Original Article

Quality of life of head and neck cancer patient: validation of the European organization for research and treatment of cancer QLQ-C30 and European organization for research and treatment of cancer QLQ-H&N35 in Indian patients

Chaukar DA, Das AK, Deshpande MS, Pai PS, Pathak KA, Chaturvedi P, Kakade AC, Hawaldar RW, D' Cruz AK
TATA Memorial Hospital, Parel, Mumbai, India.

Correspondence to: D. A. Chaukar, E-mail: dchaukar@rediffmail.com

Abstract

AIMS: To present the first cross-culture validation of the European organization for research and treatment of cancer (EORTC) quality of life questionnaires, the EORTC-QLQ-C30, and the QLQ-H&N35 in India.
SETTINGS AND DESIGN: These questionnaires were translated into two vernacular languages and pilot test was done on 15 patients. Two hundred head and neck cancer patients completed the QLQ-C30 and the QLQ-H&N35 at two time points during their treatment. Psychometric evaluation of the structure, reliability, and validity of the questionnaire was undertaken. **RESULTS:** The data support the reliability of the scales. Validity was tested by item-scale, scale-scale correlation, and by performing known group comparisons. The results demonstrated that the items correlated with their respective scale and no significant correlation was found between scales. The questionnaire was responsive to change over a period of time. **SUMMARY:** These data suggest that the EORTC QLO-C30 and the QLQ-H&N35 are reliable and valid questionnaires when applied to a sample of head and neck cancer patients in India.

Key Words: Head and neck cancer, quality of life, validation

Cancer outcomes are traditionally measured in terms of overall survival, disease free survival, time to disease progression, and other disease variables. Although these outcomes remain essential, there is general recognition of the need to assess the impact of cancer, and its treatment on patient's health-related quality of life (HRQOL).[1–3] The field of HRQOL has burgeoned in the past two decades. Inclusion of HRQOL measure in medical research is common in the West,[4] but there are few studies from developing countries. The questionnaires measuring QOL have been developed in the western countries, which are culturally different from India.

Head and neck cancer is a major health problem in India as opposed to the West, where it accounts for only 5% of all cancers. Also, head and neck cancers are mainly seen in the low socio economic strata, making it difficult to apply the same quality of life parameters measured in the developed countries with better standards of living. So far, there are no longitudinal studies on QOL reported from India. Although

illiteracy and poverty are the often-cited reasons for our neglect toward measurement of QOL, it is actually the lack of physician awareness, and the nonavailability of valid tools that are responsible for the paucity of QOL studies in India. Moreover, India is a country of 27 official languages, making the translation procedure difficult. Hindi is the national language and is spoken by a large number of patients. Hence, it was advisable to translate the questionnaire into Hindi and Marathi (the local language of the state of Maharashtra). India is a developing country with a literacy rate of only 57%.[5] Most of our patients are socio-economically backward and would accept the treatment decided by the physician. Needless to say that the various treatment options and alternatives are seldom discussed. Most of them are too scared to report emotional, mental, and physical problems. In such a scenario, the onus lies on the physician to study treatment effects and try, and improve patients QOL. A modular approach is necessary to assess the multidimensional aspects of QOL. A general module, which assesses symptoms commonly experienced by cancer patients, is supplemented by a site or treatment-specific module, which assesses difficulties unique to that particular type of cancer. Such an approach is embedded in the questionnaire developed by the European Organization for Research and Treatment of Cancer (EORTC), which is the most widely used questionnaire. The present study was undertaken to use and validate the EORTC QLQ-C30 (core questionnaire) and the QLQ-HN35 (head and neck-specific questionnaire) into two vernacular languages (Hindi and Marathi), and we present the first validation of these questionnaires in head and neck cancer patient population in a tertiary cancer center in India.

Patients and methods

Patients

This prospective longitudinal study was initiated in January 2004. Two hundred patients who consented for the study were recruited. Patients with head and neck carcinoma of all sites and stages were included in the study. All patients were treated with surgery as the primary modality of therapy. Patients with cognitive impairment or inability to understand the language were excluded from the study.

Instrument

The EORTC QLQ-C30 and the QLQ-HN35 were translated from the original English version into two vernacular languages (Hindi and Marathi) using a "forward-backward" translation procedure approved by the EORTC QOL Study Group. The translated questionnaire was pilot-tested on 15 patients before commencing the study. Patients were asked if they found any question difficult to understand, confusing, upsetting, or if they found any word difficult to understand. They were also asked to reframe any question if they found it difficult or upsetting. Based on the suggestions from these interviews, the necessary modifications were made to the questionnaire. Once these modifications were incorporated, the translated questionnaire were ready to be validated.

