

# **Satisfactory results of a psychometric analysis and calculation of minimal clinically important differences of the World Health Organization Quality of Life-BREF questionnaire in an observational cohort study with lung cancer and mesothelioma patients**

M. de Mol,  
S. Visser,  
J.G.J.V. Aerts,  
P. Lodder,  
J. de Vries,  
B.L. den Oudsten

*BMC Cancer 2018;18(1):1173*

## ABSTRACT

**Background:** To determine the psychometric properties and minimal clinically important differences (MCIDs) of the World Health Organization Quality of Life-BREF (WHOQOL-BREF) in advanced-stage lung cancer patients.

**Methods:** Patients (n=153) completed the WHOQOL-BREF and the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 (EORTC QLQ-C30). Confirmatory factor analysis (CFA) was performed and reliability and construct validity determined. MCIDs were estimated with two distribution-based methods (0.5 standard deviation (SD) and 1 standard error of measurement (1 SEM)).

**Results:** CFA confirmed WHOQOL-BREF domain structure. All domains demonstrated good internal consistency ( $\alpha > 0.70$ ), except social relationships ( $\alpha = 0.57$ ). Nineteen of the 24 WHOQOL-BREF items had correlations of  $\geq 0.40$  with their intended domain. Four items had higher correlations with a domain other than their intended domain. Moderate to strong correlations were observed for corresponding domains of the two questionnaires, except for the social domains ( $r = 0.07$ ). For 0.5 SD, MCIDs ranged from 0.88 to 1.55, and for 1 SEM MCIDs ranged from 1.76 to 2.72.

**Conclusions:** The WHOQOL-BREF has satisfactory psychometric properties in patients with advanced-stage lung cancer, whereas the observed MCIDs provide a method for interpretation of scores.

## BACKGROUND

In general, chemotherapy in patients with advanced-stage lung cancer is associated with small survival benefits [1,2]. In addition, Quality of Life (QoL) may be reduced in patients with lung cancer [3]. This emphasizes the importance of maintaining patients' QoL at an acceptable level by early identification of treatment-induced changes. QoL is evaluated by questionnaires of which the European Organisation for Research and Treatment of Cancer Quality of Life-Core 30 questionnaire (EORTC QLQ-C30) is one of the most frequently applied in cancer [4]. However, this instrument is considered to be a Health-Related Quality of Life (HRQoL) questionnaire (i.e., it focusses on those aspects of QoL related to the disease and its treatment) and measures to a lesser extent patients' opinions of the other aspects of QoL [4]. Therefore, the WHO formulated a comprehensive definition of QoL based on extensive research. In 2004, they released the World Health Organization Quality of Life instrument-BREF (WHOQOL-BREF) to enable rapid QoL assessment in epidemiological surveys and clinical studies [5].

Recently, a study performed in Taiwanese patients diagnosed with stage I to IV lung cancer reported satisfactory psychometric properties of the WHOQOL-BREF. However, the 28-item Taiwanese version of the WHOQOL-BREF (the original WHOQOL-BREF holds 26 items [5]) was used and specific results concerning patients with advanced-stage lung cancer were not reported. Therefore, further psychometric validation of the WHOQOL-BREF in this group may be mandatory for several reasons. First, patients with advanced-stage lung cancer form a well-defined group due to their poor prognosis compared to patients with stage I or II lung cancer and the population in the WHOQOL-BREF field trial [5]. Second, apart from the symptoms of lung cancer, treatment is in most patients with advanced-stage lung cancer associated with substantial adverse events which can directly influence (HR)QoL. Third, although some studies have reported results of the WHOQOL-BREF in lung cancer [6,7], the application of this questionnaire in patients starting treatment with chemotherapy was not reported. Fourth, as correct interpretation of the minimal clinically important difference (MCID) depends on the psychometric characteristics of the instrument and the patient population from which it is derived, the determination of a reliable MCID in lung cancer ideally requires evaluation of the reliability and validity of the WHOQOL-BREF in these patients.

Given these considerations, additional research is needed to enable implementation of the WHOQOL-BREF in future trials investigating therapeutic regimens in lung cancer and to facilitate the interpretation of individual scores. To contribute to these goals the objective of our study focused on two main aspects of the WHOQOL-BREF: 1) to test the reliability and validity of the WHOQOL-BREF in patients with advanced-stage lung cancer, and 2) to assess the MCIDs of the WHOQOL-BREF domain scores. We expected that the 4-domain structure of the WHOQOL-BREF would be confirmed and that the internal consistencies

of all domains were at least acceptable, except for social relationships [5]. Moreover, we hypothesized that all items of the WHOQOL-BREF would have an acceptable positive correlation (i.e., correlation coefficient  $\geq 0.40$ ) with their intended domains and that all items would have higher positive correlations with their intended domain than with the other three domains [8]. Considering construct validity, we expected significant differences in mean domain scores between known groups according to ECOG performance score and EORTC QLQ-C30 global Health Status/QoL score [9]. In addition, construct validity was assessed by correlating the domains of the WHOQOL-BREF with the scales of the EORTC QLQ-C30 [9]. We hypothesized that all domains would have at least moderate correlations (i.e., correlation coefficient  $\geq 0.50$ ) with their corresponding scales of the EORTC QLQ-C30 [10]. In this study, we expected no floor or ceiling effects for domain scores of the WHOQOL-BREF.

## METHODS

### Study population

PERSONAL is a prospective observational multi-center cohort study of patients with non-squamous non-small cell lung carcinoma (NSCLC) and unresectable mesothelioma receiving pemetrexed. Patients were recruited from October 2012 to November 2014 from three teaching hospitals (Erasmus University Medical Center, Amphia Hospital, and Sint Franciscus Gasthuis hospital) and a regional hospital (Bravis hospital) in the Netherlands. For this study, which is part of an ongoing analysis of PERSONAL, data of 191 enrolled patients was available. Patients were enrolled if they met the following criteria: were aged eighteen years or older, had a cytological or histological confirmed diagnosis of non-squamous NSCLC or unresectable malignant pleural mesothelioma and started treatment with pemetrexed monotherapy or in combination with a platinum compound. Patients were excluded when they were not able to read Dutch or could not complete the questionnaires because of a physical or mental condition. A sample size of at least 50 patients was needed in order to perform a validation study [9]. Informed consent was obtained from all individual participants included in the study. This multi-center study was approved by the Institutional Review Board of the Erasmus University Medical Center in Rotterdam, the Netherlands.

