similarly important when determining the appropriateness of instruments to support HTA. Here, the widely used EQ-5D-3L [24, 51] served as a comparator measure.

Overall, our findings further support that the EORTC QLU-C10D has good criterion and construct validity, confirming previous findings [64–68]. Hereby, hypothetically corresponding and diverging domains were defined a priori and correlations of the utilities and the domain scores between the two measures were estimated. Correlations of the utilities were above 0.7 for all but one utility decrement set. The threshold of 0.7 for the correlation coefficient is suggested to establish the criterion validity [35]. For the

hypothetically corresponding domains, higher correlations were observed than for the diverging domains. The highest correlation was found for the domain pair Pain–Pain/Discomfort ( $r = 0.679$ ). All other hypothetically corresponding domains had a higher correlation than 0.5 except for the Social Functioning–Usual Activities domain pair (max.  $r = 0.376$  across all value sets) and for Emotional Functioning–Anxiety/Depression as well as Pain–Pain/Discomfort when using the Dutch utility decrements. All the hypothetically diverging domain pairs reached a correlation coefficient of maximum 0.391.

When scrutinising the Bland-Altman plots, i.e. the agreement between the measures, the systematically lower utility scores (except for the Netherlands) of the EORTC QLU-C10D are noticeable. This is in line with findings reported for utilities of patients with neuroendocrine tumours [69], for patients undergoing esophagectomy [64], and for the QALYs of patients undergoing laryngectomy [70]. In our study, the mean difference between QLU-C10D utilities and EQ-5D-3L utilities ranged from 0.033 for the Netherlands to  $-0.082$  for Canada. As a crude measure to interpret the difference, the minimal clinically important difference (MCID) of the EQ-5D-3L in cancer patients served as a reference. The MCID lies between 0.07 and 0.08 for lung cancer patients using the United Kingdom value sets [71]. Thus, only for one country (namely Canada) the mean difference of utility scores exceeded the MCID. This indicates that only for the Canadian tariff there is a mean measurement difference that goes beyond a threshold that is considered “minimally important”. Additionally, to the mean score difference, the difference between the utilities varied across the measurement continuum of the two PBMs. At the low and mid range of the scales, the difference is larger (both positive and negative differences), whereby these differences appear



**Fig. 1** Bland-Altman plots displaying the level of agreement between the EORTC QLU-C10D and the EQ-5D-3L for different country tariffs. The blue line indicates the mean difference of EORTC QLU-C10D and EQ-5D-3L utility scores. Red lines indicate the level of

agreement of the utility scores of the two instruments. *AUS* Australia, *CAN* Canada, *EORTC* European Organisation for Research and Treatment of Cancer, *IT* Italy, *NL* the Netherlands, *PL* Poland, *QLU-C10D* Quality of Life Utility-Core 10 dimensions, *UK* United Kingdom



**Table 4** Agreement between the EORTC QLU-C10D and the EQ-5D-3L

Country tariff	Mean difference [95% CI]	SD	Upper LoA <sup>a</sup>	Lower LoA <sup>a</sup>
AUS	-0.032 [-0.040 to -0.024]	0.173	0.308	-0.372
CAN	-0.082 [-0.090 to -0.074]	0.176	0.263	-0.427
IT	-0.069 [-0.077 to -0.061]	0.166	0.256	-0.394
NL	0.033 [0.0243 to 0.0417]	0.184	0.393	-0.326
PL	-0.070 [-0.078 to -0.062]	0.166	0.255	-0.394
UK	-0.002 [-0.011 to 0.00722]	0.196	0.382	-0.386

AUS Australia, CAN Canada, CI confidence interval, EORTC European Organisation for Research and Treatment of Cancer, IT Italy, LoA level of agreement, NL the Netherlands, PL Poland, QLU-C10D Quality of Life Utility-Core 10 dimensions, SD standard deviation, UK United Kingdom

<sup>a</sup>LoA defined as mean  $\pm$  1.96  $\times$  SD

**Table 5** Ceiling and floor effects of the EORTC QLU-C10D and the EQ-5D-3L utilities

Country tariff	QLU-C10D		EQ-5D-3L	
	Floor (%)	Ceiling (%)	Floor (%)	Ceiling (%)
AUS	0.06	3.57	0.06	22.93
CAN	0.06	2.94	0.23	22.93
IT	0.06	3.23	0.23	22.93
NL	0.06	4.67	0.23	22.93
PL	0.06	3.57	0.23	22.93
UK	0.06	2.94	0.23	22.93

AUS Australia, CAN Canada, EORTC European Organisation for Research and Treatment of Cancer, IT Italy, NL the Netherlands, PL Poland, QLU-C10D Quality of Life Utility-Core 10 dimensions, UK United Kingdom