The questionnaire was served twice. Once preoperatively and again at the 15th postoperative day, giving a total of 400 filled questionnaires. A study co-ordinator, not concerned with patient care, served the questionnaires and the patients were asked to fill and return the questionnaire within 48 hours. The questionnaires were checked and corrected for missing values in the patients, presence. The relevant clinical details including age, gender, level of education, site, and stage of the tumor, and the surgical details were recorded from the hospital case files.

Assessment of validity

Reliability

Internal consistency of the multi item questionnaire was assessed by use of Cronbach's α coefficient,[6] which is used as an indicator of scale reliability. A low alpha value suggests that some items either have very high variability or that the items are not measuring the same thing. It has been recommended that internal consistency should be 0.70 or higher when scales are used for group comparisons.

Construct validity

Construct validity, which was assessed by discriminant and convergent validity, was explored by use of psychometric techniques of analysis of scaling, correlations between items and scales (It-Sc), and correlations between scales (Sc-Sc).

Within scale correlation

It was found by performing Pearson's correlation. Ideally, we looked items within a scale to be moderately highly correlated (> 0.4) with their own scale (corrected for overlap) as a support for convergent validity in items presumed to assess related or similar constructs. Support for item discriminant validity was based on a comparison of the magnitude of the correlation of an item with its own scale as compared with other scales. Scaling successes were defined as those cases in which an item correlated significantly higher with its own scale than with another scale.

Between scale correlation

Correlations between the different scales were explored using Pearson's correlation. A consistently high correlation (>0.70) may indicate that two scales assess the same or high related-constructs.

Criterion validity

Because of the lack of a reference standard, we performed known-groups comparisons at baseline by comparing the mean scores of patients with different clinical characteristics, such as site and stages. A *P*-value suggests significant difference between groups. If the *P*-value given by using Mann–Whitney *U* and Kruskal–Wallis tests was less than 0.05, then it indicated a significant change. For example, we expected that patients with early-stage (I and II) tumors to have a better QOL than in patients with advanced-stage (III and IV) tumors.

Responsiveness

Another important aspect of the clinical validity of a questionnaire is the ability to assess small difference within groups over time. We compared the mean scores pre- and postoperatively to test the responsiveness of the questionnaire. Wilcoxon signed rank test was used to know how many patients showed a positive or negative response for different functional and symptom scale, and whether there was any significant difference between them.

Statistical analysis

The statistical software program SPSS 11.5 for windows was used for the descriptive analysis. The It-Sc and Sc-Sc correlations were found by using Pearson's correlation. Internal consistency in questions was found using Cronbach's alpha coefficient. Tests of differences between groups were performed using nonparametric tests (Mann–Whitney *U* and Kruskal–Wallis tests) because of skewed data. Responsiveness was found using Wilcoxon signed ranks test.

Results

Before commencing the study the translated questionnaire was pilot-tested on 15 patients. The average time to fill the QLQ-C30 was 8 minutes and for QLQ-H and N35 it was 7 min. The patients found it easy to fill and no major changes were necessary.

Two hundred patients completed the questionnaire pre- and postoperatively, giving a total of 400 completed questionnaires. Thirty-three patients filled the English version whereas 167 (83.5%) patients filled the translated vernacular version (104 Hindi, i.e., 52%; 63% Marathi, i.e., 31.5%). Except for a brief explanation of the purpose, content, and layout of the questionnaire, most patients were able to complete the questionnaire without assistance.

There were 148 (74%) males and 52 (26%) females. One hundred and thirty two patients (66%) had completed their education at least till matriculation. The site of primary tumor was the buccal mucosa and the gingivo buccal sulcus (52%), tongue and floor of mouth (16%), larynx and hypopharynx (12.5%), thyroid (12.5%), and maxilla (3%). Miscellaneous tumors including tumors of the salivary gland and neck nodes metastasis of unknown origin (MUO) were grouped together (4%). Fifty-five patients (27.5%) had early-stage tumor (I and II), whereas 117 (58.5%) patients presented in an advanced stages (III and IV). Twenty-eight patients (14%) (thyroid, salivary glands, MUO) could not be staged as per the TNM classification.