### Study measures

The WHOQOL-BREF [5,11] is a well-established generic QoL instrument intended for use in a wide range of chronic diseases and cancer [5]. It comprises 24 items divided over four domains plus two items of the general facet describing overall QoL and general health. The domains represent physical health (seven items), psychological health (six items), social

relationships (three items), and environment (eight items) and are scored on a 4-20 scale with higher scores indicating a better QoL [11]. The general facet is scored on a 2-10 scale. Previous studies have demonstrated good psychometric properties of the WHOQOL-BREF in patients with lung cancer [12] and in patients with chronic diseases or different forms of cancer [5].

The EORTC QLQ-C30 is a cancer-specific HRQoL instrument with demonstrated psychometric properties [13]. It consists of 30 items and incorporates a global Health Status/QoL scale, five functional scales and several single items assessing additional symptoms or problems. The functional scales represent physical functioning (five items), cognitive functioning (two items), emotional functioning (four items), role functioning (two items), and social functioning (two items). EORTC QLQ-C30 scales are scored on a 0-100 scale, with higher scores on the functional scales being indicative of better HRQoL, whereas higher scores on the symptom scales are reflective of worse symptoms [4].

All questionnaires were completed after diagnosis and before the first cycle of chemotherapy. In addition to completing the questionnaires, we collected sociodemographic information (i.e., age, gender, educational level, ethnicity, employment, partner status) and clinical information (i.e., cancer stage, type of tumour, line of therapy and the Eastern Cooperative Oncology Group (ECOG) performance status). At day 1 of the first cycle of chemotherapy we assessed, according to Common Terminology Criteria for Adverse Events (CTCAE) version 3.0, the severity and number of different cancer related and, if applicable, treatment related adverse events that patients experienced.

## Statistical analysis

The response distributions of item and domain scores of the WHOQOL-BREF were assessed by using two methods. As proposed in the validation paper of the WHOQOL-BREF, skewness was observed if less than 10% of responses fell in each of two adjacent scale points of an item at the extreme ends of the scale [5]. Floor and ceiling effects of domain scores of the WHOQOL-BREF were considered to be present if more than 15% of the respondents achieved the lowest (i.e., floor effect) or highest (i.e., ceiling effect) possible score [9].

The multi-trait/multi-method methodology, as proposed by Campbell and Fiske [14] and later adapted by Ware et al., was used to study item-domain relations [15]. Analyses were performed with MAP-R software which examines the correlations between items and domains and corrects the correlation of each item with its intended domain for overlap [15]. For the multi-trait/multi-item analyses, missing values are replaced by the mean score of the other items of the corresponding domain if at least half of the items are completed. According to Trask et al., item-convergent validity was defined as a correlation coefficient  $\geq 0.40$  between questionnaire items and their intended domains [8]. Item-divergent validity was supported when items had higher correlations with their intended domain than with other domains of the questionnaire [8].

Construct validity was evaluated by correlating the WHOQOL-BREF domains with the corresponding scales of the EORTC QLQ-C30 using Pearson's correlation coefficient. According to Hinkle, correlations of 0.00 to 0.30 were regarded as negligible, 0.30 to 0.50 as low, 0.50 to 0.70 as moderate, 0.70 to 0.90 as high, and correlations of 0.90 to 1.00 as very high [10]. In addition, known-groups validity comparisons were made for the WHOQOL-BREF domains in relation to the total number of different adverse events, the number of different grade 3 or 4 adverse events, the ECOG performance status and the global Health Status/QoL score of the EORTC QLQ-C30 to assess construct validity. One-way ANOVA was used to determine whether there were any significant differences between the means of the groups.

Internal consistency reflects the capability of items within a domain to measure the same concept. To evaluate internal consistency, first the four-factor design of the WHOQOL-BREF was analysed with confirmatory factor analysis (CFA) using structural equation modelling. Missing values were replaced by expectation-maximization imputation for the CFA. The original model is demonstrated in Figure 1. Goodness of fit was assessed by the Comparative Fit Index (CFI) and the Root Mean Square Error Approximation (RMSEA). A satisfactory to good fit is defined when  $CFI > 0.90$  and  $RMSEA < 0.06$  [16,17]. For the resulting domains, Cronbach's coefficient alpha was calculated to express internal consistency. A coefficient of 0.70 or higher was considered to be acceptable [9].