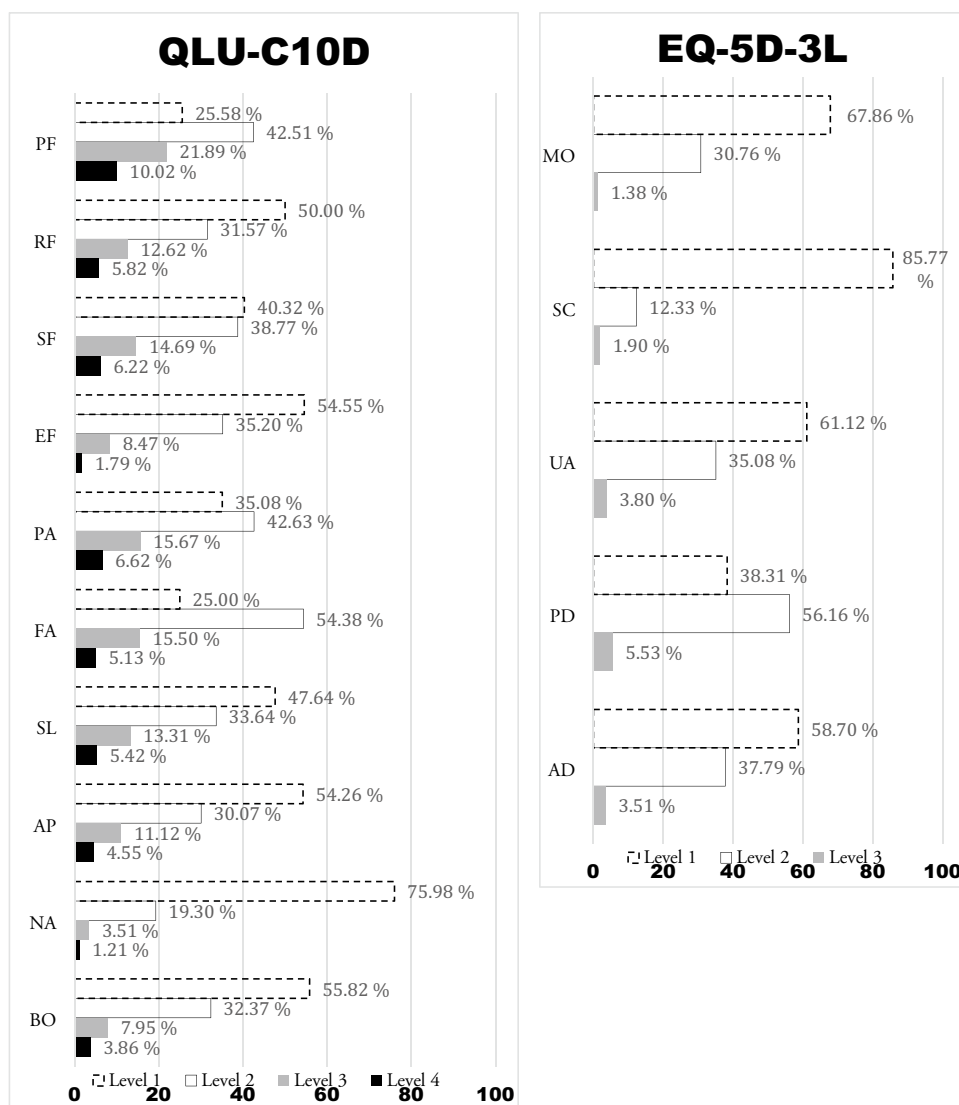
to become smaller at the upper end of the scale (for patients in good health states). This suggests that a proportional bias may be present when comparing EORTC QLU-C10D and EQ-5D-3L scores, as found for patients with gastric cancer [67]. Lastly, the minimal floor and ceiling effects (highest ceiling effect was 4.67% using the Dutch utility decrements) contribute to the claim that the EORTC QLU-C10D has good measurement properties, also in the upper and lower continuum of the questionnaire.

A few findings are worth mentioning further. Firstly, while the convergent validity of hypothetically corresponding domain pairs was not as strong as hypothesised, correlations were mostly higher for convergent domains compared to divergent domains. Secondly, several convergent correlations were smaller than expected. Specifically, we were surprised that the correlation between the EQ-5D Mobility domain and QLU-C10D Physical Functioning domain was moderate rather than high, given that the latter is essentially about mobility (ability to take long and short walks). Perhaps this is because the QLU-C10D taps the higher end of this construct (highest level is “no trouble taking a long walk” versus the EQ-5D’s “no problems in walking about”), while the EQ-5D taps the lower end of this construct (lowest level

is “I am confined to bed” versus the QLU-C10D’s “quite a bit or very much trouble taking a short walk outside the house”). Also, the correlations between the Social Functioning and Usual Activities domains across all countries were weak rather than the moderate we hypothesised. This might partly be explained by the varying concepts assessed by these domains; where the EORTC QLU-C10D inquires about “interference with social or family life”, the EQ-5D-3L refers to the “performance of usual activities”, which are distinct concepts in the international classification of functioning [72, 73]. Furthermore, data were collected in a hospital-based setting, which might influence the patients understanding of “normal” (social) activities, especially considering response shift [74]. Additionally, a weak correlation between the Emotional Functioning and Anxiety/Depression domains was observed when using the Italian and Dutch utility weights. Partly, this might again be explained by the differing concepts investigated, where the EORTC QLU-C10D inquires about depression alone, while the EQ-5D-3L assesses anxiety as well as depression. In addition, a high influence of the country-specific scoring algorithm must be surmised as the weak correlation is only present for the Italian and Dutch country tariffs. Similar findings were reported for the Italian tariffs in patients with myelodysplastic syndrome [65] and for Dutch cancer patients [75].