Reliability

Internal consistency was measured using the Cronbach's alpha coefficient (Table 1). Most of the scales in both QLQ-C30 and HN35 demonstrated a high alpha coefficient (>0.70) at both the first and second visit.

Cognitive functioning and the nausea vomiting scale (QLQ-C30) demonstrated a low alpha coefficient at

Table 1: Cronbach's alpha coefficient

QLQ-C30	Items	Pretreatment	Post-treatment
Functional scales			
Physical	1–5	0.7186	0.8213
Role	6, 7	0.6897	0.8658
Emotional	21–24	0.8297	0.8716
Cognitive	20, 25	0.3200	0.6341
Social	26, 27	0.7990	0.8025
Global QOL	29, 30	0.9381	0.9431
Symptom scales			
Fatigue	10, 12, 18	0.8037	0.8296
Nausea/vomiting	14, 15	0.5320	0.5553
Pain	9, 19	0.7120	0.7356
QLQ-H&N35 symptom scales			
Pain	31–34	0.6606	0.7612
Swallowing	35–38	0.7906	0.8450
Senses problems	43, 44	0.3651	0.7013
Speech problems	46, 53, 54	0.6615	0.7587
Trouble with social	49–52	0.9121	0.9292
eating			
Trouble with social	48, 55–58	0.8281	0.9011
contact			
Less sexuality	59, 60	0.9304	0.9198

both visits whereas head and neck senses scale (H& N35) demonstrated a low value at the first visit.

Construct validity

All the items in both the questionnaires (QLQ-C30, QLQ-H&N35) had Pearson's correlation (>0.40) with their own scale, which supports item-convergent validity, and none of the items correlated with other scales more than there own scales, which supports discriminant validity. Pearson's correlation between scales was calculated at both visits (Tables 2 and 3). There was no high correlation found between different scales at the first visit (all were <0.70). However, at the second visit the physical functioning and pain scale showed a correlation (>0.70) with the fatigue scale (0.746 and 0.707) in the core questionnaire. A high correlation was not observed in any of the QLQ-H & N35 scales.

Table 2: Correlations between scales in the QLQ-C30

Scale	PF	RF	EF	CF	SF	QL	FA	NV	PA
PF		0.62	0.51	0.60	0.50	0.40	0.75	0.48	0.64
RF	0.46		0.47	0.49	0.45	0.41	0.51	0.36	0.53
EF	0.36	0.34		0.57	0.57	0.48	0.69	0.41	0.60
CF	0.27	0.26	0.50		0.59	0.36	0.59	0.42	0.48
SF	0.25	0.49	0.50	0.35		0.38	0.54	0.31	0.45
QL	0.38	0.38	0.43	0.38	0.45		0.48	0.27	0.46
FA	0.61	0.43	0.53	0.34	0.36	0.46		0.46	0.71
NV	0.20	0.17	0.38	0.32	0.25	0.20	0.41		0.48
PA	0.37	0.39	0.49	0.39	0.36	0.44	0.58	0.28	

Table 3: Correlations between scales in QLQ-H&N35

Scale	PF	RF	EF	CF	SF	QL	FA	NV	PA
Scale	HNPA	HNSW	HNSE	HNSP	HNSO	HNSC	HNSX		
HNPA	-	**0.51**	**0.40**	**0.52**	**0.47**	**0.46**	**0.17**		
HNSW	0.57	-	0.47	0.52	0.68	0.44	0.28		
HNSE	0.34	0.46	-	0.48	0.51	0.58	0.36		
HNSP	0.43	0.53	0.33	-	0.55	0.65	0.41		
HNSO	0.59	0.61	0.36	0.54	-	0.60	0.37		
HNSC	0.42	0.43	0.41	0.49	0.56	-	0.54		
HNSX	0.24	0.23	0.32	0.32	0.21	0.41	-		

Note: In Tables 2 and 3, items to the left of the dashes represent values before the start of the treatment (first visit);
Items to the right of the dashes represent values 15 days after the start of treatment (visit 2).
Abbreviations:
EORTC QLQ-C30: PF, physical function; RF, role function; EF, emotional function; CF, cognitive function; SF, social function; QOL, global quality of life; FA, fatigue; NV, nausea and vomiting function; PA, pain.
EORTC QLQ-H&N35: HNPA, pain; HNSW, swallowing; HNSE, senses problems; HNSP, speech problems;
HNSO, trouble with social eating; HNSC, trouble with social contact; HNSX, less sexuality;