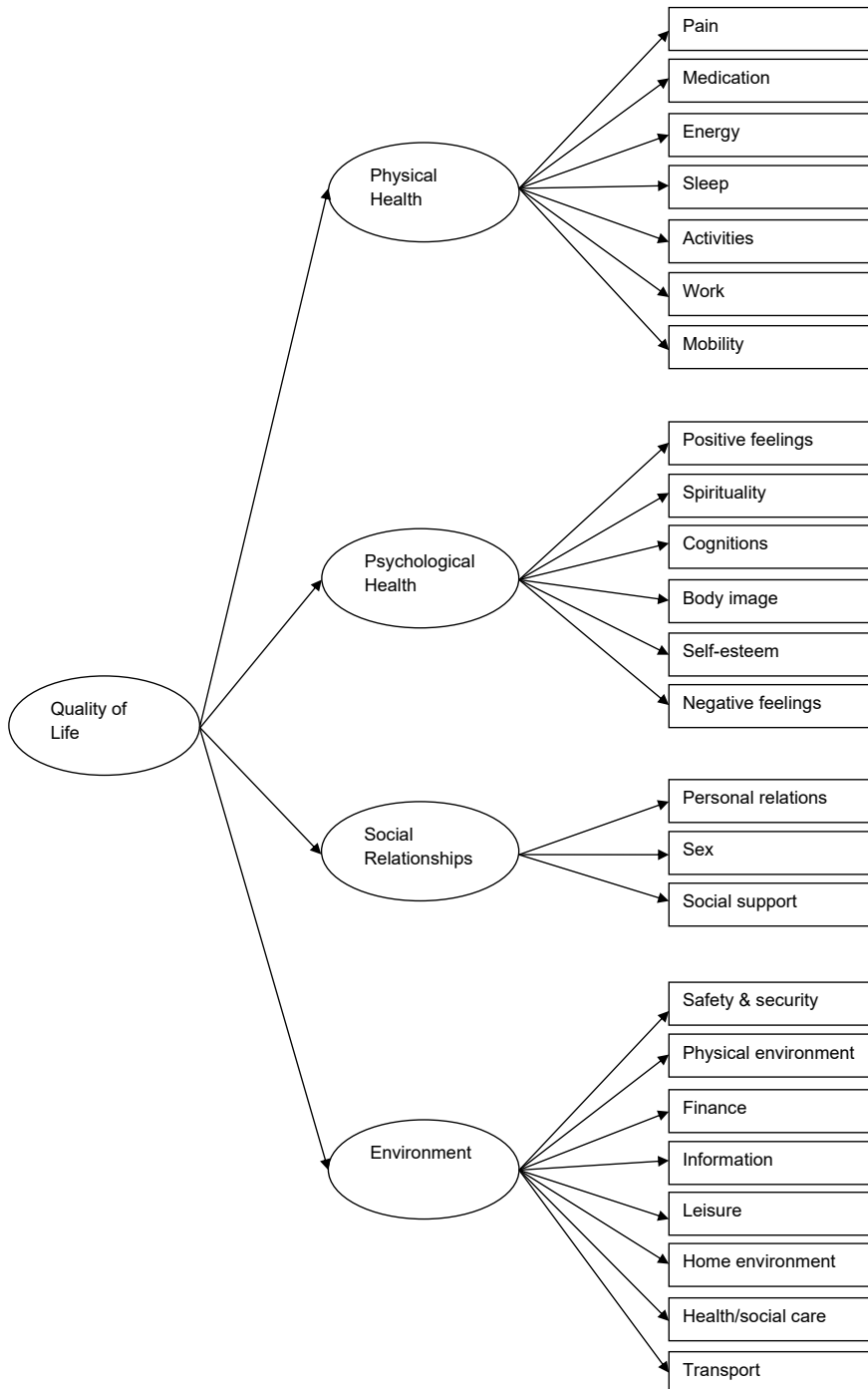
For each WHOQOL-BREF domain, the MCID was calculated using two distribution-based methods (i.e., the 0.5 SD [18] and 1 standard error of measure (SEM) [19-21]). MCID is the smallest change in an outcome that a patient would identify as important. The 0.5 SD benchmark of an outcome measure means that patients improving more than 0.5 of the outcome score's SD have reached a minimal clinically important difference [22]. As we lacked a test-retest reliability coefficient, we used the conservative lower bound of the 95% confidence intervals of the Cronbach's alphas of the four domains to calculate the SEM. Thus, the SEM was calculated with an altered version of the SEM formula [23]:  $SD \times \sqrt{(2 \times (1 - \text{lower bound 95\% Confidence Interval Cronbach's alpha}))}$ .

A p-value of  $p < 0.05$  was considered to be statistically significant. Analyses were performed using SPSS version 21.0, except for the CFA (AMOS version 22.0) and the calculation of the 95% confidence intervals of the four domains of the WHOQOL-BREF (R, version 3.2.5).

## RESULTS

### Patient characteristics

Of the 191 enrolled patients, 153 patients (80.1%) completed the questionnaires to a sufficient degree. Table 1 summarizes the patient characteristics of these patients.

**Fig. 1** Four factor model of the WHOQOL-BREF

Abbreviations: WHOQOL-BREF, World Health Organization Quality of Life-BREF questionnaire.

**Table 1. Characteristics of study population**

Characteristic	Overall sample (N=153)
Age, years	
Mean (SD)	63.4 (9.2)
Min, max	37, 83
Sex	
Male	83 (54.2)
Ethnicity	
White/Caucasian	143 (93.5)
Asian	3 (2.0)
Negroid	2 (1.3)
Other	5 (3.3)
Education <sup>a</sup>	
Low	113 (73.9)
High	33 (21.6)
Unknown	7 (4.6)
Employment	
Yes	40 (26.1)
No	112 (73.2)
Unknown	1 (0.7)
Partner status	
Married/cohabiting	116 (75.8)
Unmarried partners/not cohabiting	6 (3.9)
Divorced/separated	14 (9.2)
Widowed/partner died	10 (6.5)
Single	6 (3.9)
Unknown	1 (0.7)
Cancer stage <sup>b</sup>	
Locally advanced (IIIB)	19 (12.4)
Metastatic (IV)	119 (77.8)
Other	14 (9.2)
Unknown	1 (0.7)
Type of tumor <sup>b</sup>	
Adenocarcinoma	141 (92.2)
Large cell carcinoma	4 (2.6)
Mesothelioma	7 (4.6)
Large cell neuroendocrine carcinoma	1 (0.7)
Line of therapy	
First line	134 (87.6)
Second line	10 (6.5)
Third line	1 (0.7)



**Table 1. Characteristics of study population (continued)**

Characteristic	Overall sample (N=153)
Adjuvant	8 (5.2)
ECOG performance status <sup>a</sup>	
Grade 0	39 (25.5)
Grade 1	99 (64.7)
Grade 2	11 (7.2)
Grade 3	2 (1.3)
Unknown	2 (1.3)

Values are given in numbers (percentages) unless stated otherwise.

<sup>a</sup>Low education: persons whose highest level of education is primary education, lower general education or lower vocational education. High education: persons whose highest level of education is higher general education, higher vocational education or university.

<sup>b</sup>Measured at baseline.

Abbreviations: N, number of patients; SD, standard deviation; ECOG, Eastern Cooperative Oncology Group.

### Mean scores, floor and ceiling effects, and skewness

The WHOQOL-BREF domain scores are shown in Table 2. The mean general facet score was 5.9 (1.8). Mean scores of the four domains ranged from 12.9 (SD 3.1; physical health) to 16.2 (SD 2.6; social relationships). Floor and ceiling effects of the domain scores were below the limit of 15%. Fourteen of the 26 items demonstrated skewed response distributions with responses <10% in each of two adjacent scale points at the extreme lower end of the scale indicating that most of the information was distributed over the other scale points (Table 2). These items were positive feelings, spirituality, cognitions, self-esteem, and body image for the psychological health domain and personal relations and social support for the social relationships domain. In addition, all items of the environment domain, except leisure, demonstrated few responses at the extreme lower end of the scale. One item, negative feelings exhibited responses <10% in each of two adjacent scale points at the extreme upper end of its scale.