It is worth noting that the correlation coefficients and the agreement between the measures vary somewhat when using different country-specific utility decrements. For example, only for the Dutch tariffs does the EORTC QLU-C10D produce higher utility values compared to the EQ-5D-3L, as previously found in the Dutch cancer population. Considering the consistent methodology and the concurrent process when developing the various EORTC QLU-C10D country-specific value sets (valuation studies) [42–50], it becomes apparent that the factors “methodology” and “timelines” do not influence the variability of utility decrements. Subsequently, cultural preferences for certain health states or semantic differences in the translated versions of the questionnaires are possible influencing factors in the evaluation

**Fig. 2** Relative frequencies per response level for the EORTC QLU-C10D and the EQ-5D-3L. *AD* Anxiety/Depression, *AP* Appetite, *BO* Bowel Problems, *EF* Emotional Functioning, *EORTC* European Organisation for Research and Treatment of Cancer, *FA* Fatigue, *MO* Mobility, *NA* Nausea, *PA* Pain, *PD* Pain/Discomfort, *PF* Physical Functioning, *QLU-C10D* Quality of Life Utility-Core 10 dimensions, *RF* Role Functioning, *SC* Self-Care, *SF* Social Functioning, *SL* Sleep, *UA* Usual Activities



of the preference-based health states [76, 77]. Therefore, the observed differences across the various value sets are likely to be genuine.

Taking this into consideration, a further discussion can be raised when performing CUA. Hereby, the current state of the art is to apply the value sets from the decision makers country on obtained data, even when the data are collected in multicentric trials or entirely from another country [78]. Still our discussion highlights that values sets and data from one-and-the-same country might be the best match taking cultural and linguistic aspects into consideration. Thus, we would like to argue that proper justification has to be provided when applying weights from one country on data obtained from another country.

The main limitation of our study is its retrospective nature, i.e. it uses data from studies that were not designed to assess the validity of the QLU-C10D. Although we drew on robust data from international trials, the LUX-Lung studies

mainly included patients with advanced (Union internationale contre le cancer (UICC) stage III and IV) lung cancer. Thus, our finding relates to the validity of the EORTC QLU-C10D for this patient population and may not generalise to other cancer types and early-stage lung cancer. Also, the validation relied on the comparison to the EQ-5D-3L only, whereby alternative generic PBMs, such as the EQ-5D-5L [79], the SF-6D [21], or the Health Utility Index [80] were not available. The 5-level version of the EQ-5D is known to have improved psychometric properties compared to the 3-level version (such as lowered ceiling effects) in general populations [81] and in cancer patients [82] and is now among the most widely used PBMs worldwide [83]. Thus, having the 5-level version of the instrument as a comparator is highly recommended in future studies [68]. This will bring light as to what drives differences between the instruments, the measurement range only or also domain content.

We acknowledge that this paper addresses a narrow range of validity types but note that these are three core measurement properties of health state measures [35]. We maintained this focus in order to also investigate the impact of applying different country tariffs to this set of validity types. Assessing clinical validity, i.e. sensitivity to clinical differences between groups, and predictive validity (responsiveness) for health state changes was beyond the scope of this paper. The clinical validity, such as sensitivity to known groups and predictive validity for change, of the QLU-C10D has been demonstrated previously in a number of contexts and patient population [64–68], but not yet lung cancer patients. In a subsequent paper, we will utilise the LUX-Lung data to assess these important aspects of validity in lung cancer. Furthermore, algorithms to map EORTC QLU-C10D scores to the EuroQol measurement system would be useful so should be developed in future studies.

## 5 Conclusion

The EORTC QLQ-C30, which the EORTC QLU-C10 relies on, was originally developed and validated in a lung cancer population [40]. This fact underlines the importance of the analysis in the lung cancer population, whereby the hypothesised good criterion and construct validity of the EORTC QLU-C10D in the lung cancer population could be demonstrated. Even though the presented results showed significant proof for strong criterion and construct validity, further investigations must be undertaken to evaluate the clinical validity of the EORTC QLU-C10D in lung cancer and other cancer patient populations. Furthermore, the development of cross-walks or score mappings between the EORTC QLU-C10D and other PBMs [68] is necessary to further aid the comparability of the scores in health economic evaluations.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s41669-024-00484-9>.