Table 4: QLQ-C30 mean scores between clinical stages at visit 1

	I and II	III and IV	
	Mean	Mean	*P* value
PF	84.2424	85.3561	0.406
RF	87.8788	85.7550	0.408
EF	70.9091	70.6553	0.900
CF	91.8182	83.1909	0.005
SF	76.0606	78.3476	0.962
QL	66.2121	64.2450	0.590
FA	24.4444	26.3058	0.616
NV	2.1212	6.4103	0.048
PA	27.2727	31.0541	0.181
DY	9.0909	13.3903	0.105
SL	18.7879	21.3675	0.418
AP	13.9394	21.0826	0.162
CO	18.1818	11.3960	0.190
DI	3.0303	5.6980	0.147
FI	39.3939	39.6011	0.824

Table 5: QLQ-H&N35 mean scores between clinical stages at visit 1

	I and II	III and IV	
	Mean	Mean	*P*-value
HNPA	23.0303	26.5670	0.000
HNSW	14.3939	15.0285	0.019
HNSE	6.3636	8.1197	0.164
HNSP	17.1717	20.1330	0.001
HNSO	19.5455	27.3504	0.004
HNSC	12.0000	14.9288	0.450
HNSX	21.2121	16.8091	0.166
HNTE	32.7273	33.9031	0.000
HNOM	26.0606	27.0655	0.000
HNDR	11.5152	20.7977	0.063
HNSS	19.3939	25.0712	0.055
HNCO	17.5758	17.9487	0.000
HNFI	20.0000	27.3504	0.128
HNPK	58.1818	57.2650	0.014
HNNU	21.8182	22.2222	0.605
HNFE	5.4545	5.9829	0.148
HNWL	25.4545	40.1709	0.579
HNWG	7.2727	9.4017	0.715

Clinical validity—known-group comparisons

The ability of the questionnaires to detect difference between known groups was assessed at baseline (visit 1). Contrary to our expectations, the QLQ-C30 questionnaire did not reveal statistical difference between early stage and advanced tumors except for cognitive functioning and nausea and vomiting scale

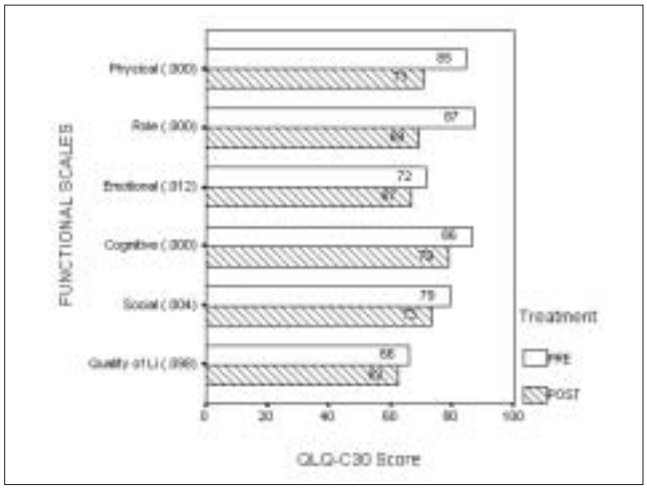

Figure 1: Mean score between functional scales

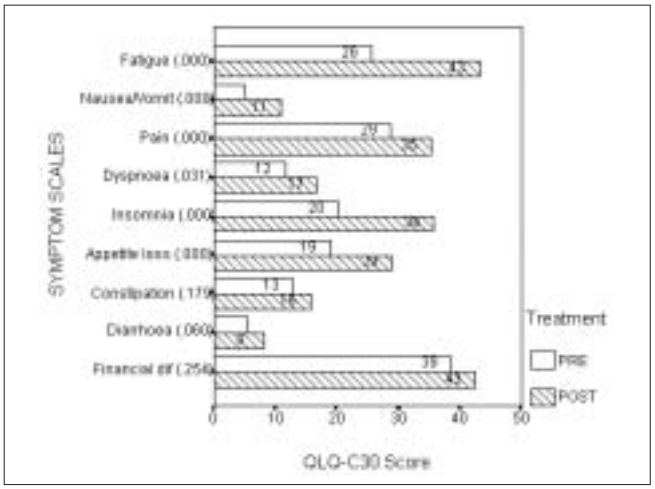

Figure 2: Mean score between symptom scales

(Table 4). The QLQ-H&N35 questionnaire demonstrated a significantly worse score for eight of the 18 symptoms scales in advanced tumors as clinically expected (Table 5).