### Confirmatory Factor Analysis

CFA with the use of structural equation modelling was conducted to analyse the four-factor structure of the WHOQOL-BREF. Inspection of the modification indices revealed two possible modifications to improve the model fit of the original model. After adding error covariances between the measurement error of the items 1 (pain) and 2 (medication) and between 8 (positive feelings) and 9 (spirituality) model fit improved. The CFI increased from 0.854 to 0.896 whereas the RMSEA decreased from 0.069 to 0.058 approaching both of the criteria for a satisfactory to good fit (CFI > 0.90 and RMSEA < 0.06).

**Table 2. Frequency responses for items of the WHOQOL-BREF**

Items/ domains	Description	N	Mean (SD)	Floor effect (%)	Ceiling effect (%)	Scale points <sup>a</sup>				
						1	2	3	4	5
General facet										
	overall QoL	150	5.9 (1.8)	4 (2.7)	3 (2.0)	3.3	14.4	24.2	44.4	12.4
	general health					21.6	35.9	26.1	13.1	2.6
Physical health										
3	pain	153	12.9 (3.1)	1 (0.7)	1 (0.7)	28.1	28.8	19.0	20.3	3.3
4	medication					11.8	26.8	25.5	24.8	10.5
10	energy					3.9	19.6	35.9	25.5	14.4
15	sleep					8.5	23.5	26.1	26.1	15.0
16	activities					7.8	34.6	24.8	26.8	5.9
17	work					12.4	39.2	21.6	20.3	6.5
25	mobility					4.6	11.8	13.1	35.3	35.3
Psychological health										
5	positive feelings	153	14.4 (2.5)	0 (0.0)	2 (1.3)	<b>3.9</b>	<b>7.8</b>	39.2	41.2	7.8
6	spirituality					<b>2.6</b>	<b>5.2</b>	32.7	43.8	15.0
7	cognitions					<b>4.6</b>	<b>9.8</b>	44.4	22.2	19.0
11	body image					<b>0.7</b>	<b>4.6</b>	28.1	26.1	39.9
18	self esteem					<b>0.7</b>	<b>9.8</b>	30.7	43.8	13.7
26	negative feelings					14.4	32.7	43.1	<b>9.8</b>	<b>0.0</b>
Social relationships										
19	personal relations	153	16.2 (2.6)	0 (0.0)	20 (13.1)	<b>0.7</b>	<b>0.7</b>	<b>6.5</b>	40.5	50.3
20	sex					5.9	13.7	36.6	24.8	16.3
21	social support					<b>0.7</b>	<b>1.3</b>	<b>9.2</b>	34.0	54.9
Environment										
8	safety & security	153	15.9 (2.3)	0 (0.0)	1 (0.7)	<b>0.7</b>	<b>3.9</b>	26.8	32.0	36.6
9	physical environment					<b>0.7</b>	<b>2.0</b>	26.8	35.3	34.0
12	finance					<b>2.0</b>	<b>6.5</b>	37.3	21.6	32.0
13	information					<b>0.0</b>	<b>0.7</b>	36.6	37.3	24.8
14	leisure					3.9	13.1	24.2	35.3	23.5
22	home environment					<b>2.6</b>	<b>5.2</b>	12.4	39.9	39.9
23	health/social care					<b>1.3</b>	<b>2.0</b>	15.7	49.0	32.0
24	transport					<b>1.3</b>	<b>2.0</b>	<b>9.2</b>	43.8	43.8

<sup>a</sup>Values are given in percentages.

Values in bold represent skewed distributions of the frequency of responses of patients.

Abbreviations: WHOQOL-BREF, World Health Organization Quality of Life-BREF questionnaire; N, number of patients; SD, standard deviation; QoL, Quality of Life.

## Multi-trait/multi-item analyses and internal consistency

Multi-trait/multi-item analyses demonstrated that all items, except those representing sleep, body image, sex, physical environment, and finance had a correlation of  $\geq 0.40$  with their intended domain (Table 3). Four items showed higher correlations with other domains than their own. The item sleep of the physical health domain had a higher positive correlation with the psychological health domain, whereas the items body image and self-esteem of the psychological health domain had higher positive correlations with the environment domain and the physical health domain respectively. In addition, the item personal relationships of the social relationships domain showed a higher positive correlation with the environment domain than its intended domain. For all domains, except the social relationships domain, Cronbach's alpha was higher than 0.70 (i.e., physical health: 0.81, 95% CI 0.76-0.85; psychological health: 0.77, 95% CI 0.71-0.83; environment: 0.77, 95% CI 0.70-0.82; social relationships: 0.57, 95% CI 0.43-0.68).

## Construct validity

Table 4 presents the correlations between the domains of the WHOQOL-BREF and the EORTC QLQ-C30 domains/symptom scales. In general, low correlations were observed between WHOQOL-BREF domains and EORTC QLQ-C30 domains/symptom scales. Only for physical health, moderate to high correlations were observed with the EORTC QLQ-C30 domains except for the correlation with cognitive functioning. The lowest correlations were found between social relationships and the EORTC QLQ-C30 domains/symptom scales. The observed negative correlations between the WHOQOL-BREF and EORTC QLQ-C30 symptom scales indicate that a higher score of the WHOQOL-BREF domains corresponded with less worse symptoms.

Table 5 shows the known-groups validity comparisons for the WHOQOL-BREF domains and general facet in relation to the number of different adverse events, the number of different grade 3 or 4 adverse events, the ECOG performance status and the global Health Status/QoL score. Significant differences were detected regarding the general facet score, physical health, and psychological health among ECOG grades 0, 1, and 2 or higher. Similar results were observed for the general facet and the WHOQOL-BREF domain scores according to global Health Status/QoL as measured by the EORTC QLQ-C30 except for social relationships. For all of the observed significant differences except one (i.e., psychological health based on ECOG performance score), effect sizes were medium to large.

## Minimal clinically important differences

Table 6 demonstrates the distribution-based estimates of the MCIDs for the different domains of the WHOQOL-BREF.