**Acknowledgments** This publication is based on research using data from data contributors Boehringer Ingelheim that has been made available through Vivli, Inc. Vivli has not contributed to or approved and is not in any way responsible for the contents of this publication. The research in this article was done during AB's tenure at EORTC. AB is no longer affiliated with EORTC.

## Declarations

**Funding** This work was supported by an EORTC grant (#12/2016). The EORTC Quality of Life Group business model involves charges for commercial companies using EORTC instruments. Academic use of EORTC instruments is free of charge.

**Conflict of Interest** The authors have no conflict of interest to declare.

**Ethics Approval** Ethical approval for this study was provided by the ethics board of the Medical University of Innsbruck as an amendment to the study AN 2015-0016.

**Consent to Participate** All patients provided written informed consent when participating in the original trials this retrospective analysis relies on. The original study protocols were developed in accordance with the Declaration of Helsinki, and ethical approval was sought at participating centres.

**Consent for Publication** Not applicable

**Availability of Data and Material** The results presented in this manuscript rely on the secondary analysis of trial data proprietary of the data contributor Boehringer Ingelheim. The authors of this article do not have the permission to make the original trial data or any other trial materials available for public use.

**Code Availability** Not applicable.

**Authors' Contributions** Conceptualisation: Gamper E.M., King M.K., Norman R. Acquisition of data: Pilz M.J., Gamper E.M. Analysis and interpretation of data: Pilz M.J., Gamper E.M., Seyringer S., Al-Naesan I. Drafting of the manuscript: Pilz M.J., Seyringer S., Al-Naesan I., Gamper E.M. Critical revision of the manuscript: Bottomley A., King M.K., Norman R., Gamper E.M. Statistical analysis: Schlosser L., Hell T. Provision of study materials or patients: Bottomley A., King M.K., Norman R. Obtaining funding: Gamper E.M. Administrative, technical, or logistic support: King M.K., Norman R., Gamper E.M. Supervision: Gamper E.M. All authors have read and approved the manuscript.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

## References

1. International Agency for research on cancer. Cancer today: Estimated number of new cases in 2020, World, both sexes, all ages (excl. NMSC). [https://gco.iarc.fr/today/online-analysis-table?v=2020&mode=cancer&mode\\_population=continents&population=900&populations=900&key=asr&sex=0&cancer=39&type=0&statistic=5&prevalence=0&population\\_group=0&ages\\_group%5B%5D=0&ages\\_group%5B%5D=17&group\\_cancer=1&include\\_nmssc=0&include\\_nmssc\\_other=1](https://gco.iarc.fr/today/online-analysis-table?v=2020&mode=cancer&mode_population=continents&population=900&populations=900&key=asr&sex=0&cancer=39&type=0&statistic=5&prevalence=0&population_group=0&ages_group%5B%5D=0&ages_group%5B%5D=17&group_cancer=1&include_nmssc=0&include_nmssc_other=1). Accessed 20 Jan 2023.
2. Schwartz AG, Cote ML. Epidemiology of lung cancer. *Adv Exp Med Biol.* 2016;893:21–41.
3. Migliorino MR, Santo A, Romano G, Cortinovis D, Galetta D, Alabiso O, Carteni G, Vari S, Fasola G, Pazzola A, Giuffrida D, Zaniboni A, Caprioli A, Longo F, Acciai V, de Marinis F. Economic burden of patients affected by non-small cell lung cancer (NSCLC): the LIFE study. *J Cancer Res Clin Oncol.* 2017;143(5):783–91.