Patients with oral cancer reported higher problems with the teeth, mouth opening, pain, and social contact as compared with patients with laryngeal and hypopharyngeal cancers who reported higher problems with speech and cough.

Responsiveness

The ability of the questionnaire to detect a clinically significant change over a period of time was assessed using the Wilcoxon signed rank test. In the QLQ-C30 questionnaire there was a statistically significant change in the post-treatment scores as compared with pretreatment scores except for the constipation, diarrhea, financial difficulty, and QL score. (Figures 1 and 2.)

In the QLQ-H&N35 there was a statistically significant deterioration in the post-treatment scores as compared with the pretreatment scores for all the scales, except for problems with teeth and weight gain (Figure 3).

Discussion

Head and neck cancer is a major health problem worldwide. It is a bigger problem in developing countries like India, where it accounts for 30% of all cancer in males, and 13% of cancer in females.[7] At the Tata Memorial Hospital, which is a tertiary cancer center, it accounts for approx 25%[8] of the patient load. The head and neck region includes numerous delicate,

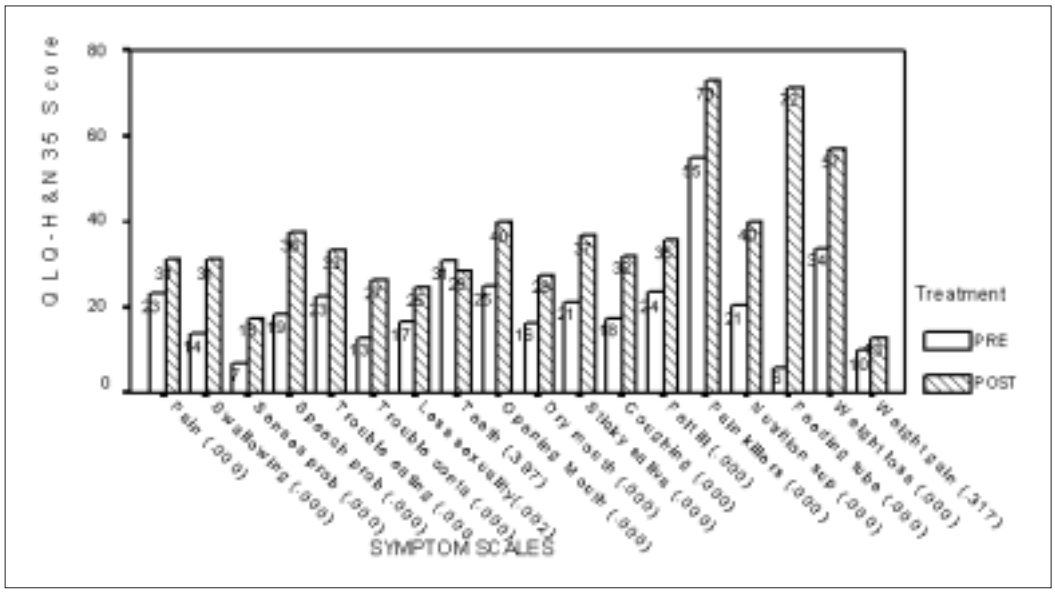

Figure 3: Mean score between symptom scales

intricately organized structures essential for basic physiological functions, and critical for appearance, expression, and social interaction. Depending on the site, size, and patterns of spread, head and neck tumors can cause varying degrees of structural deformations, and functional handicaps compromising well-being, self-esteem, and social integration. In addition, treatment of head and neck tumors can induce additional mutilation, thereby worsening the quality of life. Hence, QOL is an important end-point in evaluating treatment results of head and neck cancers.

In planning a study that proposes a standardized evaluation of the impact on QOL of a treatment, two options are available. The first involves designing an instrument *ex novo*. However, this is not an easy solution, because a new instrument can be used in a clinical study only after a validation procedure, which may take four to six years.[9] The second option requires us to translate one of the existing and validated instruments, the majority of which are in English. This study is the first to translate and validate the EORTC-QLQ-C30 and the QLQ-H&N35 modules for use in the Indian population with head and neck cancer. The advantage of translating a validated tool into a local language instead of developing a new one is primarily the amount of time that is needed to establish the reliability for the translated tool. The present study was performed to test the applicability of a western questionnaire in a different cultural background like ours and also to validate the translated version. Once the questionnaire is validated, it can be used for multicentric studies across the nation. We are presenting the results of a validation study of the EORTC-QLQ-C30 and the QLQ-H & N35 questionnaire in Indian patients with head and neck cancer.