**Table 3. Multi-trait/Multi-item item-domain correlation for the WHOQOL-BREF (N=153)**

Items/domains	Description	Physical health	Psychological health	Social relationships	Environment
Physical health					
3R	pain	.53*	.41	<b>.20</b>	<b>.35</b>
4R	medication	.40*	<b>.24</b>	<b>.01</b>	<b>.21</b>
10	energy	.57*	.53	<b>.13</b>	.42
15	sleep	.27*	.32	<b>.06</b>	.27
16	activities	.72*	.56	<b>.17</b>	<b>.46</b>
17	work	.70*	<b>.54</b>	<b>.17</b>	<b>.43</b>
25	mobility	.61*	.47	<b>.24</b>	<b>.45</b>
Psychological health					
5	positive feelings	.52	.67*	<b>.33</b>	<b>.47</b>
6	spirituality	<b>.34</b>	.53*	<b>.24</b>	<b>.36</b>
7	cognitions	.44	.48*	<b>.20</b>	.48
11	body image	.32	.34*	.30	.45
18	self esteem	.60	.58*	<b>.38</b>	.49
26R	negative feelings	.43	.52*	<b>.21</b>	.43
Social relationships					
19	personal relations	<b>.16</b>	.30	.43*	.47
20	sex	<b>.15</b>	.31	.32*	.31
21	social support	<b>.16</b>	.28	.43*	.42
Environment					
8	safety & security	.46	.57	<b>.36</b>	.60*
9	physical environment	<b>.12</b>	.35	.31	.38*
12	finance	.23	.29	.26	.30*
13	information	<b>.24</b>	<b>.33</b>	<b>.36</b>	.53*
14	leisure	.41	.34	<b>.19</b>	.46*
22	home environment	.46	.55	.41	.56*
23	health/social care	.36	.41	.32	.41*
24	transport	<b>.39</b>	.44	.43	.57*

\*Pearson item-scale correlations corrected for overlap.

Correlations in bold represent correlations between items and domains that differ more than two standard errors from their correlations with their own domains.

Abbreviations: SD, standard deviation; N, number of patients; QoL, Quality of Life; WHOQOL-BREF, World Health Organization Quality of Life-BREF questionnaire.

## DISCUSSION

Patients with advanced-stage lung cancer are prone to a decrease in QoL due to poor prognosis and cancer and treatment-related adverse events. Unfortunately, trials investigating

**Table 4. Correlations of the WHOQOL-BREF with the EORTC QLQ-C30 domains (N=153)**

EORTC QLQ-C30 domains/ Items	WHOQOL-BREF domains/items				
	General facet	Physical health	Psychological health	Social relationships	Environment
Physical functioning	0.43*	0.73*	0.37*	0.04	0.26*
Role functioning	0.46*	0.73*	0.46*	0.08	0.26*
Emotional functioning	0.49*	0.51*	0.61*	0.19**	0.43*
Cognitive functioning	0.33*	0.49*	0.47*	0.08	0.40*
Social functioning	0.42*	0.59*	0.47*	0.07	0.32*
Global Health Status/QoL	0.67*	0.73*	0.58*	0.12	0.39*
Fatigue	-0.39*	-0.69*	-0.44*	-0.06	-0.33*
Nausea and vomiting	-0.32*	-0.29*	-0.24*	-0.16	-0.05
Pain	-0.32*	-0.62*	-0.28*	-0.06	-0.26*
Dyspnea	-0.28*	-0.30*	-0.15	0.05	-0.10
Insomnia	-0.29*	-0.49*	-0.35*	-0.04	-0.32*
Appetite loss	-0.36*	-0.38*	-0.22*	0.02	-0.05
Constipation	-0.17**	-0.23*	-0.18**	-0.06	-0.24*
Diarrhea	0.00	-0.08	-0.04	-0.10	-0.06
Financial problems	-0.03	-0.34*	-0.22*	-0.11	-0.36*

Pearson correlation coefficients.

\* $P < 0.01$ .

\*\* $P < 0.05$ .

Abbreviations: WHOQOL-BREF, World Health Organization Quality of Life-BREF questionnaire; EORTC QLQ-C30, European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; N, number of patients; QoL, Quality of Life.

new therapies and treatment modalities in lung cancer often assess the impact on QoL with the use of HRQoL instruments [2,24-26]. This is unfortunate as the WHOQOL-BREF may facilitate a more comprehensive evaluation of QoL. Given the importance of a comprehensive evaluation of QoL, the present study assessed the psychometrics and MCIDs of the WHOQOL-BREF in patients with advanced-stage lung cancer to facilitate adequate QoL monitoring in clinical practice and lung cancer trials. In general, our study demonstrated that the WHOQOL-BREF is a reliable and valid instrument in patients with advanced-stage lung cancer.

We found that the general health item of the general facet was more positively skewed in our study compared with the WHOQOL-BREF field trial reflecting higher frequencies of patients with worse general health [5]. This is as expected given the frequent occurrence of adverse events and poor prognosis of advanced-stage lung cancer. However, the patients in this study indicated better QoL for several items of the psychological health, social relationships and environment domains than the patients included in the field trial. Moreover, an additional seven items of these three domains were negatively skewed in our patients

**Table 5. Known-groups comparisons for the WHOQOL-BREF general facet and domain scores (n=153)**

	General facet			Physical health			Psychological health			Social relationships			Environment		
	N	Mean (SD)	P-value (effect size)*	N	Mean (SD)	P-value (effect size)*	N	Mean (SD)	P-value (effect size)*	N	Mean (SD)	P-value (effect size)*	N	Mean (SD)	P-value (effect size)*
Number of different adverse events <sup>a</sup>															
0-5	84	6.1 (1.7)	0.10	87	13.6 (3.1)	<0.001 (0.08)	87	14.5 (2.5)	0.47	87	16.1 (2.7)	0.37	87	15.8 (2.2)	0.81
more than 5	66	5.6 (1.8)		66	11.8 (2.9)		66	14.2 (2.6)		66	16.4 (2.4)		66	15.9 (2.4)	
Number of different adverse events with CTCAE grade 3 or 4 <sup>a</sup>															
0	112	6.0 (1.8)	0.16	115	13.1 (3.1)	0.06	115	14.4 (2.4)	0.93	115	16.1 (2.6)	0.36	115	15.8 (2.3)	0.82
1	27	6.0 (1.8)		27	12.8 (3.0)		27	14.4 (2.8)		27	16.5 (2.6)		27	16.0 (2.3)	
2 or more	11	4.9 (1.4)		11	10.8 (2.6)		11	14.7 (2.9)		11	17.1 (2.2)		11	16.1 (2.6)	
ECOG performance score <sup>b</sup>			<0.001 (0.08)			<0.001 (0.23)			0.04 (0.04)			0.05			0.38
0	37	6.6 (1.5)		39	15.1 (2.3)		39	15.1 (2.2)		39	16.3 (2.5)		39	16.1 (1.8)	
1	98	5.7 (1.8)		99	12.3 (3.0)		99	14.3 (2.6)		99	16.5 (2.3)		99	15.9 (2.5)	
2 or more	13	4.8 (1.1)		13	10.2 (2.0)		13	13.1 (2.2)		13	14.7 (3.5)		13	15.1 (2.3)	
Global Health Status/QoL EORTC QLQ-C30 score															
0-50	67	4.9 (1.5)	<0.001 (0.30)	69	10.8 (2.7)	<0.001 (0.36)	69	13.1 (2.6)	<0.001 (0.22)	69	16.0 (2.8)	0.28	69	15.1 (2.4)	<0.001 (0.10)
more than 50	79	6.8 (1.4)		80	14.6 (2.3)		80	15.5 (1.9)		80	16.5 (2.3)		80	16.6 (2.0)	

P-values calculated with one-way ANOVA unless stated otherwise.