4. Gridelli C, Rossi A, Carbone DP, Guarize J, Karachaliou N, Mok T, Petrella F, Spaggiari L, Rosell R. Non-small-cell lung cancer. *Nat Rev Dis Primers*. 2015;21(1):15009.
5. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin*. 2018;68(1):7–30.
6. Eberhardt WE, De Ruyscher D, Weder W, Le Péchoux C, De Leyn P, Hoffmann H, Westeel V, Stahel R, Felip E, Peters S, Panel Members. 2nd ESMO Consensus Conference in Lung Cancer: locally advanced stage III non-small-cell lung cancer. *Ann Oncol*. 2015;26(8):1573–1588.
7. Al Dayel F. EGFR mutation testing in non-small cell lung cancer (NSCLC). *J Infect Public Health*. 2012;5 Supplement:S31–34.
8. Zhang YL, Yuan JQ, Wang KF, Fu XH, Han XR, Threapleton D, Yang ZY, Mao C, Tang JL. The prevalence of EGFR mutation in patients with non-small cell lung cancer: a systematic review and meta-analysis. *Oncotarget*. 2016;48(7):78985–93.
9. Aguiar F, Fernandes G, Queiroga H, Machado JC, Cirnes L, Souto Moura C, Hespanhol V. Overall survival analysis and characterization of an EGFR mutated non-small cell lung cancer (NSCLC) population. *Arch Bronconeumol*. 2018;54(1):10–7.
10. Hsu WH, Yang JC, Mok TS, Loong HH. Overview of current systemic management of EGFR-mutant NSCLC. *Ann Oncol*. 2018;1(Suppl. 1):i3–9.
11. Ding PN, Lord SJ, GebSKI V, Links M, Bray V, Gralla RJ, Yang JC, Lee CK. Risk of treatment-related toxicities from EGFR tyrosine kinase inhibitors: a meta-analysis of clinical trials of gefitinib, erlotinib, and afatinib in advanced EGFR-mutated non-small cell lung cancer. *J Thorac Oncol*. 2017;12(4):633–43.
12. Normando SR, Cruz FM, Del Giglio A. Cumulative meta-analysis of epidermal growth factor receptor-tyrosine kinase inhibitors as first-line therapy in metastatic non-small-cell lung cancer. *Anti-cancer Drugs*. 2015;26(9):995–1003.
13. Zhang Y, Sheng J, Yang Y, Fang W, Kang S, He Y, Hong S, Zhan J, Zhao Y, Xue C, Ma Y, Zhou T, Ma S, Gao F, Qin T, Hu Z, Tian Y, Hou X, Huang Y, Zhou N, Zhao H, Zhang L. Optimized selection of three major EGFR-TKIs in advanced EGFR-positive non-small cell lung cancer: a network meta-analysis. *Oncotarget*. 2016;12(7):20093–108.
14. Sequist LV, Yang JC, Yamamoto N, O'Byrne K, Hirsh V, Mok T, Geater SL, Orlov S, Tsai CM, Boyer M, Su WC, Bannouna J, Kato T, Gorbunova V, Lee KH, Shah R, Massey D, Zazulina V, Shahidi M, Schuler M. Phase III study of afatinib or cisplatin plus pemetrexid in patients with metastatic lung adenocarcinoma with EGFR mutations. *J Clin Oncol*. 2013;31(27):3327–34.
15. Wu YL, Zhou C, Hu CP, Feng J, Lu S, Huang Y, Li W, Hou M, Shi JH, Lee KY, Xu CR, Massey D, Kim M, Shi Y, Geater SL. Afatinib versus cisplatin plus gemcitabine for first-line treatment of Asian patients with advanced non-small-cell lung cancer harbouring EGFR mutations (LUX-Lung 6): an open-label, randomised phase 3 trial. *Lancet Oncol*. 2014;15(2):213–22.
16. Yang JC, Wu YL, Schuler M, Sebastian M, Popat S, Yamamoto N, Zhou C, Hu CP, O'Byrne K, Feng J, Lu S, Huang Y, Geater SL, Lee KY, Tsai CM, Gorbunova V, Hirsh V, Bannouna J, Orlov S, Mok T, Boyer M, Su WC, Lee KH, Kato T, Massey D, Shahidi M, Zazulina V, Sequist LV. Afatinib versus cisplatin-based chemotherapy for EGFR mutation-positive lung adenocarcinoma (LUX-Lung 3 and LUX-Lung 6): analysis of overall survival data from two randomised, phase 3 trials. *Lancet Oncol*. 2015;16(2):141–51.
17. Griebisch I, Palmer M, Fayers PM, Ellis S. Is progression-free survival associated with a better health-related quality of life in patients with lung cancer? Evidence from two randomised trials with afatinib. *BMJ Open*. 2014;4(10):1–9.