The translations were performed according to the guidelines laid down by the EORTC. Because of familiarity with English language in our country, we did not encounter much difficulty in the translation procedure and we could frame the sentences with the same meaning as the original version. In most instances the translated versions were a mirror image of the original version. The translated versions were pilot-tested on 15 patients. This is an essential step in cross-culture validation as it gauges the acceptability of the questionnaire.

The original and the translated versions were tested for psychometric properties. Both the scales demonstrated acceptable reliability. The internal consistency coefficients of all scales of the EORTC-QLQ-H&N35 were satisfactory (>0.70). In the QLQ-C-30, the cognitive function and the nausea and vomiting scale had a low alpha coefficient at the first visit, which improved at the second visit. Similar findings have been reported on other studies.[10,11] The cognitive function scale has also shown limitations in the first assessment of the EORTC validation study ($\alpha = 0.56$)[12] The low internal consistency coefficient in cognitive function, nausea, and vomiting might have resulted from the diversity of the patient conditions.

In both the questionnaires, the items correlated with their respective scales and did not show a high correlation with any other scale. The small-to-moderate correlation between scales from the QLQ-C30 and the QLQ-H&N35 indicate that the two modules tap relatively different dimensions of QOL, as intended. Within the core questionnaire (QLQ-C30) the physical functioning and the pain scale correlated with the fatigue scale at the second visit. This can be explained by the fact that clinically the items measuring these scales are linked and correlate with each other. The tests for validity based on comparisons between patient subgroups known to differ in clinical status yielded generally consistent results. The QLQ-C30 was less successful in discriminating between patients with different stages of disease. The H&N35 was more sensitive in detecting the difference in symptoms for early- *vs* advanced-stage disease. An essential property of QOL instrument intended for use in clinical trials is that it should be responsive to changes in patients' health status over time. In the current study, statistically significant changes in functional and symptom levels were observed in the expected direction. Postoperative patients had a significantly worse score in most of the scales as compared wirh preoperative patients. We conclude that the EORTC-QLQ-C30 and the QLQ-H&N35 have satisfactory psychometric properties when applied to a sample of Indian population with head and neck cancer. The translated versions have satisfactory levels of reliability and validity. The instrument can now be used for multicentric studies and can also be used in prospective clinical trials.

References

1. OsobaD. The quality of life committee of the clinical trials group of the National Cancer Institute of Canada: organization and functions. Qual Life Res. 1992;1:203–11.
2. Ganz PA, Bernhard J, Hurny C. Quality of life and psychosocial oncology research in Europe: state of the art. J Psychosoc Oncol 1991;9:1–22.
3. Nayfield SG, Hailey BJ, editors. Quality of Life Assessment in Cancer Clinical Trials. Bethesda, MD:1991.
4. Testa, MA, Simonson DC. Assessment of quality-of-life outcomes. N Engl J Med 1996; 334:835–40.
5. Literacy rate of India: Census India 2001, www.censusIndia.net/

(acssess on 03.03.2005).
6. Bland JM, Altman DG, Statistics notes: Cronbach's alpha, BMJ 1997;314:527.
7. Sanghvi LD, Rao DN, Joshi S. Epidemiology of Head and Neck cancers. Semin Surg Oncol 1989;5:305–9.
8. Hospital Cancer Registry, Annual Report 2000, Tata Memorial Hospital, www.tatamemorialcentre.com
9. M. Tamburini. Health-related quality of life measures in cancer. An Oncol 2001;12:7–10.
10. De Boer JB, Sprangers MAG, Aaronson NK, Lange MJA, Van Dam F. The feasibility, validity and reliability of the EORTC QLQ-C30 in assessing the Quality of Life of patients with symptomatic HIV infection or AIDS. Psychol Health 1994;9:65–77.
11. Ringdal GI, Ringdal K. Testing the EORTC Quality of Life questionnaire on cancer patients with heterogeneous diagnoses. Qual Life Res 1993;2:129–41.
12. Aaronson NK, Ahemezdai S, Bergman B, *et al*. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality of life instrument for use in intentional clinical trials. J Natl Cancer Inst 1993;85:365–76.