\*Effect sizes ( $\eta^2$ ) were only shown where one-way ANOVA was significant ( $P < 0.05$ ).

<sup>a</sup>Reported adverse events: as reported at and before day 1 of the first cycle of chemotherapy.

<sup>b</sup>Post hoc analyses with Tukey HSD test of significant differences: for the general facet, between ECOG 0 and 1 and also 0 and 2 or higher; for physical health, between all ECOG categories; for psychological health, between ECOG 0 and ECOG 2 or higher.

Abbreviations: WHOQOL-BREF, World Health Organization Quality of Life-BREF questionnaire; SD, standard deviation; N, number of patients who completed the questionnaire; CTCAE, Common Terminology Criteria for Adverse Events; ECOG, Eastern Cooperative Oncology Group; QoL, Quality of Life; EORTC QLQ-C30, European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30.

**Table 6. Estimates of minimal clinically important differences on WHOQOL-BREF domains**

Domains	0.5 SD	1 SEM
General facet	0.876	
Physical health	1.545	2.155
Psychological health	1.259	1.914
Social relationships	1.274	2.716
Environment	1.142	1.761

Abbreviations: WHOQOL-BREF, World Health Organization Quality of Life-BREF questionnaire; SD, standard deviation; SEM, standard error of measure.

indicating also better QoL. One item (i.e., negative feelings) was positively skewed demonstrating that most patients rarely experienced negative feelings while the WHOQOL-BREF field trial observed higher frequencies in the scale points that corresponded with increased negative feelings. As this higher level of QoL was not related to physical QoL, which is in general determined by universal factors (i.e., the cancer and its treatment), but rather to the other domains of the WHOQOL-BREF, this may be explained by several reasons. Given the negative skewness of seven of the eight items of the environment domain, it is likely that the high standard of care and the high level of prosperity in the Netherlands may be, at least in part, responsible for this observation. In addition, patients with lung cancer may experience less psychological distress compared to patients with other types of cancer. A meta-analysis by Krebber et al. found that the prevalence of depression as diagnosed by a structural interview was the lowest (3%) in lung cancer patients compared with other forms of cancer. The prevalence of depression as diagnosed by self-report instruments (20%) was also lower or comparable with other forms of cancer [27].

Prior to testing the reliability and validity of the WHOQOL-BREF, we performed a first order CFA to analyse if the proposed four-factor model matched with the patients in the present study. Before (i.e., RMSEA) and after (i.e., CFI and RMSEA) adding error covariances between the measurement errors of items pain and medication and between items positive feelings and spirituality, the observed fit indices indicated a slightly better model fit than the field trial of the WHOQOL-BREF [5]. However, as we were not able to calculate 95% confidence intervals for the observed fit indices and Skevington et al. did not report them [5], we could not determine if the CFI and RMSEA observed in the present study were significantly different. Moreover, if they are different, it is likely that the differences in fit indices are explained by the differences between patient populations of both studies. In the present study a homogeneous sample of patients with advanced-stage lung cancer was used whereas the patient population of the WHOQOL-BREF field trial consisted of patients with different diseases [5]. Also the statistical differences between the present study and that of Skevington et al. impair the direct comparison of model fit.

Similarly as observed by Skevington et al., the internal consistency of the social relationships domain was below the commonly accepted value of 0.7 [5] whereas the other domains all had a Cronbach's alpha > 0.70. As Cronbach's alpha is in part dependent of the number of items of a domain, a reason for this low alpha possibly lies in the fact that the social relationships domain consists of just three items. In a recent Taiwanese validation study of the WHOQOL-BREF which did not report specific results of patients with advanced-stage lung cancer (i.e., overall results of Rasch analyses of patients with stage I to IV disease were reported), the inclusion of one extra item (i.e., being respected) in the social relationship domain resulted in a Cronbach's alpha of 0.67 [12], which is higher than observed in this study, although comparable with the alpha found in the field trial (0.68) [5]. Explanations for the lower observed internal consistency of the social relationships domain in our study in contrast with the other two reports could be the homogeneity of the patient sample or the decreased ability of the combined items to reflect the underlying construct in patients with advanced-stage lung cancer compared to those with limited disease stage or other forms of cancer or chronic diseases. Furthermore, one of the three items (i.e., personal relations) had a higher correlation with the environment domain than with its own hypothesized domain in this study which indicates that this item may not be completely representative for the construct of social relationships. In addition to the relatively low Cronbach's alpha, this result further hampers the interpretation of analyses with this domain and raises the question if the three items should be assessed separately.

After performing multi-trait/multi-item analyses we observed similar cross domain correlations as the field trial did. While the self-esteem item of the psychological health domain in the field trial was strongly related with the other three domains [5], we observed a stronger correlation with the physical health domain than with its own domain. This is not only in contrast with the results of the field trial [5], but also with patients with other forms of cancers. One study in cervical cancer survivors reported that self-esteem was related to the mental component summary score and not with the physical component summary score of the Short Form 36 QoL questionnaire [28]. A reason for this result could be the considerable impact advanced-stage lung cancer can have on physical abilities. This may lead to dependence of others which could affect self-esteem. In the field trial of the WHOQOL-BREF the centre specific analyses revealed that the items safety & security and energy often had higher correlations with domains other than their own [5] whereas we found that this was the case for the items sleep, body image, self-esteem, and personal relations. These differences in cross-correlation could be explained by some reasons. For instance, as the sample size of this study was relatively small, the observed differences may reflect mere chance than a true observation. Also methodological differences and the specific characteristics of patients with advanced disease (e.g., poor prognosis, prone to cancer-related adverse events) are, at least in part, responsible for these findings.