18. Pratt-Chapman MBA. Patient-reported outcomes in health economic decision-making: a changing landscape in oncology. *Recent Results Cancer Res*. 2019;213:67–83.
19. Trippoli S, Vaiani M, Lucioni C, Messori A. Quality of life and utility in patients with non-small cell lung cancer. *Quality-of-life Study Group of the Master 2 Project in Pharmacoeconomics*. *Pharmacoeconomics*. 2001;19(8):855–63.
20. McDonough CMTA. Measuring preferences for cost-utility analysis: how choice of method may influence decision-making. *Pharmacoeconomics*. 2007;25(2):93–106.
21. Brazier J, Usherwood T, Harper R, Thomas K. Deriving a preference-based single index from the UK SF-36 Health Survey. *J Clin Epidemiol*. 1998;51(11):1115–28.
22. King MT, Costa DS, Aaronson NK, Brazier JE, Cella DF, Fayers PM, Grimison P, Janda M, Kemmler G, Norman R, Pickard AS, Rowen D, Velikova G, Young TA, Viney R. 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.
23. Roudijk B, Donders ART, Stalmeier PFM. Setting dead at zero: applying scale properties to the QALY model. *Med Decis Making*. 2016;38(6):627–34.
24. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med*. 2001;33(5):337–43.
25. Feeny D, Furlong W, Boyle M, Torrance GW. Multi-attribute health status classification systems. *Health Utilities Index*. *Pharmacoeconomics*. 1995;7(6):490–502.
26. Rowen D, Brazier J, Ara R, Azzabi ZI. The role of condition-specific preference-based measures in health technology assessment. *Pharmacoeconomics*. 2017;35(Suppl. 1):33–41.
27. Drummond M, Sculpher M. Common methodological flaws in economic evaluations. *Med Care*. 2005;43(Suppl. 7):5–14.
28. Goodwin EGC. A systematic review of the literature on the development of condition-specific preference-Based Measures of Health. *Appl Health Econ Health Policy*. 2016;14(2):161–83.
29. Langworth L, Yang Y, Young T, Mulhern B, Hernández Alava M, Mukuria C, Rowen D, Tosh J, Tsuchiya A, Evans P, Devianee Keetharuth A, Brazier J. 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;9(18):1–224.
30. Versteegh MM, Leunis A, Uyl-de Groot CA, Stolk EA. Condition-specific preference-based measures: benefit or burden? *Value Health*. 2012;15(3):504–13.
31. Neumann PJ, Goldie SJ, Weinstein MC. Preference-based measures in economic evaluation in health care. *Annu Rev Public Health*. 2000;21:587–611.
32. Brazier JE, Rowen D, Mavranzeouli I, Tsuchiya A, Young T, Yang Y, Barkham M, Ibbotson R. Developing and testing methods for deriving preference-based measures of health from condition-specific measures (and other patient-based measures of outcome). *Health Technol Assess*. 2016;16(32):1–114.
33. Richardson J, Iezzi A, Khan MA, Chen G, Maxwell A. Measuring the sensitivity and construct validity of 6 utility instruments in 7 disease areas. *Med Decis Making*. 2016;36(2):147–95.
34. Pettitt D, Raza S, Naughton B, Roscoe A, Ramakrishnan A, Ali A, Davies B, Dopson S, Hollander G, Smith J, Brindley D. The limitations of QALY: a literature review. *J Stem Cell Res Ther*. 2016;6(4):1–7.
35. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34–42.
36. Miller VA, Hirsh V, Cadranel J, Chen YM, Park K, Kim SW, Zhou C, Su WC, Wang M, Sun Y, Heo DS, Crino L, Tan EH, Chao TY, Shahidi M, Cong XJ, Lorence RM, Yang JC. Afatinib versus placebo for patients with advanced, metastatic non-small-cell lung cancer after failure of erlotinib, gefitinib, or both, and one or two