In general, low correlations were observed between WHOQOL-BREF domains and EORTC QLQ-C30 domains/symptom scales. This is probably related to differences in constructs and concepts between the questionnaires. Whereas the WHOQOL-BREF is a generic questionnaire, the EORTC QLQ-C30 is a cancer specific questionnaire. Moreover, items of the WHOQOL-BREF are positively phrased while those of the EORTC QLQ-C30 are often negatively phrased. In this regard, the EORTC QLQ-C30 may not be regarded as a gold standard to evaluate construct validity of the WHOQOL-BREF. In addition, this also points to the additional value of the WHOQOL-BREF in QoL analyses in cancer patients.

Both the field trial of the WHOQOL-BREF and the recent Taiwanese study did not report MCIDs to facilitate the clinical application of the WHOQOL-BREF [5,12]. In the present study, we were able to report statistically derived MCIDs for the four WHOQOL-BREF domains. Because we were not able to perform a test-retest reliability analysis, we used the conservative lower bound of the 95% confidence intervals of each of the Cronbach's alphas of the WHOQOL-BREF domains for the calculation of the 1 SEM MCIDs. Considering the vulnerability of patients with advanced-stage lung cancer for treatment and cancer-related adverse events and the short period of three weeks between chemotherapy cycles, we expected it to be difficult to define an appropriated interval between completions of the WHOQOL-BREF for two reasons. 1) If the interval between completions of the WHOQOL-BREF would be too short, patients could recall their earlier answers. 2) If the interval between completions of the WHOQOL-BREF would be too long, it is likely that the occurrence of therapy and cancer-related adverse events would have biased WHOQOL-BREF scores. However, by taking the lower bound of the confidence interval, we expected that patients who experience a larger difference over time than the observed SEM estimates are likely to have a true change. Considering that the 0.5 SD MCIDs depend on the variance of test scores, which are expected to be relatively small in a homogenous patient population as in the present study, the larger 1 SEM MCIDs may provide a more conservative method for the interpretation of individual scores. However, 1 SEM MCIDs depend on the reliability of a questionnaire. A questionnaire with a limited reliability may result in a relatively large 1 SEM. This could result in an overestimation of the true MCID which decreases sensitivity. Given these considerations, we calculated MCIDs according to both methods and recommended to base the choice for either of the two approaches on the homogeneity of the patient sample and the reliability of the questionnaire in the particular population.

Another limitation is that the present study used CFA in combination with the multi-trait/multi-method methodology [14,15] which is in contrast with the increased application of Rasch analysis in recent years to assess psychometric properties of QoL questionnaires in cancer [29-32]. However, we chose the same methodology for the analyses to enable precise comparisons of the psychometric properties observed in this study with those reported by the original field trial of the WHOQOL-BREF.

Lastly, the sample size of our study could be considered a limitation. Although we included less than recommended 200 patients by Boomsma and Hoogland [33], we still observed an acceptable model fit which demonstrated that our data suited the simple design of the model [34].

## CONCLUSIONS

This study demonstrated that the WHOQOL-BREF has satisfactory reliability and validity in patients diagnosed with advanced-stage lung cancer. Moreover, we identified and proposed MCIDs to facilitate application of the WHOQOL-BREF not only in studies investigating new therapies and treatment modalities, but also in daily clinical practice.

## ACKNOWLEDGEMENTS

We acknowledge the assistance of John E. Ware, Jr., PhD and Barbara Gandek, PhD for providing the MAP-R software which facilitated the internal consistency and construct validity analyses of the WHOQOL-BREF in this study.

## REFERENCES

1. Group NM-AC (2008) Chemotherapy in addition to supportive care improves survival in advanced non-small-cell lung cancer: a systematic review and meta-analysis of individual patient data from 16 randomized controlled trials. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 26 (28):4617-4625. doi:10.1200/JCO.2008.17.7162
2. Spiro SG, Rudd RM, Souhami RL, Brown J, Fairlamb DJ, Gower NH, Maslove L, Milroy R, Napp V, Parmar MK, Peake MD, Stephens RJ, Thorpe H, Waller DA, West P, Big Lung Trial p (2004) Chemotherapy versus supportive care in advanced non-small cell lung cancer: improved survival without detriment to quality of life. *Thorax* 59 (10):828-836. doi:10.1136/thx.2003.020164
3. Chabowski M, Polanski J, Mazur G, Janczak D, Rosinczuk J (2017) Sociodemographic and Clinical Determinants of Quality of Life of Patients with Non-small Cell Lung Cancer. *Adv Exp Med Biol* 1022:1-10. doi:10.1007/5584\_2017\_36
4. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, Filiberti A, Flechtner H, Fleishman SB, de Haes JC, et al. (1993) The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *Journal of the National Cancer Institute* 85 (5):365-376
5. Skevington SM, Lotfy M, O'Connell KA, Group W (2004) The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 13 (2):299-310
6. Mohan A, Mohan C, Bhutani M, Pathak AK, Pal H, Das C, Guleria R (2006) Quality of life in newly diagnosed patients with lung cancer in a developing country: is it important? *European journal of cancer care* 15 (3):293-298. doi:10.1111/j.1365-2354.2006.00654.x
7. Mohan A, Mohan C, Pathak AK, Pandey RM, Guleria R (2007) Impact of chronic obstructive pulmonary disease on respiratory status and quality of life in newly diagnosed patients with lung cancer. *Respirology* 12 (2):240-247. doi:10.1111/j.1440-1843.2006.01038.x
8. Trask PC, Tellefsen C, Espindle D, Getter C, Hsu MA (2008) Psychometric validation of the cancer therapy satisfaction questionnaire. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* 11 (4):669-679. doi:10.1111/j.1524-4733.2007.00310.x
9. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC (2007) Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of clinical epidemiology* 60 (1):34-42. doi:10.1016/j.jclinepi.2006.03.012
10. Hinkle DE WW, Jurs SG (2003) *Applied statistics for the behavioral sciences*. 5th ed Boston, Mass: Houghton Mifflin;
11. Development of the World Health Organization WHOQOL-BREF quality of life assessment. The WHOQOL Group (1998). *Psychological medicine* 28 (3):551-558
12. Lin CY, Yang SC, Lai WW, Su WC, Wang JD (2015) Rasch models suggested the satisfactory psychometric properties of the World Health Organization Quality of Life-Brief among lung cancer patients. *Journal of health psychology*. doi:10.1177/1359105315603474
13. Luckett T, King MT, Butow PN, Oguchi M, Rankin N, Price MA, Hackl NA, Heading G (2011) Choosing between the EORTC QLQ-C30 and FACT-G for measuring health-related quality of life in cancer clinical research: issues, evidence and recommendations. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 22 (10):2179-2190. doi:10.1093/annonc/mdq721
14. Campbell DT, Fiske, D. W. (1959) Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological bulletin* 56 (2):81-105

15. Ware JE, Harris, W.J., Gandek, B., et al. (1997) MAP-R for Windows: Multitrait./Multi-Item Analysis Program-Revised User's Guide. Health Assessment Lab; Boston, MA:
16. Den Oudsten BL, Van Heck GL, Van der Steeg AF, Roukema JA, De Vries J (2009) The WHOQOL-100 has good psychometric properties in breast cancer patients. *Journal of clinical epidemiology* 62 (2):195-205. doi:10.1016/j.jclinepi.2008.03.006
17. Hu L BP (1999) Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model* 6:1-55
18. Norman GR, Sloan JA, Wyrwich KW (2003) Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 41 (5):582-592. doi:10.1097/01.MLR.0000062554.74615.4C
19. Wyrwich KW, Nienaber NA, Tierney WM, Wolinsky FD (1999) Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. *Med Care* 37 (5):469-478
20. Wyrwich KW, Tierney WM, Wolinsky FD (1999) Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol* 52 (9):861-873. doi:S0895435699000712 [pii]
21. Wyrwich KW, Tierney WM, Wolinsky FD (2002) Using the standard error of measurement to identify important changes on the Asthma Quality of Life Questionnaire. *Qual Life Res* 11 (1):1-7
22. King MT (2011) A point of minimal important difference (MID): a critique of terminology and methods. *Expert Rev Pharmacoecon Outcomes Res* 11 (2):171-184. doi:10.1586/erp.11.9
23. Beaton DE, Bombardier C, Katz JN, Wright JG (2001) A taxonomy for responsiveness. *Journal of clinical epidemiology* 54 (12):1204-1217
24. Rittmeyer A, Gorbunova V, Vikstrom A, Scherpereel A, Kim JH, Ahn MJ, Chella A, Chouaid C, Campbell AK, Barlesi F (2013) Health-related quality of life in patients with advanced nonsquamous non-small-cell lung cancer receiving bevacizumab or bevacizumab-plus-pemetrexed maintenance therapy in AVAPERL (MO22089). *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer* 8 (11):1409-1416. doi:10.1097/JTO.0b013e3182a46bfc
25. Gridelli C, de Marinis F, Pujol JL, Reck M, Ramlau R, Parente B, Pieters T, Middleton G, Corral J, Winfree K, Melemed S, Zimmermann A, John W, Beyrer J, Chouaki N, Visseren-Grul C, Paz-Ares LG (2012) Safety, resource use, and quality of life in paramount: a phase III study of maintenance pemetrexed versus placebo after induction pemetrexed plus cisplatin for advanced nonsquamous non-small-cell lung cancer. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer* 7 (11):1713-1721. doi:10.1097/JTO.0b013e318267cf84
26. Spigel DR, Patel JD, Reynolds CH, Garon EB, Hermann RC, Govindan R, Olsen MR, Winfree KB, Chen J, Liu J, Guba SC, Socinski MA, Bonomi P (2015) Quality of life analyses from the randomized, open-label, phase III PointBreak study of pemetrexed-carboplatin-bevacizumab followed by maintenance pemetrexed-bevacizumab versus paclitaxel-carboplatin-bevacizumab followed by maintenance bevacizumab in patients with stage IIIB or IV nonsquamous non-small-cell lung cancer. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer* 10 (2):353-359. doi:10.1097/JTO.0000000000000277
27. Krebber AM, Buffart LM, Kleijn G, Riepma IC, de Bree R, Leemans CR, Becker A, Brug J, van Straten A, Cuijpers P, Verdonck-de Leeuw IM (2014) Prevalence of depression in cancer patients: a meta-analysis of diagnostic interviews and self-report instruments. *Psycho-oncology* 23 (2):121-130. doi:10.1002/pon.3409

28. Bartoces MG, Severson RK, Rusin BA, Schwartz KL, Ruterbusch JJ, Neale AV (2009) Quality of life and self-esteem of long-term survivors of invasive and noninvasive cervical cancer. *Journal of women's health* 18 (5):655-661. doi:10.1089/jwh.2008.0959
29. Ashley L, Smith AB, Jones H, Velikova G, Wright P (2014) Traditional and Rasch psychometric analyses of the Quality of Life in Adult Cancer Survivors (QLACS) questionnaire in shorter-term cancer survivors 15 months post-diagnosis. *Journal of psychosomatic research* 77 (4):322-329. doi:10.1016/j.jpsychores.2014.07.007
30. Shih CL, Chen CH, Sheu CF, Lang HC, Hsieh CL (2013) Validating and improving the reliability of the EORTC qlq-c30 using a multidimensional Rasch model. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* 16 (5):848-854. doi:10.1016/j.jval.2013.05.004
31. Winstanley JB, Saw R, Boyle F, Thompson J (2013) The FACT-Melanoma quality-of-life instrument: comparison of a five-point and four-point response scale using the Rasch measurement model. *Melanoma research* 23 (1):61-69. doi:10.1097/CMR.0b013e32835c7dd9
32. Holzner B, Bode RK, Hahn EA, Cella D, Kopp M, Sperner-Unterweger B, Kemmler G (2006) Equating EORTC QLQ-C30 and FACT-G scores and its use in oncological research. *European journal of cancer* 42 (18):3169-3177. doi:10.1016/j.ejca.2006.08.016
33. Boomsma A, Hoogland, J.J. (2001) The robustness of LISREL modeling revisited. In Cudeck, R., Du Toit, S., Sörbom D. (Eds.), *Structural equation modelling: Present and future* (pp. 139-168). Chicago: Scientific Software.
34. Byrne BM (2005) Factor analytic models: viewing the structure of an assessment instrument from three perspectives. *Journal of personality assessment* 85 (1):17-32. doi:10.1207/s15327752jpa8501\_02