- lines of chemotherapy (LUX-Lung 1): a phase 2b/3 randomised trial. *Lancet Oncol.* 2013;13(5):528–38.
37. Sequist LV, Yang JC, Yamamoto N, O'Byrne K, Hirsh V, Mok T, Geater SL, Orlov S, Tsai CM, Boyer M, Su WC, Bennouna J, Kato T, Gorbunova V, Lee KH, Shah R, Massey D, Zazulina V, Shahidi M, Schuler M. Phase III study of afatinib or cisplatin plus pemetrexed in patients with metastatic lung adenocarcinoma with EGFR mutations. *J Clin Oncol.* 2013;31(27):3327–34.
  38. Schuler M, Yang JC, Park K, Kim JH, Bennouna J, Chen YM, Chouaid C, De Marinis F, Feng JF, Grossi F, Kim DW, Liu X, Lu S, Strausz J, Vinnyk Y, Wiewrodt R, Zhou C, Wang B, Chand VK, Planchard D, LUX-Lung 5 Investigators. Afatinib beyond progression in patients with non-small-cell lung cancer following chemotherapy, erlotinib/gefitinib and afatinib: phase III randomized LUX-Lung 5 trial. *Ann Oncol.* 2016;27(3):417–23.
  39. Soria JC, Felip E, Cobo M, Lu S, Syrigos K, Lee KH, Göker E, Georgoulas V, Li W, Isla D, Guclu SZ, Morabito A, Min YJ, Ardizzoni A, Gadgeel SM, Wang B, Chand VK, Goss GD, LUX-Lung 8 Investigators. Afatinib versus erlotinib as second-line treatment of patients with advanced squamous cell carcinoma of the lung (LUX-Lung 8): an open-label randomised controlled phase 3 trial. *Lancet Oncol.* 2015;16(8):897–907.
  40. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, Filiberti A, Flechtner H, Fleishman SB, de Haes JC, 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.
  41. Gamper E-M, King MT, Norman R, Loth FLC, Holzner B, Kemmler G. The EORTC QLU-C10D discrete choice experiment for cancer patients: a first step towards patient utility weights. *J Patient Rep Outcomes.* 2022;6(1):42. <https://doi.org/10.1186/s41687-022-00430-5>.
  42. Finch AP, Gamper E, Norman R, Viney R, Holzner B, King M, Kemmler G, EORTC Quality of Life Group. Estimation of an EORTC QLU-C10 value set for Spain using a discrete choice experiment. *Pharmacoeconomics.* 2021;39(9):1085–98.
  43. Gamper EM, King MT, Norman R, Efficace F, Cottone F, Holzner B, Kemmler G. EORTC QLU-C10D value sets for Austria, Italy, and Poland. *Qual Life Res.* 2020;29(9):2485–95.
  44. Jansen F, Verdonck-de Leeuw IM, Gamper E, Norman R, Holzner B, King M, Kemmler G. Dutch utility weights for the EORTC cancer-specific utility instrument: the Dutch EORTC QLU-C10D. *Qual Life Res.* 2021;30(7):2009–19.
  45. Kemmler G, Gamper E, Nerich V, Norman R, Viney R, Holzner B, King M. 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.
  46. King MT, Viney R, Simon Pickard A, Rowen D, Aaronson NK, Brazier JE, Cella D, Costa DSJ, Fayers PM, Kemmler G, McTaggart-Cowan H, Mercieca-Bebber R, Peacock S, Street DJ, Young TA, Norman R. 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.
  47. McTaggart-Cowan H, King MT, Norman R, Costa DSJ, Pickard AS, Regier DA, Viney R, Peacock SJ. 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):1–11.
  48. Nerich V, Gamper EM, Norman R, King M, Holzner B, Viney R, Kemmler G. 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.
  49. Norman R, Mercieca-Bebber R, Rowen D, Brazier JE, Cella D, Pickard AS, Street DJ, Viney R, Revicki D, King MT, European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Group and the MAUCa Consortium. U.K. utility weights for the EORTC QLU-C10D. *Health Econ.* 2019;28(12):1385–401.
  50. Revicki DA, King MT, Viney R, Pickard AS, Mercieca-Bebber R, Shaw JW, Müller F, Norman R. United States utility algorithm for the EORTC QLU-C10D, a multiattribute utility instrument based on a cancer-specific quality-of-life instrument. *Med Decis Making.* 2021;41(4):485–501.
  51. EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy.* 1990;16(2):199–208.
  52. EuroQol Group. EQ-5D-3L | Valuation: Overview of available standard value sets for EQ-5D-3L. <https://euroqol.org/eq-5d-instruments/eq-5d-3l-about/valuation/>. Accessed 18 Jan 2023.
  53. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, Vienna, Austria; 2021.
  54. Viney R, Norman R, King MT, Cronin P, Street DJ, Knox S, Ratcliffe J. Time trade-off derived EQ-5D weights for Australia. *Value Health.* 2011;14(6):928–36.
  55. Bansback N, Tsuchiya A, Brazier J, Anis A. Canadian valuation of EQ-5D health states: preliminary value set and considerations for future valuation studies. *PLoS ONE.* 2012;7(2): e31115.
  56. Scalone L, Cortesi PA, Ciampichini R, Belisari A, D'Angiolella LS, Cesana G, Mantovani LG. Italian population-based values of EQ-5D health states. *Value Health.* 2013;16(5):814–22.
  57. Lamers LM, McDonnell J, Stalmeier PF, Krabbe PF, Busschbach JJ. The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. *Health Econ.* 2006;15(10):1121–31.
  58. Golicki D, Jakubczyk M, Niewada M, Wrona W, Busschbach JJ. Valuation of EQ-5D health states in Poland: first TTO-based social value set in Central and Eastern Europe. *Value Health.* 2010;13(2):289–97.
  59. Dolan P. Modeling valuations for EuroQol health states. *Med Care.* 1997;35(11):1095–108.
  60. Prinsen CAC, Morkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res.* 2018;27(5):1147–57. <https://doi.org/10.1007/s11136-018-1798-3>.
  61. Devlin NJBR. EQ-5D and the EuroQol group: past, present and future. *Appl Health Econ Health Policy.* 2017;12(2):127–37.
  62. COSMIN. COSMIN definitions of domains, measurement properties, and aspects of measurement properties. <https://cosmin.nl/wp-content/uploads/COSMIN-definitions-domains-measurement-properties.pdf>. Accessed 12 Aug 2023
  63. Akoglu H. User's guide to correlation coefficients. *Turk J Emerg Med.* 2018;18(3):91–3.
  64. Bulamu NB, Vissapragada R, Chen G, Ratcliffe J, Mudge LA, Smithers BM, Isenring EA, Smith L, Jamieson GG, Watson DI. 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.
  65. Gamper EM, Cottone F, Sommer K, Norman R, King M, Brecchia M, Caocci G, Patriarca A, Palumbo GA, Stauder R, Niscola P, Platzbecker U, Caers J, Vignetti M, Efficace F. 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.
  66. 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.

67. 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.* 2022;(9):885-893.
68. Shaw JW, Bennett B, Trigg A, DeRosa M, Taylor F, Kiff C, Ntais D, Noon K, King MT, Cocks K. 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.
69. Soare IA, Leeuwenkamp O, Longworth L. Estimation of Health-Related Utilities for 177Lu-DOTATATE in GEP-NET Patients Using Utilities Mapped from EORTC QLQ-C30 to EQ-5D-3L and QLU-C10D Utilities. *Pharmacoeconomics Open.* 2021;5(4):715-25.
70. Jansen F, Coupé VMH, Eerenstein SEJ, Cnossen IC, van Uden-Kraan CF, de Bree R, Doornaert P, Halmos GB, Hardillo JAU, van Hinte G, Honings J, Leemans CR, Verdonck-de Leeuw IM. Cost-utility and cost-effectiveness of a guided self-help head and neck exercise program for patients treated with total laryngectomy: results of a multi-center randomized controlled trial. *Oral Oncol.* 2021;117:1-9.
71. 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;21(5):70.
72. Cieza A, Ewert T, Ustün TB, Chatterji S, Kostanjsek N, Stucki G. Development of ICF Core Sets for patients with chronic conditions. *J Rehabil Med.* 2004;44 Suppl.:9-11.
73. Pilz MJ, Rothmund M, Lidington E, Piccinin C, Arraras JJ, Groenvold M, Holzner B, van Leeuwen M, Petersen MA, Schmidt H, Young T, Giesinger JM, EORTC Quality of Life Group. Content comparison of the EORTC CAT Core, SF-36, FACT-G, and PROMIS role and social functioning measures based on the International Classification of Functioning, Disability and Health. *Psycho-oncology.* 2023;32(9):1372-84. <https://doi.org/10.1002/pon.6188>
74. Vanier A, Oort FJ, McClimans L, Ow N, Gulek BG, Böhnke JR, Sprangers M, Sébille V, Mayo N, Response Shift - in Sync Working Group. Response shift in patient-reported outcomes: definition, theory, and a revised model. *Qual Life Res.* 2021;30(12):3309-22.
75. Pilz MJ, Seyringer S, Hallsson LR, Bottomley A, Jansen F, King MT, Norman R, Rutten MJ, Leeuw IMV, Siersema PD, Gamper EM. 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. <https://doi.org/10.1007/s10198-024-01670-6>
76. Koller M, Müller K, Nolte S, Schmidt H, Harvey C, Mölle U, Boehm A, Engeler D, Metzger J, Sztankay M, Holzner B, Groenvold M, Kuliš D, Bottomley A, EORTC Quality of Life Group. Investigating the response scale of the EORTC QLQ-C30 in German cancer patients and a population survey. *Health Qual Life Outcomes.* 2021;19(9):235.
77. Scott NW, Fayers PM, Bottomley A, Aaronson NK, de Graeff A, Groenvold M, Koller M, Petersen MA, Sprangers MA, EORTC and the Quality of Life Cross-Cultural Meta-Analysis Group. Comparing translations of the EORTC QLQ-C30 using differential item functioning analyses. *Qual Life Res.* 2006;15(6):1103-15.
78. Sonntag M, König HH, Konnopka A. The estimation of utility weights in cost-utility analysis for mental disorders: a systematic review. *Pharmacoeconomics.* 2013;31(12):1131-54.
79. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G, Badia X. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res.* 2011;20(10):1727-36.
80. Marra CA, Woolcott JC, Kopec JA, Shojania K, Offer R, Brazier JE, Esdaile JM, Anis AH. A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. *Soc Sci Med.* 2005;60(7):1571-82.
81. Agborsangaya CB, Lahtinen M, Cooke T, Johnson JA. Comparing the EQ-5D 3L and 5L: measurement properties and association with chronic conditions and multimorbidity in the general population. *Health Qual Life Outcomes.* 2014;12(1):74. <https://doi.org/10.1186/1477-7525-12-74>.
82. Zeng X, Sui M, Liu B, 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>.
83. Brazier J, Ara R, Rowen D, Chevrou-Severac H. A review of generic preference-based measures for use in cost-effectiveness models. *Pharmacoeconomics.* 2017;35(1):21-31. <https://doi.org/10.1007/s40273-017-0545-x>.

## Authors and Affiliations

Micha J. Pilz<sup>1</sup> · Simone Seyringer<sup>2</sup> · Imad Al-Naesan<sup>1</sup> · Madeleine T. King<sup>3</sup> · Andrew Bottomley<sup>4</sup> · Richard Norman<sup>5</sup> · Lisa Schlosser<sup>6</sup> · Tobias Hell<sup>6</sup> · Eva Maria Gamper<sup>2</sup> on behalf of ob behalf of the EORTC Quality of Life Group

✉ Eva Maria Gamper  
eva.gamper@tirol-kliniken.at

<sup>1</sup> University Hospital of Psychiatry II, Medical University of Innsbruck, Innsbruck, Austria

<sup>2</sup> Department of Nuclear Medicine, Medical University of Innsbruck, 6020 Innsbruck, Austria

<sup>3</sup> School of Psychology, University of Sydney, Sydney, NSW, Australia

<sup>4</sup> Quality of Life Department, European Organisation for Research and Treatment of Cancer, Brussels, Belgium

<sup>5</sup> School of Population Health, Curtin University, Perth, WA, Australia

<sup>6</sup> Department of Mathematics, University of Innsbruck, Innsbruck, Austria